TNS

Theoretical and Natural Science

Proceedings of CONF-MPCS 2024 Workshop:
Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations

Chicago, USA

October 10th, 2024

EWA Publishing

Volume 52

Editors
Anil Fernando
University of Strathclyde

Marwan Oma
Illinois Institute of Technology

ISSN: 2753-8818
ISSN: 2753-8826 (eBook)

ISBN: 978-1-83558-621-1
ISBN: 978-1-83558-622-8 (eBook)

Publication of record for individual papers is online:
https://www.ewadirect.com/proceedings/tns/home/index

This imprint is published by **EWA Publishing**
Address: John Eccles House, Robert Robinson Avenue, Oxford, England, OX4 4GP
Email: info@ewapublishing.org

Committee Members

CONF-MPCS 2024

General Chair

Anil Fernando, *University of Strathclyde*

Organizing Chair

Marwan Omar, *Illinois Institute of Technology*

Organizing Committee

Sharidya Rahman, *Monash University*
Büşra Oğuzhan, *Çukurova University*
Selda Kapan Ulusoy, *Erciyes University*
Mazhar Javed Awan, *University of Management & Technology*
Arshad Hassan Khan, *FAST NUCES Islamabad*

Technical Program Chair

Stavros Shiaeles, *University of Portsmouth*

Technical Program Committee

Bilyaminu Auwal Romo, *University of East London*
Bismark Singh, *University of Southampton*
Achintya Haldar, *University of Arizona*
Bhupesh Kumar, *University of St Andrews*
Altaf Khan, *University of Mianwali*
Yazeed Ghadi, *Al Ain University*
Jie Zhang, *University of Bath*
Mustafa Istanbullu, *Çukurova University*
Gueltoum Bendiab, *University of Frères Mentouri, Constantine*

Publicity Committee

Festus Adedoyin, *Bournemouth University*
Xiaolong Li, *Beijing University of Posts and Telecommunications*

Preface

The 2nd International Conference on Mathematical Physics and Computational Simulation (CONF-MPCS 2024) is an annual conference focusing on research areas including mathematics, physics, and simulation. It aims to establish a broad and interdisciplinary platform for experts, researchers, and students worldwide to present, exchange, and discuss the latest advance and development in mathematics, physics, and simulation.

This volume contains a selection of high-quality papers submitted to the workshop "Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations," held in collaboration with the CONF-MPCS 2024. The workshop, chaired by Dr. Marwan Omar from Illinois Institute of Technology, is part of a broader initiative to examine interdisciplinary approaches in applied mathematics, applied physics, algorithm, simulation, machine learning, and quantum computing. Each of these papers has gained a comprehensive review by the editorial team and professional reviewers. Each paper has been examined and evaluated for its theme, structure, method, content, language, and format.

Cooperating with prestigious universities, CONF-MPCS 2024 organized two more workshops in Glasgow and Constantine. Dr. Anil Fernando chaired the workshop "Unlocking Video Contextual Ad Insights: Enhancing Topic Explainability with Rich Multimodal Content Retrieval", which was held at University of Strathclyde. Dr. Gueltoum Bendiab chaired the workshop "Machine Learning: Integrating Machine Learning Techniques to Advance Network Security", which was held at University of Frères Mentouri.

Besides these workshops, CONF-MPCS 2024 also held an online session. Eminent professors from top universities worldwide were invited to deliver keynote speeches in this online session, such as Dr. Anil Fernando from University of Strathclyde and Dr. Marwan Omar from Illinois Institute of Technology. They have given keynote speeches on related topics of mathematics, physics, and simulation.

On behalf of the committee, we would like to give sincere gratitude to all authors and speakers who have made their contributions to CONF-MPCS 2024, editors and reviewers who have guaranteed the quality of papers with their expertise, and the committee members who have devoted themselves to the success of CONF-MPCS 2024.

Dr. Anil Fernando

General Chair of Conference Committee

Workshops

Workshop – Glasgow: Unlocking Video Contextual Ad Insights: Enhancing Topic Explainability with Rich Multimodal Content Retrieval

August 9th, 2024 (GMT+1)

Department of Computer and Information Sciences, University of Strathclyde

Workshop Chair: Prof. Anil Fernando, Professor in University of Strathclyde

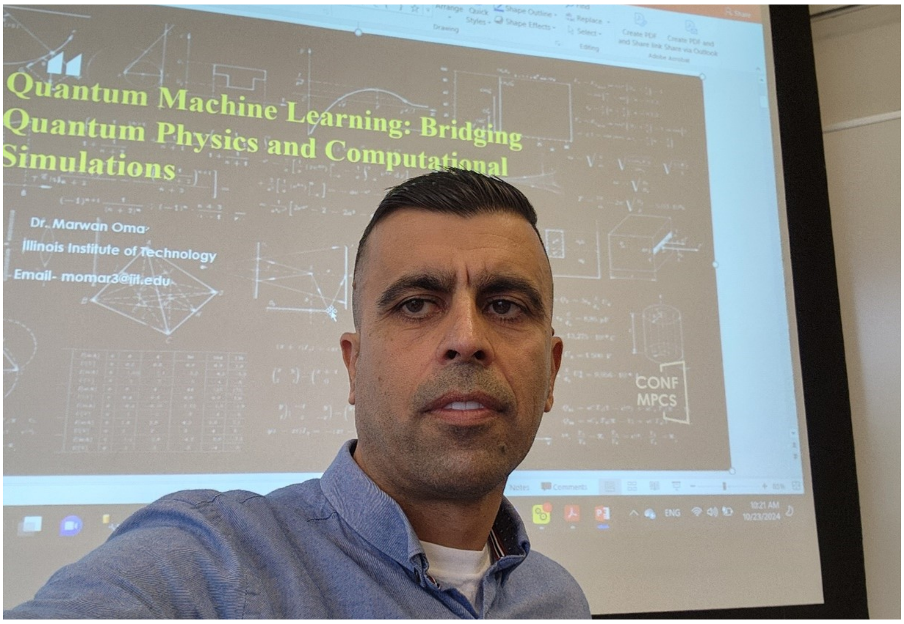
Workshop – Constantine: Machine Learning: Integrating Machine Learning Techniques to Advance Network Security

July 15th, 2024 (GMT+1)

Electronics Department, University of Frères Mentouri

Workshop Chair: Dr. Gueltoum Bendiab, Associate professor in University of Frères Mentouri

Workshop – Chicago: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations



October 10th, 2024 (UTC -5)

ITM Department, Illinois Institute of Technology

Workshop Chair: Dr. Marwan Omar, Associate Professor in Illinois Institute of Technology

CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations

CONF-MPCS 2024

Table of Contents

Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations
DOI: 10.54254/2753-8818/52/2024CH0112

Near-space radiation environment analysis and protective measures

Xuan Qi

School of Physics, Stony Brook University, Stony Brook USA

xuan.qi.1@stonybrook.edu

Abstract. As a matter of fact, near space radiation environment analysis plays a crucial role in near space physics. With this in mind, this study introduces the types of radiation damage in the near-space radiation environment as well as corresponding protective measures. To be specific, the protective strategy for Total Ionizing Dose (TID) and for Displacement Damage (DD) are similar, focusing on material shielding to mitigate the effects. According to the analysis, the guiding principle for the protection strategy against Single Event Effects (SEE) is risk management, aiming to minimize the probability of catastrophic SEE and to detect and mitigate the impact of non-destructive SEE. Based on the analysis, some measures are proposed accordingly. In fact, current strategies for future radiation protection may involve the use of biological membranes to absorb radiation or the application of quantum mechanics principles to eliminate the effects brought about by radiation. These results shed light on guiding further exploration if near space radiation.

Keywords: Single event effects, total ionizing dose effects, displacement damage effects, radiation protection.

1. Introduction

The impact of the near-space radiation environment on electronic devices is becoming increasingly evident with the improvement of device integration, and the resulting losses are becoming more severe [1]. Due to the many advantages of observation missions in near-space compared to satellites, reducing the impact of radiation on aircraft and spacecraft and developing corresponding protective measures have become a research focus for major aerospace powers [2, 3]. The radiation effects in near-space include: Total Ionizing Dose (TID) effects, Displacement Damage (DD) effects, and Single Event Effects (SEE) [4, 5]. These effects differ in the way they affect devices and the extent of their impact, and the corresponding protective measures also vary [4, 5],

In recent years, major aerospace powers have gained a thorough understanding of these effects and have developed relatively mature protective measures. This paper will introduce some basic concepts related to near-space radiation, review the current research on the near-space radiation environment, and summarize the protective measures against the radiation environment.

2. Related concepts

When studying radiation effects in aircraft and spacecraft, the focus is on the energy changes, i.e., the energy difference ΔE, when charged particles interact with matter. To accurately describe this energy

DOI: 10.54254/2753-8818/52/2024CH0112

transfer, one introduces the physical concept of Linear Energy Transfer (LET), defined as $LET = \frac{dE}{dX}$. This formula describes the rate of change of energy as the incident particle travels through the target material. When charged particles enter the target material, they gradually slow down and, if the volume of the target material is large enough, the particles will eventually come to a complete stop. The total distance that particles travel in the target material along the direction of incidence is called the range (R). To compare the effects of different mass densities of materials on the LET and range of charged particles, the concept of mass thickness is commonly used, which is calculated by multiplying the mass density of the material with its thickness or length. LET and range are important indicators for measuring the intensity of interaction between charged particles and matter, especially LET, which plays a central role in this field.

When charged particle A interacts with target material B, it transfers energy through three main mechanisms: ionization, displacement, and bremsstrahlung radiation. For particles with small mass such as electrons, bremsstrahlung radiation also occurs. For ionization, this is the process where charged particles collide with the outer electrons of target atoms, causing the electrons to be knocked away from the atoms. The linear energy transfer of this collision is denoted as LET_e. Displacement is the process where charged particles collide with the nuclei of target atoms, causing the nuclei to move. The linear energy transfer of this nuclear collision is denoted as LET_n. As for bremsstrahlung radiation, when the incident particles interact with target atoms, electromagnetic radiation is produced due to accelerated motion. The linear energy transfer of this radiation is denoted as LET_r. Considering these three mechanisms, the total linear energy transfer LET_{tot} of charged particle A in target material B is composed of the sum of the three parts of energy transfer, expressed as: $LET_{tot} = LET_e + LET_n + LET_r$

When charged particles travel through an infinitely large target material, they transfer and dissipate energy through a series of collision processes, which results in a finite range. During these collision processes, not only does the speed of the incident particles gradually decrease, but their trajectories also change due to the forces they experience, a phenomenon known as scattering. It is worth noting that heavier particles typically deflect less during scattering than lighter particles.

Ionization is the primary mode of interaction between charged particles and matter. It involves inelastic collisions between incoming incident particles (with energy E_a) and the outer electrons of the target atoms (with binding energy I_b, on the order of 10eV) through electrostatic Coulomb interaction. When $E_a > I_b$, the outer electrons of the target atom gain enough energy to break free and become free electrons with a certain energy E_{fe}, leaving the target atom as a positive ion, while also causing energy loss and deceleration of the incident particle. If E_{fe} is still greater than I_b, it can continue to cause secondary ionization. Displacement occurs when an incoming particle (with mass M_a and energy E_a) undergoes a dynamic elastic collision with a stationary target atomic nucleus (with mass M_b), causing the target atom to be ejected along the direction of the incident particle at an angle θ with energy:

$$E_b = \frac{4M_aM_b}{(M_a+M_b)^2}cos^2\theta E_a \tag{1}$$

thereby moving away from its original position and creating vacancies and interstitial atoms in the atomic lattice, introducing defects in the irradiated material. If E_b is large enough, it can continue to collide with other target atoms, causing secondary displacements.

When high-energy charged particles (with atomic number Z_a, mass number A_a, and energy E_a) strike the spacecraft material (with atomic number Z_b), they are decelerated by obstruction, and with the change in velocity, they radiate photons (with a continuous energy spectrum up to E_a), which is known as bremsstrahlung radiation. The power of bremsstrahlung radiation $\propto (Z_aZ_b)^2/A_a^2$, so for actual charged particles, electrons, due to having the same charge number as protons but a mass of only about $1/1837$ of that of protons, have a bremsstrahlung radiation power that is 10^6 times that of protons under the same conditions, which is particularly prominent. Therefore, for space radiation particles mainly composed of electrons, protons, and heavy ions, the bremsstrahlung radiation effect of electrons is primarily considered [3]. In the space radiation environment, when electrons pass through materials such as satellite skins and electronic device casings, they are decelerated by obstruction, producing X-rays. These secondary X-rays have strong penetrating power and continue to penetrate electronic device casings or device encapsulations, producing ionizing radiation doses at critical devices, leading to Total Ionizing Dose effects [2, 3].

3. Types of radiation damage in near-space

The types of radiation in near-space are mainly divided into three types, i.e., Total Ionizing Dose (TID) effects, Displacement Damage (DD) effects, and Single Event Effects (SEE) [2, 3]. The Total Ionizing Dose effect (also known as TID effect) is the impact caused by the ionizing action of space radiation particles, which is positively correlated with the cumulative dose. These radiations produce additional electron-hole pairs through the ionization process in the material of the device, which move, evolve, and are captured within the device, ultimately affecting the performance of the device. For spacecraft, the TID effect on the operational devices is the most prominent, with the cumulative ionizing charge causing the device's operating parameters to gradually degrade and eventually fail. On April 5, 2010, the "Galaxy 15" satellite of the World Teleport Communication Organization suffered a failure. Through verification, this failure was caused by deep charge discharge leading to device damage and satellite failure [1].

As shown in the Fig. 1, the Displacement Damage effect is caused by the elastic collision between space radiation particle A and the atom B of the target material, causing the target material atom to move as shown in the figure below, including the target atom leaving its original position to become an interstitial atom and leaving a vacancy at the original position of the target atom. The degree of displacement of the target material atom is related to the elastic collision energy received by the target material per unit mass of radiation particles, denoted as Displacement Damage Dose (DDD), which is also called Displacement Damage effect. This effect can be caused by the bombardment of a single particle.

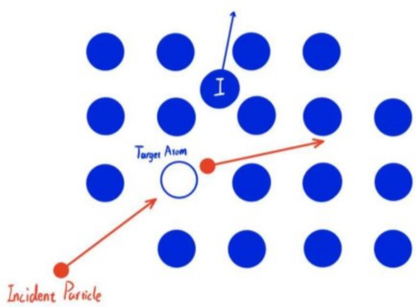


Figure 1. Displacement Damage Effects (Photo/Picture credit: Original).

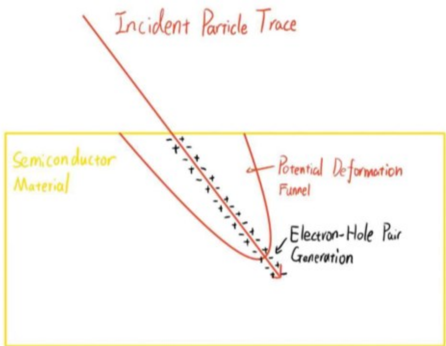


Figure 2. Single Event Effects (Photo/Picture credit: Original).

DOI: 10.54254/2753-8818/52/2024CH0112

There are various types of Single Event Effects, and their most fundamental process involves three stages. A single charged particle travels through the device material, forming an instantaneous ionization action that creates an electron-hole pair column along its trajectory. The sensitive nodes in the device collect the ionizing charge to form a transient current pulse. The transient current pulse impacts the specific functional parts of the device circuit, inducing various types of Single Event Effects. A sketch is shown in Fig. 2. The main Single Event Effects include:

- Single-Event Upset (SEU). A logic error in the circuit caused by a potential error due to a single particle strike.
- Single-Event Latch-up (SEL). When a high-energy particle strikes a CMOS device, it may trigger a state known as latch-up, where some parts of the device consume abnormally high current, causing device damage or performance degradation.
- Single Event Transient (SET). In logic circuits, a single particle strike may cause a brief change in current or voltage, and this transient effect may affect the normal operation of the circuit.
- Single Event Functional Interrupt (SEFI). A high-energy particle strike may temporarily interrupt some functions of the system or device.
- Single Event Burnout (SEB). Due to a large current caused by a single particle, in power MOSFETs and other devices, a single high-energy particle may cause permanent damage to the device.
- Single Event Gate Rupture (SEGR). Similar to SEB, it is a device failure caused by a large current. This effect causes gate damage.

Among them, SEL, SEB, and SEGR are destructive Single Event Effects, which are highly destructive once damage occurs and are currently difficult to remotely repair. SEU, SET, and SEFI are non-destructive Single Event Effects, which can generally be restored to some extent through various software and hardware means. A more serious loss occurred in 2011, when the "Phobos-Grunt" Mars probe launched by Russia, after reaching low orbit and only about 3 hours after launch, suffered a single particle strike on a key device causing a Single-Event Latch-up, preventing the execution of the expected ignition command, and ultimately the satellite fell into the Earth's atmosphere [6].

4. Total Ionizing Dose (TID) Effect Protection Design

In specific circuit design, tolerance design can be used to ensure that the circuit system can operate normally within a wider range of electrical parameters, or even when there is a certain degree of deviation. However, in general, when conducting TID irradiation experiments, tolerance has often been considered already, so there is not much room for tolerance design. Redundant backup is a frequent practice to improve system reliability. For redundancy design actively carried out due to TID reasons, or redundancy design that improves the TID level by taking advantage of basic reliability backups, it is necessary to determine through experiments and research which redundancy mode (cold backup, hot backup, warm backup, etc.) has the least TID impact or is safe. Based on the overall design of the spacecraft, the impact of TID effects can be reduced, for example, by arranging more critical devices in the center of the spacecraft.

Individual equipment can assess the radiation dose through a simple method. First, consider the thickness of the equipment case and the thickness of the satellite shell (usually 0.8 mm aluminum material), and then use the dose-thickness relationship curve obtained from experimental data to estimate the radiation dose that the equipment may withstand. Through this assessment, it can be determined whether the TID tolerance of the electronic devices used meets the required standards. If the assessment results show that the tolerance of the devices is insufficient, the thickness of the case can be increased within the weight limit of the equipment, or shielding materials can be locally added to key electronic devices to improve their radiation resistance and meet the requirements for radiation protection. For key devices, the shielding contribution of the internal PCB board of the case can also be further considered, as shown in the Fig. 3. For the device shown in the figure above, in addition to the shielding of the case thickness d_{case} on all six sides, considering that the area of the internal PCB board is large and the shielding blockage solid angle for key devices is large, ignoring the lack of shielding in the direction parallel to the PCB board, it is artificially assumed that the one-dimensional shielding thickness received by the key device from the case and the PCB board (d_i) is $d_{case} + \frac{1}{2}(d_1 + d_2 + d_3 + \cdots + d_n)$. The radiation dose under this shielding thickness can be evaluated according to the dose-thickness curve to see if it meets the requirements.

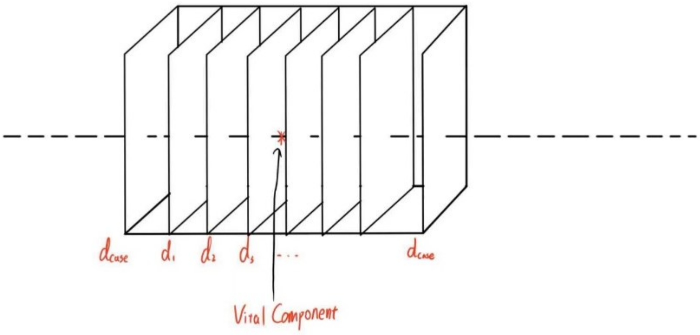


Figure 3. The One-Dimensional Shielding Design (Photo/Picture credit: Original).

The above solid sphere one-dimensional shielding design can quickly give a safe shielding design for satellites without mass cost burden, mainly considering the shielding of the satellite shell and the individual equipment case that the spacecraft internal single machine designer is concerned about, and ignoring the shielding effect of other instruments, parts, structures, etc., inside the spacecraft on the single machine equipment of interest, which is a conservative design. For satellites with high TID index and limited mass resources (such as MEO satellites), it is necessary to comprehensively analyze the shielding effects of those details ignored by the one-dimensional model, and carry out additional shielding design on this basis. Therefore, it is necessary to perform three-dimensional shielding analysis and design for satellites with complex material composition and structural layout. For the calculation of radiation dose, two sets of experimental data are generally required: the shielding thickness statistical distribution data $\mathrm{F}(d_i)$ and the one-dimensional dose-thickness curve data $\mathrm{D}(d_i)$ of the spacecraft mission. The unit of thickness uses equivalent thickness of aluminum. The following are examples of the two sets of curves in Fig. 4 [7]. Using the following formula, the more realistic radiation dose after the complex spacecraft structure layout can be calculated.

$$Dose = \frac{1}{4\pi}\sum_1^n D(d_i) \cdot F(d_i) \tag{2}$$

For the model of the spacecraft, modeling software such as ProE or COSMIC can be used for modeling, and the protection plan can be analyzed and specifically designed for the dose.

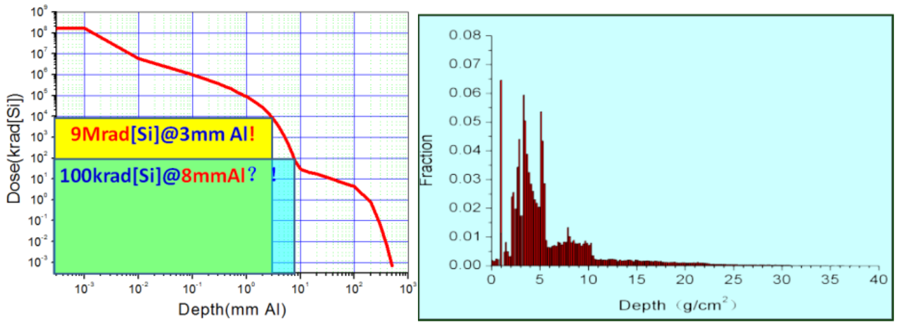


Figure 4. One-dimensional dose-depth curve $\mathrm{D}(\boldsymbol{d_i})$ of a certain MEO satellite (left) and shielding thickness statistical distribution chart $\mathrm{F}(\boldsymbol{d_i})$ [7].

In the aforementioned chapters, for the transmission and shielding calculation of space charged particles in the spacecraft material, the spacecraft material is simplified to aluminum material, and the common electronic equipment case also uses aluminum alloy. For the local shielding of key devices, in order to save space, materials with larger mass density and atomic number, such as lead or tantalum, are used. High-density materials can better absorb high-energy photons (such as X-rays) compared to low-density materials. Therefore, after the low-density material decelerates the electrons, the high-density material can absorb the bremsstrahlung radiation (X-rays) caused by the deceleration of electrons to reduce radiation damage. For materials with a medium thickness (about 3-10mm), composite materials made of aluminum and tantalum in a certain ratio (20%-30% aluminum content of tantalum-aluminum composite material) can better shield radiation. For the protection of key devices, considering the space saving and the good protection performance of pure tantalum under a thinner thickness (less than 3mm), pure tantalum is generally used for protection. For large thickness parts, due to the payload requirements of aerospace, pure aluminum is generally used [3, 8].

A research result in August 2018 mentioned the current protection measures for TID effects [4]. According to the calculation results of the radiation protection software, for a specific optical system structure model, a protective cover can be added in front of the lens to enhance protection. This protective cover design is simple and can be flat or concave, and should have the same engineering installation characteristics as the light-blocking cylinder components in front of the lens in the model. Following the principle of minimizing the mass of the satellite, it is recommended that the thickness of the protective cover be between 1 and 3 millimeters, and to achieve the best protective effect, the protective cover should be located no more than 20 millimeters in front of the lens [4]. The quantitative data calculated by the software shows that after the shielding of the protective cover with a thickness of 1 to 3 millimeters, the contribution of the radiation dose to the most severe environmental parts through the exposed window of the Earth's radiation belt can be reduced to one to ten percent of the original radiation dose, that is, the radiation dose is reduced by two orders of magnitude [4]. If the on and off time ratio of the remote sensor in orbit has been further considered, the radiation dose at the main mirror can be further reduced. For extremely high-energy galactic cosmic ray particles, the dose change before and after protection is not significant. However, due to the very small dose contribution of such particles, they are usually negligible. Considering the analysis above, the protection scheme using a protective cover can significantly reduce the radiation absorption dose of the remote sensor in the most severe environment, bringing the radiation absorption dose of this part to the same order of magnitude as other parts of the optical system, thereby effectively enhancing the radiation protection performance of the system [4].

5. Displacement damage protection design

The Displacement Damage effect is similar to the Total Ionizing Dose effect, both being caused by the cumulative effect of low to medium energy but high flux radiation belt particles and solar protons leading to the degradation of electrical parameters. Therefore, the protective design concepts for the Total Ionizing Dose effect are also applicable to the protection design for Displacement Damage effects. However, when designing proton shielding, it is necessary to consider that the lower the proton energy, the greater the displacement damage effect it causes. Therefore, a careful analysis and calculation of the impact of lower energy protons after material shielding is required.

6. Single event effect protection design

The types of Single Event Effects (SEEs) that occur in devices and the impacts they have on circuits and equipment vary, and so do the corresponding protective design strategies and methods. The following table is a ranking of the Single Event Effects that space missions need to focus on, considering both the frequency of occurrence in orbit and the resulting hazards. The ranking is in order of concern, starting with CMOS device Single Event Latch-up (SEL), digital device Single Event Functional Interrupt (SEFI) and Single Event Upset (SEU), power device Single Event Burnout (SEB), and analog device Single Event Transient (SET) (seen from Table 1). The guiding principle of Single Event Effect (SEE)

DOI: 10.54254/2753-8818/52/2024CH0112

protection design is risk management. The following are the protective design approaches and main methods for destructive and non-destructive Single Event Effects, respectively. It is necessary to avoid and reduce the probability of destructive Single Event Effects (SEEs); uses devices with a high Linear Energy Transfer (LET) threshold; derate the operating voltage of power devices (operate at less than 50% of the safe working voltage for Single Event Burnout (SEB) and Single Event Gate Rupture (SEGR)). SEL current limiting and recovery: Utilize smart circuits that autonomously detect and determine the occurrence of Single Event Latch-up (SEL), actively cut power, wait for an appropriate duration, and then reapply power; or through ground personnel, manually determine the occurrence of SEL, and manually send commands to cut and reapply power; also consider periodic automatic power cycling to eliminate potential latch-ups [9]. One also needs to implement backup designs for critical components and circuits. Besides, it is necessary to detect and mitigate the effects of non-destructive Single Event Effects (SEEs). For SEU-induced data errors, redundancy voting in hardware and software, Error-Detecting and Correcting (EDAC) codes, refreshing, reconfiguration. SEFI-induced system malfunction (detection and recovery); error counters, watchdog timers; built-in or external health diagnostic management units. Partial or full refreshing of components; system reset, power cycling SET (Single Event Transient); logic units: temporal redundancy; power circuits: Filtering, voltage regulation; restart to address potential latch-ups; linear or interface circuits: filtering, and possible SEU protection in subsequent connected digital circuits.

Table 1. Single Event Effects of Concern for Spacecraft Missions

Priority	Single Event Type	Severity (from 1 to 5)	Frequency (from 1 to 5)	Note
1	CMOS SEL	Severe 4	Frequent 4	Hard to suppress its occurrence; it can only be recovered by power cycling.
2	SEFI, SET	Medium 2	Very Frequent 5	Error correction and recovery can be performed.
3	SEB	Very Severe 5	Seldom 1	A hard error that cannot be fixed, limiting current is the only option.
4	SET	Light 1	Very Frequent 5	In a few cases, the fluctuations can pose significant hazards.

In 2021, a research team mentioned that they conducted a series of laser pulse irradiation experiments on the CMOS process SRAM model CY62167DV30LL, with the aim of exploring its Single Event Latch-up (SEL) characteristics, and quantitatively calculated the holding voltage and holding current of the SRAM in the latched state through these experiments [10]. Based on these electrical characteristics and latching conditions, the team proposed two circuit-level protection methods: voltage division resistors and power supply current limiting. Both methods are easy to implement and do not affect the normal function of the device, and have high engineering practicality. For this SRAM, the range of holding voltage V_{hold} is 1.5 to 1.6 volts, and the range of holding current I_{hold} is 9.9 to 11.2 milliamperes. For this chip, the power supply current limiting method successfully brought the device out of the Single Event Latch-up state. To achieve this, the limiting current value should be greater than the device's standard operating current of 2 milliamperes and less than the holding current of 10.3 milliamperes. Similarly, by using the method of connecting a voltage division resistor in the circuit, the device was also successfully brought out of the Single Event Latch-up state. Under standard working conditions, the minimum value of the resistor should be between 60 and 70 ohms, and the maximum value should not exceed 500 ohms. Although this circuit-level protection strategy cannot completely prevent the occurrence of Single Event Effects (SEEs) and SRAM data errors caused by large current surges, it can quickly bring the device out of the latched state. On the one hand, this can prevent the

DOI: 10.54254/2753-8818/52/2024CH0112

device from being damaged due to continuous operation in a high current state; on the other hand, it can significantly reduce the number of data anomalies and limit errors to a smaller area centered on sensitive points. Therefore, by using an interleaved storage of data, combined with traditional error detection and correction methods, the correctness of the data can be ensured [10].

7. Limitations and prospects

All current protective measures aim to mitigate radiation damage, and it is not possible to completely eliminate damage within the payload allowance at a low cost. Possible future research directions include following. For biomembrane, melanin produced by some fungi can absorb radiation very well or convert high-energy photons into low-energy photons. This may help reduce the impact of Total Ionizing Dose (TID) effects. For quantum mechanical principle devices, the devices that do not rely on electrons as charge carriers and information transmission particles (such as the polarized photons used in quantum encryption) may help reduce the impact of electron-hole pairs generated by high-energy particles on the device.

8. Conclusion

To sum up, this study introduces the types of radiation damage in the near-space radiation environment and corresponding protective measures. The types of radiation damage include Total Ionizing Dose (TID) effects, Displacement Damage (DD) effects, and Single Event Effects (SEE), with damage sources from ionization, displacement, and bremsstrahlung radiation. The protection strategy for TID involves calculating the shielding effect contributed by the materials of the spacecraft to determine the absorbed dose of the devices. For the protection against Displacement Damage, the strategy is similar, focusing on material shielding to mitigate the effects. The guiding principle for the protection strategy against SEE is risk management, aiming to minimize the probability of catastrophic SEE and to detect and mitigate the impact of non-destructive SEE. Current strategies for future radiation protection may involve the use of biomembranes to absorb radiation or the application of quantum mechanics principles to eliminate the effects brought about by radiation.

References

[1] Progress has been made in the experimental study of the failure of the YH-15 communication satellite at the Space Center. Progress made in the experimental study on the failure of the YH-15 communication satellite at the Space Center, CAS 2010 Retrieved from: https://www.cas.cn/ky/kyjz/201005/t20100513_2844669.shtml
[2] Wang N, Ma L L, Liu Q, et al. 2023 The near-space altitude experiment for satellite radiometric calibration and the first results. National Remote Sensing Bulletin 27(5) 1177-1193.
[3] Zhang Y 2018 Research of X-ray/Electrons Induced Single Event Soft Errors in 45nm SRAM. Xiangtan University.
[4] Zhang X, Zhao X and Li X 2018 Protection method for total dose effect on the optical system of typical satellite remote sensor and related design. Key Laboratory for Optical Remote Sensor Technology of CAST, Beijing Institute of Space Mechanics & Electricity 11.
[5] Wang C 1997 The Influence with Reliability of Motional Satellite by the Single-Event Phenomena. Shijiazhuang. China Electronics Technology Group Corporation thirteenth Research Institute.
[6] Han J, Feng G, Yu Y et al. 2014 Analysis of Single-event Effects Rate of K6R4016V1D Chips Applied in Low Earth Orbit. Chinese Journal of Space Science 35(1) 64-68.
[7] Zou S, Xue B, Liu Z, Zhou X and Sheng L 2017 Design of shielding for GEO satellite and analysis for its effectiveness. Spacecraft Environmental Engineering 4 410-414.
[8] Hu Z, Li Z, Zhang T and Zhang X 2003 New Development and Applications of Tantalum and Tantalum Alloys. Rare metals and hard alloys 3 34-36+48
[9] Wu H, Zhu X, Han J 2021 Research on Circuit Level Protection Design of SRAM Single Event Latch-up Effect. Atomic Energy Science and Technology 4 758-766.

[10] Hu X, Wei F, Li L et al. 2022 In-situ environmental detecting payloads onboard the floating platform in near-space. Aerospace Technology 3 95-104.

DOI: 10.54254/2753-8818/52/2024CH0106

An empirical research on the influence of house price on the consumer market based on monthly statistics in Shanghai

Xinyan Liu[1], Kejing Lyu[2,4], Zhuoyuan Wang[3]

[1]School of Mathematical Sciences, Arizona State University, Phoenix, 85281, United States of America
[2]School of Sociology, Southeast University, Nanjing, 211189, China
[3]School of Jin Cheng No.1 Middle School, Shanxi, 048000, China

[4]213234132@seu.edu.cn

Abstract. Consumption plays a crucial role in people's daily lives. In recent years, China's consumer market has faced urgent needs for consumption transformation and heavy pressure from economic downturn. As an important component of everyday consumption and wealth savings, there is a close relationship between changes in house prices and the consumer market. There has been a heated debate in the academic community regarding the positive or negative relationship between the two. The study keeps focus on the changes of consumer market, attempting to reveal the effects of house price has. This article extracts monthly consumption and housing price data in Shanghai from 2014 to 2024, strictly arranges them in chronological order, and uses the ECM model to handle the instability of time series data. The conclusion of this study is that the rise in housing prices can stimulate consumption. Also house prices have a controlling effect on the fluctuations of the consumer market.

Keywords: House price, consumption market, fixed assert.

1. Introduction

On the macro level, the variety, abundance, and novelty of consumption not only contribute to economic growth and market vitality, but also play an important role in the development of social structure, cultural exchange, and personal happiness. As an important part of consumption, real estate plays a vital role in people's lives and wealth. Changes in house prices also affect the mortgage amount that property owners can get. For those who do not have a house, rising house prices not only increase their alternative consumption costs, but also force them to choose between daily consumption and house purchase investment. They have to store some assets for future purchase, thus reducing their consumption level. In addition, although the correlation between house prices and consumption is obvious, it does not mean that there is a direct relationship between the two. House prices and consumption may be affected by a variety of factors such as interest expenses, income expectations, credit conditions and unemployment.

When it comes to the influence of house price on consumption, scholars tend to support two kinds of theory, one supposing the effect is positive and the other standing for negative impacts. Meanwhile, plenty of studies focus on the way house price affects the consumption behaviors. Wealth effect and Crowding-out effect are the widest-known theories to explain the mechanism.

DOI: 10.54254/2753-8818/52/2024CH0106

A group of researchers deem that house prices have a positive impact. Lin et al. found that for cities with different degrees of development, the impact of housing prices on consumption will be reflected in both positive and negative wealth effects [1]. Cong and Song believed that the rise of housing prices can promote the consumption demand of urban residents, and there exists a significant partial mediating effect of human capital between housing prices and urban consumption. Overall, human capital and housing prices have a positive promoting effect on urban consumption [2]. Zhang and Chong believed that the rise in housing prices has driven non-residential consumption among residents, with little impact on essential consumption. Also, it will increase unnecessary consumption through the "wealth effect" [3]. To conclude, it is believed by plenty of researchers that house price pushes the consumption.

However, the assumption that house price affects consumption positively is faced with a large number of doubts. Hong proposed that there is a negative correlation between overall fluctuations in real estate prices and consumption growth [4]. Tong proposed that the increase in household income has a promoting effect on household consumption, while changes in housing prices are negatively correlated with household consumption, and regional differentiation is severe [5]. Sun et al. found that housing prices inhibit the upgrading of residents' consumption structure, which is different in different consumption structure upgrade stages, in different levels of cities and regions [6]. Jin and Jiang believed that the rise in housing prices has a wealth effect and a substitution effect, because the rise in housing prices makes residents spend more on buying houses, which leads to a decline in daily consumption levels [7]. These scholars all define the relationship between house price and consumption as negative.

There also exist a few studies that observed the relationship in different aspects. Liu and Chen believed that the rise in housing prices has significantly increased the inequality of residents' consumption [8]. Jin and Chu believed that the fluctuations in housing prices and stock prices will precede economic fluctuations and consumption fluctuations, which shows that real estate and stocks are a kind of w in housing prices and stock prices have little impact on residents' consumption, and there is a two-way causal relationship between residents' consumption and housing price fluctuations and economic fluctuations [9]. Xu and Wu found through research that the impact of GDP on housing prices shows a nonlinear improvement effect as the population increases, and the more serious the aging of the region, the greater the possibility of housing prices falling [10].

In conclusion, most scholars agree that house prices play a significant role in the change of consumption, no matter how the effect is finished. However, a large number of studies work on the relationship between consumption behaviors and house price, ignoring the reaction of the consumer market. As a result, the study focuses on the changes of consumer market, attempting to define the way house price affects the market.

2. Methodology

2.1. Data source

The data in this study comes from the statistical data from the National Bureau of Statistics, Shanghai Municipal Bureau of Statistics, and Shanghai Municipal Bureau of Finance. Monthly data related to housing prices and consumer markets in Shanghai from 2014 to 2024 are collected, including the monthly average selling price of commercial housing, housing price index, total social retail goods, and rate of change. At the same time, the change index of fixed assets investment and the change rate of general public budget expenditure are taken into consideration. All the data mentioned above is calculated on a monthly basis. Average house prices are listed in the unit of Yuan, while the general public budget expenditures are presented in the billions of Yuan. All the data about growth rate are accumulated year-on-year.

2.2. Indicators selection and explanation

To measure the changes in the consumer market in Shanghai, this study selected the growth rate of total retail sales of social goods and the consumer index of residents as the dependent variables. The total retail sales of consumer goods refer to the amount of physical goods sold directly by enterprises to individuals or social groups through transactions for non production or non business purposes, as well as the income obtained from providing catering services. The goods included in the statistics do not include raw materials used for production investment, nor do they include expenses incurred by residents for purchasing commercial housing, and expenses incurred by farmers for purchasing agricultural production materials. The fluctuation of housing prices is displayed through the average housing price and monthly housing price index. To control dimensionality, the housing price data was logarithmically processed. The price index of second-hand and newly built houses will be used to reflect the impact of different types of housing prices on the consumer market. The influence of policy expenditure and corporate investment on the consumer market cannot be ignored. The change rate of general public budget expenditure and fixed assets investment will reflect the impact of both. After defining the names and indication of variables, a descriptive analysis is necessary to give a concise summary of the statistics. The basic qualities of the information are displayed in Table 1.

Table 1. Descriptive Analysis

Variable	Min	Max	Mean	Std. deviation
Growth Rate of Total Retails Sales of Consumer Good (GSC)	-48.300	113.500	7.408	17.954
Consumer Price Index (CPI)	-0.200	3.690	1.958	0.904
Growth Rate of House Price (GHP)	-37.530	106.970	11.588	21.088
House Price (HP)	9.593	10.806	10.266	0.325
Growth Rate of Fixed Assets Investment (GFI)	-21.200	40.500	7.329	9.408
Growth Rate of General Public Budget Expenditure (GPE)	-12.000	66.200	9.549	13.626
Growth Rate of Second-hand House Price (PR-S)	-7.500	37.400	5.658	9.965
Growth Rate of Newly-built House Price (PR-N)	-5.900	39.500	6.797	10.211

2.3. *Method introduction*

In order to take on further study on the influencing effect of house price changes on the consumer market, a correlation test should be finished at first to figure out whether the indicators are selected properly. So the study uses the method of Pearson correlation. The method is aimed to find out the linear correlation between these two variables. The range of the Pearson correlation coefficient is $[-1,1]$. If no linear correlation between the two variables, the value of the Pearson correlation coefficient is 0. The result of the correlation test is shown in Table 2.

Table 2. Pearson Correlation Test

	GCS	CPI
GHP	0.323**	0.052
HP	0.059	-0.487**
PR-S	0.062	0.438**
PR-N	0.017	0.376**
GFI	0.613**	-0.660**
GPE	0.082	0.327**
* p<0.05 ** p<0.01		

It can be learned from the table 2 that growth rate of house price and fixed asserts investment are correlated with the growth rate of total consumer goods' sales at the level of 0.05. Either of the two independent variables shows a positive influence. While all the independent variables except the growth rate of house price are correlated with the consumer price index at the level of 0.05. Among them, house price and the growth rate of fixed asserts investment present negative effects, and the other variables

DOI: 10.54254/2753-8818/52/2024CH0106

correlate CPI positively. As a result, GHP and GFI is used to explain the variable GCS, while PR-S, PR-N, GFI and GPE is used to explain the variable CPI.

Since the concerning variables are collected in Shanghai from 2014 to 2024 every month in the order of time, it is suitable to define the variables as time-series statistics. Before the further study is started, an ADF test is in need to test whether the statistics is stationary. And the result of the ADF test will decide if the variables should be differential processed before being substituted into the model and calculated. The results of the ADF test are listed in Table 3.

Table 3. Results of the ADF Test

Variable	T(Unpossessed)	P	Stationary	T(First Difference)	P	Stationary
GCS	-5.908	0.000	Yes			
CPI	-0.780	0.825	No	-5.601	0.000	Yes
GHP	-2.857	0.051	Yes			
HP	-1.813	0.374	No	-9.284	0.000	Yes
GFI	-4.052	0.001	Yes			
GPE	-1.829	0.366	No	-2.933	0.042	Yes
PR-S	-1.879	0.342	No	-4.528	0.000	Yes
PR-N	-1.648	0.458	No	-3.917	0.002	Yes

It is presented in Table 3 that the variables except GCS, GHP and GFI are not stationary. To deal with the situation, the ECM model can be taken into use, which is a model often used to cope with time-series statistics without stationarity. For non-stationary time series, it is proper to be transformed into stable sequences through differential methods. The ECM model is the EG two-step method (EG-ADF method). Firstly, OLS linear regression is established for X and Y, and residual ECM is obtained. Then, Y and X are taken as first-order differences to obtain det(Y) and det(X), and ECM is taken as first-order lag, i.e., ECM (-1). Finally, the model will establish OLS linear regression for det(Y), det(X) and ECM (-1). The constant ECM model form is performed as:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + \delta Y_{t-1} + \varepsilon \tag{1}$$

3. Result and Discussion

3.1. The ECM model of GCS

As mentioned in the former part, the variable GCS can be explained by the variables GHP and GFI. Before the statistics is adapted into the ECM model, taking a co-integration test is in schedule to find out whether the statistics should be processed. The co-integration between three variables can be learned from Table 4.

Table 4. Johansen Co-integration Test (Based on the Trace Statistics)

H0	Eigenvalue	Trace	Critical value of 10%	Critical value of 5%	Critical value of 1%
None	0.624	135.216	27.067	29.796	35.463
At most one	0.136	27.626	13.429	15.494	19.935
At most two	0.099	11.500	2.705	3.841	6.635

According to the table, the hypothesis’ None ‘, ‘At most 1 co-integration’ and ‘At most 2 co-integrations’ are rejected at the 1% level as their trace statistics all go beyond the critical values. Consequently, the concerning indexes have a co-integration relationship between each other, which means there exists a balanced connection between the variables. As a result, the statistics is allowed to be used in the ECM model.

DOI: 10.54254/2753-8818/52/2024CH0106

On the basis of the co-integration test, the next step to take is the co-integration regression model, which works on the long-term equilibrium relationships. The results of the model are presented in Table 5.

Table 5. Co-integration Regression

	Coefficient	Std. Error	t	p
c	-3.433	1.736	-1.977	0.051
GHP	0.227	0.061	3.701	0.000
GFI	1.121	0.137	8.164	0.000
F	F (2, 108)=43.521, p =0.000		Durbin-Watson	1.034
nobs	111		AIC	895.490
R2	0.446		BIC	903.618
Adjusted R2	0.436		RMSE	13.300

According to the table, the formula of the co-integration equation model is performed as:

$$GCS = -3.433 + 0.227 * GHP + 1.121 * GFI \tag{2}$$

The R value of the cointegration equation model is 0.446. From the perspective of the co-integration equation test, the co-integration equation model passed the F-test (F=43.521, p=0.000<0.1). For the long-term equilibrium equation parameter test, the coefficient of housing price growth rate was 0.227, showing a significant 0.01 level (t=3.701, p=0.000<0.01), indicating that housing price growth rate will have a positive long-term equilibrium relationship with the growth rate of total retail sales. Meanwhile, the coefficient value of fixed assets investment growth rate is 1.121, showing a significant 0.01 level (t=8.164, p=0.000<0.01), which means that the growth rate of fixed assets investment will have a positive long-term equilibrium impact on the growth rate of total retail sales.

When finishing the co-integration regression, an ECM model follows, studying the short-term volatility relationship, which includes the first-order difference component of the sequence and the previous period error value, i.e. ECM (-1). Table 6 explains the adaption of the ECM model.

Table 6. ECM Model (Error Corrected) of GCS

	Coefficient	Std. Error	t	p
c	-0.205	0.898	-0.229	0.820
det(GHP)	0.189	0.060	3.161	0.002
det(GFI)	3.101	0.247	12.576	0.000
ECM(-1)	-0.777	0.074	-10.465	0.000
F	F (3, 105) =75.827, p =0.000		Durbin-Watson	1.542
nobs	109		AIC	801.055
R2	0.684		BIC	811.821
Adjusted R2	0.675		RMSE	9.197

From the table 6, it can be learned that the ECM model formula is presented as:

$$GSC = -0.205 + 0.189 * \det(GHP) + 3.101 * \det(GFI) - 0.777 * ECM\ (-1) \tag{3}$$

The R value of the ECM model is 0.684. From the perspective of model testing, the ECM model passed the F-test (F=75.827, p=0.000<0.1). Regarding the impact of short-term fluctuations, the coefficient of the difference in housing price growth rate was 0.189, showing a significant 0.01 level (t=3.161, p=0.002<0.01), indicating that the current fluctuation of housing price growth rate will have a positive adjustment on the current fluctuation of the total retail sales growth rate of goods.

The coefficient value of the growth difference sub item of fixed assets investment is 3.101, showing a significant 0.01 level (t=12.576, p=0.000<0.01), which means that the current fluctuation of the growth rate of fixed assets investment will have a positive adjustment to the current fluctuation of the total retail sales growth rate.

DOI: 10.54254/2753-8818/52/2024CH0106

Besides, the value of the error correction coefficient is -0.777, showing a significant level of 0.01 (t=-10.465, p=0.000<0.01), indicating that when short-term fluctuations deviate from long-term equilibrium, the non-equilibrium state will be pulled back to equilibrium with an adjustment force of -0.777.

3.2. The ECM model of CPI

Similarly, the variables to be used in the ECM model, including CPI, HP-S, HP-N, GFI and GPE are supposed to take a co-integration test in the beginning. According to the statistical results, the hypothesis that there exist at most 3 co-integration relationship cannot be refused. Consequently, the concerning statistics has to be differenced at first order. The descriptive analysis of the processed data can be seen in Table 7.

Table 7. Descriptive Analysis of Variables at First Difference

Variable	Min	Max	Mean	Std.Deviation
Diff1(CPI)	-1.600	1.240	-0.025	0.271
Diff1(HP-S)	-10.300	8.600	-0.142	2.017
Diff1(HP-N)	-6.700	6.900	-0.085	1.938
Diff1(GFI)	-14.600	17.900	0.046	4.138
Diff1(GPE)	-34.200	50.100	-0.120	9.771

After differencing the variables lack of co-integration, the processed data is supposed to pass the next co-integration test. The results are shown in Table 8. The further analysis is aimed to base on the co-integration relationship between the concerning variables.

Table 8. Johansen Co-integration Test (Based on the Trace Statistics)

H0	eigenvalue	Trace	10%	5%	1%
None	0.619	229.162	65.820	69.819	77.820
At most 1	0.343	124.907	44.493	47.855	54.681
At most 2	0.311	79.565	27.067	29.796	35.463
At most 3	0.230	39.332	13.429	15.494	19.935
At most 4	0.098	11.145	2.705	3.841	6.635

Regarding the hypothesis of ‘up to 4 co-integrations’, its trace statistic value is 11.145, and the 1% critical value is 6.635, which means that the hypothesis is rejected at the 1% level. The result indicates that the co-integration exists among the variables, which allows the adaption to the ECM model to continue.

Similar to the progress finished in the former part, the data of the five variables is substituted into the co-integration regression model to find out the long-term relationship. The first step--the co-integration regression is presented in Table 9, whose result reveals the influencing effect of the independent variables and attempts to predict further changes.

Table 9. Co-integration Regression of CPI

	Coefficient	Std. Error	t	p
c	-0.017	0.019	-0.940	0.349
Diff1(HP-S)	0.045	0.019	2.340	0.021
Diff1(HP-N)	-0.015	0.020	-0.739	0.462
Diff1(GPE)	0.000	0.002	0.123	0.902
Diff1(GFI)	-0.045	0.005	-10.047	0.000
F	F (4,105)=27.200,p =0.000		Durbin-Watson	1.820

Table 9. (continued).

nobs	110	AIC	-44.265
R2	0.509	BIC	-30.763
Adjusted R2	0.490	RMSE	0.189

The table 9 above shows the adaption to the co-integration regression model, which is also called the long-term equilibrium equation. The formula of the co-integration equation model is:

$$Diff1(CPI) = -0.017 + 0.045\ Diff1(HPS) - 0.015\ Diff1(HSN) - 0.045\ Diff1(GFI) \quad (4)$$

The R square value of the co-integration equation model is 0.509. From the perspective of the co-integration equation test, the co-integration equation model passed the F-test (F=27.200, p=0.000<0.1). For the parameter test of the long-term equilibrium equation, the coefficient value of the Diff1(HP-S) was 0.045, showing a significant 0.01 level (t=2.340, p=0.021<0.05), indicating that the Diff1 second-hand housing price index will have a positive long-term equilibrium relationship with the Diff1(CPI).On the contrary, the coefficient of the growth rate of Diff1(GFI) is -0.045, showing a significant 0.01 level (t=-10.047, p=0.000<0.01), which means that the growth rate of Diff1(GFI) will have a negative long-term equilibrium impact on Diff1(CPI).

While another two variables fail to influence CPI in the long term. The coefficient value of the Diff1(PR-N) is -0.015. There is no significant difference (t=-0.739, p=0.462>0.05).The coefficient value of Diff1(GPE) is 0.000, which does not show significant (t=0.123, p=0.902>0.05). Consequently, the effects of the variables PR-S and GFI on CPI are much stronger than that of PR-N and GPE.

On the basis of the co-integration regression step, the study focuses on the short-term changes which are predicted to take place among the several variables. The error corrected formula and the influencing effects are worked out with the first-differenced data. The result is shown in Table 10.

Table 10. ECM Model (Error Corrected) of CPI

	Coefficient	Std. Error	t	p
c	-0.000	0.018	-0.027	0.979
Diff2(HP-S)	0.036	0.017	2.148	0.034
Diff2(HS-N)	-0.021	0.022	-0.961	0.339
Diff2(GPE)	0.000	0.001	0.243	0.809
Diff2(GFI)	-0.051	0.004	-11.908	0.000
ECM(-1)	-0.924	0.098	-9.402	0.000
F	F (5,102)=44.725,p =0.000		Durbin-Watson	1.926
nobs	108		AIC	-44.29
R2	0.687		BIC	-28.20
Adjusted R2	0.671		RMSE	0.186

It can be learned from Table-10 that the model is able to explain 67.1% part of the CPI changes, and the ECM model has passed the F-test (F=44.725, p=0.000<0.1). The formula is presented as: $Diff1(CPI) = 0.036Diff2(PRS) - 0.021Diff2(PRN) - 0.051\ Diff2(GFI) - 0.924\ ECM\ (-1)$.

Regarding the impact of short-term fluctuations, the coefficient value of the difference component of the Diff1(PR-S) is 0.036, showing a significant level of 0.05 (t=2.148, p=0.034<0.05), indicating that the current fluctuation of the Diff1 (PR-S) will have a positive adjustment to the current fluctuation of the Diff1(CPI). Also, the coefficient value of Diff1(GFI) is -0.051, showing a significant 0.01 level (t=-11.908, p=0.000<0.01), which means that the current fluctuation of Diff1(GFI) will negatively adjust

DOI: 10.54254/2753-8818/52/2024CH0106

the current fluctuation of Diff1(CPI). It is worthy of consideration that the impact of the variable PR-N shows no significance in the changes of CPI.

The error correction coefficient value is -0.924, showing a significant level of 0.01 ($t=-9.402$, $p=0.000<0.01$). The result shows that when short-term fluctuations deviate from long-term equilibrium, the non-equilibrium state will be pulled back to equilibrium with an adjustment force of -0.924.

3.3. Discussion

According to the results of the GCS model, the growth rate of housing prices has a positive equilibrium effect on the growth rate of total retail sales of social goods. This result supports the theory that rising housing prices drive consumption. The calculation scope of the total retail sales of social goods does not include the purchase of housing, but mainly reflects the purchasing power of the consumer market. The research results support the close correlation between changes in housing prices in Shanghai and the level of activity in the consumer market. At the same time, the growth rate of house prices and the growth rate of social fixed assets investment can jointly pull the volatile GCS back to equilibrium. Based on this, it can be seen that changes in housing prices have a certain controlling effect on social purchasing power, and this result has certain reference value in formulating policies for consumption. When there is abnormal or weak growth in social consumption levels, investment and regulation in the real estate market may help improve the situation.

From the research results, it can be seen that in the long-term and short-term fluctuations of CPI, changes in second-hand housing prices have a stronger impact compared to the new housing price index. This phenomenon largely reflects the housing market situation and residents' housing consumption concepts in Shanghai. The housing prices in Shanghai are at a relatively high level, and the cost of purchasing a house is relatively high. The purchase of expensive new houses often requires a long-term storage and preparation process, so the impact on monthly price indices is not significant. The transaction of second-hand houses relies on high cost-effectiveness and investment value, which has a long-term impact on the monthly price index changes. Its price fluctuations can effectively pull the price index back to the original benchmark model.

4. Conclusion

According to the research results, there is a positive equilibrium relationship between the growth rate of housing prices in Shanghai and the growth rate of social fixed assets investment and the growth rate of total retail sales of social goods. When the growth rate of total retail sales of social goods deviates from long-term equilibrium, the two independent variables are adjusted with a force of -0.777 to pull it back to equilibrium. Second hand housing price growth rate and social fixed assets investment rate have positive and negative impacts on social consumer price index respectively. The current fluctuations of both will pull the unbalanced CPI back to the original model with a force of -0.924.

The results of this study suggest a positive relationship between housing prices and changes in the consumer market. The rise in housing prices has driven up the purchase amount and purchasing power of residents. The research results also indicate that the price of second-hand houses in Shanghai has a significant positive impact on the consumer price index, while the impact of changes in the price of newly built houses is relatively weak.

This study observes the volatility of the consumer market through the growth rate of total retail sales of social goods and the consumer price index of residents. There are still certain shortcomings in the research, such as vague explanations of the impact of housing prices and insufficient grasp of macro data. In recent years, academic research on housing prices and consumption has often focused on consumers, with less attention paid to changes in the consumer market. In the future, further research on the relationship between the consumer market and housing prices can deepen the understanding of the regulatory power of real estate, and provide guidance for formulating policies to stimulate consumption in the context of downward pressure in the consumer market.

DOI: 10.54254/2753-8818/52/2024CH0106

Authors contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Wang W, Shen J and Wang W P 2020 Housing prices and consumer inequality. Journal of Shandong University (Philosophy and Social Sciences Edition), 6, 88-101.
[2] Cong Y and Song X W 2022 Housing Prices, Human Capital, and Urban Resident Consumption: An Empirical Study Based on 35 Large and Medium sized Cities in China. Financial Issues Research, 5, 67-76.
[3] Zhang C 2017 Analysis of the Impact of Rising House Prices on Resident Consumption: Mechanism Analysis, Model Construction, and Empirical Testing. Northeastern University of Finance and Economics.
[4] Hong T 2006 Real estate price fluctuations and consumption growth: empirical analysis and theoretical explanation based on Chinese data. Nanjing Social Sciences, 5, 54-58
[5] Shi T 2022 The Impact of House Prices on Household Consumption: An Empirical Study Based on Four Cities in Jiangsu Province. Construction Technology, 6, 105-108.
[6] Sun H, Wang Z H and Yao J 2022 The impact of housing prices on the upgrading of consumption structure: mechanisms and empirical evidence. Journal of Shanghai University of Finance and Economics, 24 (02), 61-77.
[7] Jin T and Jiang K 2012 Research on the Relationship between Chinese Residents’ Consumption and House Prices from the Perspective of Generalized Virtual Economy. Research on Generalized Virtual Economy, 3(3), 35-41
[8] Liu J and Chen B K 2021 Has the rise in housing prices widened consumer inequality in China? Economics (Quarterly). 21 (04), 1253-1274.
[9] Jin T and Chu M 2015 Revealing the Impact of China’s House and Stock Price Fluctuations on Economic Fluctuations and Resident Consumption. Research on Generalized Virtual Economy.
[10] Xu Y H and Wu L Y 2019 Dynamic relationship between population characteristics and housing price fluctuations in different regions of China. Statistical Research, 36 (1), 28-38.

DOI: 10.54254/2753-8818/52/2024CH0111

Analysis of the factors affecting the price of gold based on multiple linear regression

Jiaqi Zhang

Beijing NO.8 High School, Beijing, 100033, China

1811040403@stu.hrbust.edu.cn

Abstract. Gold has a variety of properties, in addition to the commodity properties, gold also has a monetary function. There are many investment values, so the price of gold is also much attention by investors. Many factors may affect the price of gold. In this paper, relevant data from December 2011 through December 2018 have been selected from a number of sources in order to provide as comprehensive a picture as possible of the various factors that may affect the price of gold. An Ordinary Least Squares regression model was used for the analysis. It was found that U.S. stock market conditions, gold mining conditions, U.S. dollar exchange rate, and crude oil-related indices have different correlation results from various perspectives of the index analysis. Besides, there is a significant positive relationship between the price of other precious metals and Gold Price. Meanwhile, The U.S. Economic Related Index has a significant negative impact relationship with Gold Price.

Keywords: OLS regression model, gold price, influencing factors.

1. Introduction

Gold has several properties, before it had the purchasing power, gold played the value of ordinary commodities, with commodity attributes. After gold is fixed to act as a general equivalent, it evolves into money. Gold naturally plays a monetary function in this situation [1]. In addition, gold also has other special investment values, like risk-avoiding, inflation-proof, and so on. Because of this, the price of gold also got more attention. However, the price of gold often fluctuates, leading to the impact of gold investment, and its characterization has also laid a theoretical analysis for the analysis of the gold reserve function [2]. Therefore, the study of the influence factors of gold price is significant.

There are many factors affecting the price of gold, such as the dollar index, inflation, and so on. Li used Vector Error Correction Model to study the inflation situation, the dollar index, and the dollar interest rate index on the impact of the international gold price. The research pointed out that inflation on the gold price has a significant negative correlation. In a certain period, the dollar interest rate index and the gold price have a negative correlation, but in the long term, there is a significant positive correlation [3]. Yan pointed out that the dollar value index has both a significant impact in the short term and the long term on the price of gold [4]. Zhang analyzed the mechanism of how factors affect China's spot gold price in the role of gold price. The study concluded that the RMB exchange rate has a positive correlation with the international gold price and the price of gold. The Chinese stock price index has a negative correlation with the U.S. Stock Price Index and China's spot gold price. After using the Vector Autoregression Model to analyze the Chinese stock price index, the U.S. Stock Price Index, and China's spot gold price, it concluded that the price of gold itself has the greatest impact, followed by interest rates [5]. Tu used the Generalized Method of Moments and Error Correction Model to analyze the data of gold price and money supply for a total of 117 months from 2002 to 2012 and found that money supply had a significant impact on gold price [6]. In addition to the economic factors mentioned above, many other factors affect the gold price.

Monetary policy and social conditions also have an impact on the gold price. Shang studied the role of the three influencing factors of the asset portfolio balance channel, signaling channel, and liquidity channel under different monetary policies. And found that the tight monetary policy will reduce the gold price; the steady monetary policy has less influence on the gold price; the loose monetary policy will increase the gold price, which shows that the monetary policy is also one of the influencing factors [7]. Lu studied the impact of the gold price on the GDP growth rate, unemployment rate, and other factors. The research found that its influence is not significant in general, but when the announced GDP growth rate, unemployment rate, and other data do not match the market expectations, the price of gold will change [8]. In the long term, the gold price is influenced by the following factors.

In addition, the gold price will be affected by some special events. Wu analyzed the change in the price of gold before and after four events in the trade friction between China and the United States. Only the event that the Chinese and U.S. governments both implement the measures to increase tariffs has a significant negative correlation with the gold price. Other three events that the G20 summit of the U.S. and China consultations, the Trump administration declared to raise tariffs on China, and China and the United States signed a trade agreement all have a positive correlation with gold prices [9]. Wu and Yang selected the data of 10 years before and after the financial crisis. Found that at the end stage of the financial crisis, there was a significant effect on the short-term changes in the price of gold, but the impact of shocks at the outbreak stage was not significant [10].

This paper will use the Ordinary Least Squares to analyze the gold price and other data that may be relevant to the price of gold in recent years. Examining the factors that influence the price of gold in the long run.

2. Methodology

2.1. Data source and description

There are many factors that may affect the price of gold, such as the U.S. stock index, the U.S. dollar index, the exchange rate between the U.S. dollar and other currencies, the price of precious metal futures, and the price of crude oil futures. Therefore, the data from December 2011 to December 2018 are selected for this paper in many aspects. There are sixteen data items in total. Includes Gold Price (GP), S&P 500 Index (SP), Dow Jones Index (DJ), Eldorado Gold Corporation (EG), EUR USD Exchange rate (EU), Brent Crude oil Futures (OF), Crude Oil WTI USD (OS), Silver Futures (SF), US Bond Rate (USB), Platinum Price (PLT), Palladium Price (PLD), Rhodium Prices (RHO), US dollar Index Price (USDI), Gold Miners ETF (GDX), Oil ETF USO (USO) and Cobalt Price (CO).

The research will use this data to describe as fully as possible the various factors that may affect the price of gold. The frequency of each data is the same, and they are all measured in days. The data are all from the Kaggle database.

2.2. Indicators selection

In this paper, SP and DJ are chosen to represent the situation of the U.S. stock market. EG and GDX represent the situation related to gold mining. EU represents the exchange rate between the U.S. dollar and other currencies. OF, OS, and USO represent the situation related to crude oil. SF, PLT, PLD, RHO, and CO represent the situation of precious metal futures. USB and USDI represent the situation related to the U.S. economy (Table 1).

The mean value of gold price in the sample is \$127.323, the minimum is \$100.92, and the maximum is \$173.2, while the standard deviation is 17.527, which indicates that the price of gold fluctuates a lot in the period selected in this paper. The standard deviation of the EUR USD Exchange rate is only 0.101,

DOI: 10.54254/2753-8818/52/2024CH0111

the standard deviation of the US Bond Rate is only 0.433, and the standard deviation of the Cobalt Price is only 0.493, which indicates that its volatility is small, while all other variables are more volatile especially Dow Jones Index and Silver Futures.

Table 1. Descriptive Analysis.

Items	Min	Max	Mean	Std. Deviation
GP	100.920	173.200	127.323	17.527
SP	122.060	293.090	204.49	43.832
DJ	11769.210	26833.471	18161.094	3889.752
EG	2.770	80.200	28.277	20.326
EU	1.039	1.393	1.208	0.101
OF	27.880	126.220	77.505	27.401
OS	26.550	110.300	70.153	23.472
SF	33170.000	65292.000	43284.478	7530.704
USB	1.358	3.239	2.263	0.433
PLT	775.600	1737.600	1183.915	273.842
PLD	470.450	1197.500	766.805	148.307
RHO	0.000	2600.000	1130.442	570.013
USDI	78.300	103.288	89.809	7.516
GDX	12.700	57.520	26.747	10.621
USO	7.820	41.600	22.113	11.431
CO	1.936	3.954	2.938	0.493

2.3. Research method

This paper aims to investigate the factors that affect the price of gold. Therefore, it will be analyzed using ordinary least squares regression (OLS) analysis, which can be used to study the effect of independent variables on dependent variables. The basic formula of ordinary least squares regression analysis is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_P x_P + \varepsilon \quad (1)$$

Where y is the dependent variable, x_i is the independent variable, β_0 is the regression constant, β_p represents the regression coefficient, and ε is the random error.

3. Results and discussion

3.1. Correlation analysis

According to the results of correlation analysis, all the data have a significant correlation with GP, so all of them are selected as indicators of this study. Based on the positive and negative correlation coefficients in the graph it can be concluded whether there is a positive or negative correlation between the variables. The results show that GP has a significant negative correlation with SP, DJ, USB, PLD, and USDI. Besides, GP has a significant positive correlation with EG, EU, OF, OS, SF, PLT, PLD, RHO, GDX, USO, and CO (Figure 1).

DOI: 10.54254/2753-8818/52/2024CH0111

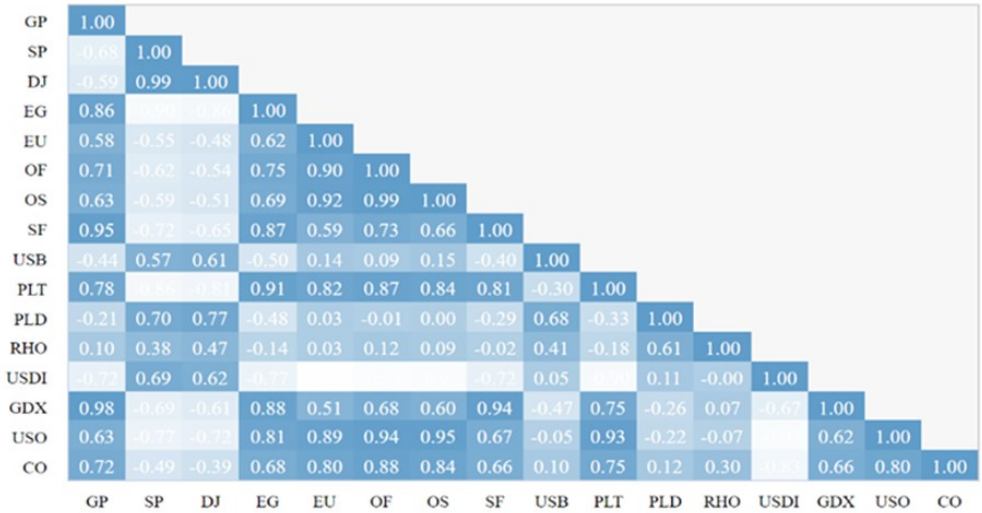


Figure 1. Correlation results.

In this study, GP was used as the dependent variable, and SP, DJ, EG, EU, OF, OS, SF, USB, PLT, PLD, RHO, USDI, GDX, USO, and CO were the independent variables. Based on this, the formula of the OLS model built is:

$$GP = 201.209 - 0.216 * SP + 0.002 * DJ - 0.049 * EG - 33.247 * EU + 0.095 * OF + 0.051 * OS + 0.000 * SF - 4.320 * USB + 0.016 * PLT + 0.003 * PLD + 0.001 * RHO - 0.903 * USDI + 0.783 * GDX - 0.843 * USO + 4.068 * CO \tag{2}$$

3.2. OLS regression results

In this study, GP was used as the dependent variable, and SP, DJ, EG, EU, OF, OS, SF, USB, PLT, PLD, RHO, USDI, GDX, USO, and CO were the independent variables. By using OLS regression, it can be seen that the model R2=0.987, which indicates that each independent variable can explain the cause of 98.7% of the change in GP. The F-test of the model was conducted and it was found that the model passed the F-test (F=8446.244, p=0.000<0.05). It shows that at least one of SP, DJ, EG, EU, OF, OS, SF, USB, PLT, PLD, RHO, USDI, GDX, USO, and CO will have an impact relationship on GP.

The regression coefficient is a parameter that indicates the magnitude of the effect of the independent variable on the dependent variable. The larger the value of the regression coefficient, the greater the influence of the independent variable on the dependent variable. A positive regression coefficient indicates that the dependent variable increases as the independent variable increases, which means that there is a positive influence relationship. A negative regression coefficient indicates means that there is a negative influence relationship.

The variable presents significance when P<0.01, which means that there is a significant influence relationship between the variable on the dependent variable. Conversely, it means that the variable does not have a critically effect relationship with the dependent variable. Among the indicators, all of them showed significance with p<0.01 except for OS with p=0.056>0.01.

From the analysis, it can be seen that DJ, OF, SF, PLT, PLD, RHO, GDX, CO will have a significant positive influence relationship on GP. and SP, EG, EU, USB, USDI, USO will have a significant negative effect on GP. However, OS does not have an impact on GP.

DOI: 10.54254/2753-8818/52/2024CH0111

Table 2. OLS model results.

	Coefficient	Standard Error	t	p	95% CI
Constant	201.209	14.474	13.902	0.000**	172.841 ~ 229.577
SP	-0.216	0.015	-13.964	0.000**	-0.246 ~ -0.185
DJ	0.002	0.000	9.030	0.000**	0.002 ~ 0.002
EG	-0.049	0.018	-2.762	0.006**	-0.083 ~ -0.014
EU	-33.247	5.635	-5.900	0.000**	-44.291 ~ -22.202
OF	0.095	0.020	4.814	0.000**	0.056 ~ 0.134
OS	0.051	0.027	1.916	0.056	-0.001 ~ 0.103
SF	0.000	0.000	14.533	0.000**	0.000 ~ 0.000
USB	-4.320	0.340	-12.712	0.000**	-4.987 ~ -3.654
PLT	0.016	0.001	14.955	0.000**	0.014 ~ 0.018
PLD	0.003	0.001	3.080	0.002**	0.001 ~ 0.006
RHO	0.001	0.000	8.148	0.000**	0.001 ~ 0.001
USDI	-0.903	0.089	-10.176	0.000**	-1.076 ~ -0.729
GDX	0.783	0.026	30.073	0.000**	0.732 ~ 0.834
USO	-0.843	0.067	-12.526	0.000**	-0.975 ~ -0.711
CO	4.068	0.291	13.994	0.000**	3.498 ~ 4.637
R2	0.987				
Adj R2	0.987				
F	F (15,1702)=8446.244,p=0.000				
D-W	0.274				
Dependent Variable: GP					
* p<0.05 ** p<0.01					

4. Conclusion

This paper finds that in the long run Dow Jones Index, Brent Crude oil Futures, Silver Futures, Platinum Prices, Palladium Prices, Rhodium Prices, Gold Miners ETF, and Cobalt Prices have a significant positive impact on GP. S&P 500 Index, Eldorado Gold Corporation, EUR USD Exchange rate, US Bond Rate, US Dollar Index Price, Oil ETF USO significantly negatively impact gold price.

From the results, it can be seen the overall situation of the U.S. stock market, the situation related to gold mining, and the exchange rate between the U.S. dollar and other currencies. the situation related to crude oil, the situation of precious metal futures, and the situation related to the U.S. economy all have an impact on the price of gold. However, only the situation of precious metal futures has a positive relationship with the Gold Price. Only the situation related to the U.S. economy has a negative relationship with Gold Prices. Other aspects of the data from different perspectives have different results, still need to follow up on this aspect of a more in-depth and comprehensive study.

This paper explores the impact of various factors on the price of gold, and the results of this paper can help the research on gold price prediction. Some factors that can be based on the data selected in this paper can be certain conclusions, but there is a need to carry out more in-depth research to confirm. Some factors need to be analyzed by selecting a wider range of data to obtain more certain conclusions.

References

[1] Wang R 2014 Influencing factors and trend analysis of international gold price. Shandong University of Finance and Economics.
[2] Zou Q 2014 Research on the financial function of gold. Shanghai Academy of Social Sciences.
[3] Li J T 2021 Analysis of Factors Affecting the Trend of Gold Price. University of International Business and Economics.

DOI: 10.54254/2753-8818/52/2024CH0111

[4] Yan H X 2022 Empirical Analysis and Suggestions on the Influencing Factors of Gold Price. Shanghai University of Finance and Economics.
[5] Wu Z H 2023 Study on the impact of China-US trade friction on international gold price changes. Yunnan University of Finance and Economics.
[6] Tu Z 2014 Research on the impact of China's money supply on gold price. Southwest University of Finance and Economics.
[7] Shang T Y 2022 Research on the impact of monetary policy on gold price. Guizhou University of Finance and Economics.
[8] Lu Q Y 2022 Analysis of factors influencing the international gold price. Shanghai Academy of Social Sciences.
[9] Zhang Q 2021 Research on the influencing factors of Chinese spot gold price. Shenyang University of Science and Technology.
[10] Wu Y M and Yang F 2014 Dynamic analysis of influencing factors of gold price before and after financial crisis. Gold, 35(8).

DOI: 10.54254/2753-8818/52/2024CH0102

Research on the factors affecting diabetes mellitus based on logistic regression

Junxi Li

School of Applied Statistics, Northeastern University, Qinhuangdao, 066000, China

202119298@stu.neuq.edu.cn

Abstract. In recent years, more and more diabetic patients have appeared all over the world, and people have begun to pay more and more attention to this kind of health problems, and it is a necessary thing to understand the influencing factors related to diabetes mellitus. This study explores the key factors that influence the occurrence of diabetes using multivariate logistic regression analysis and can be used to predict diabetes in individuals. The data for the study was obtained from the Kaggle website and various factors affecting diabetes were analyzed and a multivariate logistic regression model was developed to assess the impact of different factors on the risk of developing diabetes. The study found that good lifestyle habits and better basic personal circumstances the lower the risk of developing diabetes. These findings emphasize the importance of individuals focusing on their daily habits and improving their quality of life, which can help individuals reduce their risk of diabetes, and for those who are potentially at risk of developing diabetes, personal information can be used to make predictions and provide appropriate advice to help them change their bad habits.

Keywords: Diabetes, multiple logistic regression, risk factor, predictive model.

1. Introduction

Diabetes is one of the most prevalent chronic diseases globally, affecting a large population each year and imposing a significant financial burden on the economy. Ten years ago, experts predicted that by 2030, the global population suffering from diabetes would exceed 366 million, with China reaching 42.3 million. However, data from 2008 show that the number of diabetic patients had already reached 94.5 million, far exceeding the previously projected value for twenty years later [1]. This alarming increase in incidence underscores the urgency of understanding the factors contributing to diabetes. Identifying these factors is crucial for several reasons. Firstly, understanding the risk factors can aid in the prevention and control of diabetes, facilitating the development and implementation of effective health policies [2]. Secondly, recognizing these factors allows for the personalization of treatment regimens, which can improve treatment outcomes and enhance the quality of life for patients.

Many factors influencing diabetes are closely related to lifestyle, such as diet, physical activity, and smoking. However, diabetic patients often struggle with self-management, showing the best compliance with medication but the worst with blood glucose monitoring. Age, gender, self-efficacy, and diabetes-related knowledge are significant contributors to diabetes [3]. Therefore, managing diabetic patients requires not only personal effort but also substantial public healthcare resources, support from society, government, and hospitals, and increased awareness-raising efforts [4]. Therefore, all people need to pay attention to this issue, not only individuals need to understand the factors that cause diabetes, but also the relevant authorities should take certain measures. Healthcare professionals should prioritize regular assessment of the debilitating conditions of diabetic patients and develop professional, multifaceted, and safe intervention programs. These programs should consider the specific conditions of elderly patients with diabetes and aim to standardize and improve overall care [5]. Regular phone and SMS follow-ups for type 2 diabetes patients have been shown to improve lifestyle and behavioral habits, contributing to better disease control and rehabilitation [6].

For the data selection in this study, factors that may lead to diabetes were chosen, including lifestyle habits such as smoking, physical activity, daily consumption of vegetables, mental health, physical health and fruits and some personalized basic information such as age, education level and income. Personalized diet and exercise interventions for pregnant women, for instance, can effectively control weight gain during pregnancy, reduce the incidence of gestational diabetes mellitus (GDM), and ensure the health of pregnant women [7]. Wang highlighted that people who remain sedentary for long periods, do not engage in physical activity, and consume large amounts of alcohol are more likely to develop diabetes [8]. While moderate alcohol consumption may reduce diabetes prevalence, the overall trend shows an increase in diabetes prevalence with higher alcohol intake.

This study will analyze the impact of these factors on diabetes, providing insights that can help individuals reduce harmful habits in their daily lives. Additionally, the financial burden of diabetes on patients, families, and society is enormous. By studying the influencing factors of diabetes, valuable references can be provided for reducing the burden of disease and improving the efficiency of medical resource utilization.

2. Methodology

2.1. Data source and description

The data is derived from the Diabetes Health Indicators Dataset, a pure set of CDC's (Centers for Disease Control and Prevention) BRFSS2015 survey responses. The Behavioral Risk Factor Surveillance System (BRFSS) is an annual CDC survey on health-related issues. Each year, the survey collects responses from over 400,000 Americans on health-related risk behaviors, chronic health conditions, and the use of preventative services. It has been conducted every year since 1984. A CSV of the data set available on Kaggle for the year 2015 has been used for this project. This dataset consists of 441,455 individuals and has 330 features. In it I screened out 44,009 data and studied 12 of them. These are either questions that the participants are asked directly, or they are calculated on the basis of the responses of the participants.

2.2. Index selection and description

The target variable Diabetes_012 has 3 classes. 0 is for no diabetes or only during pregnancy, 1 is for prediabetes, and 2 is for diabetes. There are also a number of independent variable indicators. As shown in Table 1.

In the table it is possible to see the names of the variables that influence the factors of diabetes, as well as each variable their type: binary and continuous. In the last column is a description of each variable, explaining in detail what each variable corresponds to.

DOI: 10.54254/2753-8818/52/2024CH0102

Table 1. Variable interpretation.

Variable	Type	Meaning
HighCoal	Binary	0 for no high cholesterol 1 for high cholesterol
Smoker	Binary	Respondents smoked at least 100 cigarettes in their entire lifeRespondents smoked at least 100 cigarettes in their entire life.0 for no and 1 for yes
HeartDiseasorAttack	Binary	coronary heart disease (CHD) or myocardial infarction (MI). 0 for no and 1 for yes
PhysActivity	Binary	physical activity in past 30 days (not including job). 0 for no and 1 for yes
Fruits	Binary	Consuming fruit 1 or more times per day 0 for no and 1 for yes
Veggies	Binary	Consuming fruit 1 or more times per day 0 for no and 1 for yes
PhysHlth	Continuous	How many days during the past 30 days when physical health not good
MentHlth	Continuous	How many days during the past 30 days when mental health not good
Sex	Binary	0 for female and 1 for male
Education	Continuous	Education level ranging from 1 to 6
Income	Continuous	Income scale ranging from 1 to 8 1 for less than $10,000, 5 for less than $35,000 and 8 for $75,000 or more

2.3. Method introduction
In this study, Multiple Logistic Regression Analysis was employed to investigate the factors influencing diabetes [9, 10]. This method is suitable for the research as it allows people to assess the relationship between multiple independent variables and a categorical dependent variable. The first step is to clean the data, processing and removing outliers to ensure that the retained data is viable for this analysis. The basic information of the independent variables was then analyzed with descriptive statistics, examined the mean and distribution of each variable, roughly analyzed the relationship between each variable and the presence or absence of diabetes through scatter plots. After that, the logistic regression model was specified with diabetes status as the dependent variable and the selected independent variables as predictors.

The logistic regression model was fitted using maximum likelihood estimation. This process involved iteratively adjusting the model parameters to maximize the likelihood of observing the given data.

3. Results and discussion

3.1. Descriptive analysis
Analyzed through descriptive statistics, it can be seen that the dependent variable diabetes_012 is a categorical variable with three values, while the dependent variables HighChol, smoker, PhysActivity, HeartDiseasorAttacker, Fruits, and Vegies are binary variables that take values between 0 and 1 between 0 and 1, while the other variables are continuous variables. The highest frequency of no diabetes was found in the sample with 37,107 or 84.32% of the total, while 6,001 or 13.64% of the total were diabetic. Slightly more people did not have high cholesterol than those with high cholesterol, 57.87% and 42.13% respectively. Similarly, non-smokers outnumbered smokers by about 10%. Significantly more respondents did not have heart disease than those with heart disease, who accounted for less than 10 percent of the total. The study also found that more than 77% of the respondents exercised daily and

that the number of people who ate vegetables and fruits daily was also higher than the number of people who did not eat fruits daily.

The scatterplot (Figure 1) roughly show that cholesterol, smoking and heart disease are positively correlated with the prevalence of diabetes, while physical activity, healthy eating habits, education level and income are negatively correlated with the prevalence of diabetes. The results of these preliminary analyses suggest that healthy lifestyles and diets are more prevalent in the sample diabetes.

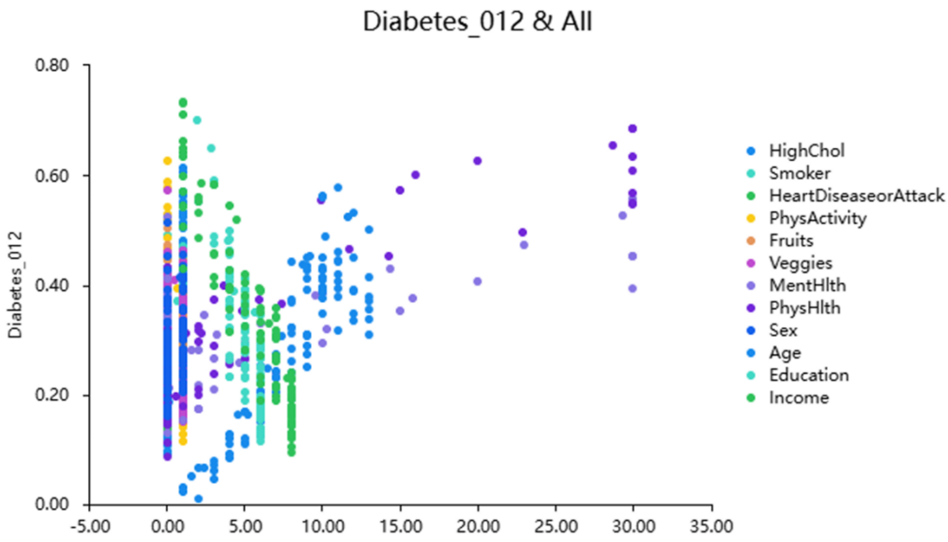


Figure 1. Scatterplot of variables

3.2. Model results

Multiple logistic regression (Table 2) analyses were acted to further validate the hypotheses of the model and to explore the effect of the respective variables on diabetes. The model used no diabetes (diabetes_012 equal to 0) as the control group. The first equation 1.0/0.0 means that, relative to 0.0 (no diabetes), with 1.0 (prediabetes). The second equation 2.0/0.0 means that, relative to 0.0 (no diabetes), at 2.0 (with diabetes).

The table 2 leads to the equation of multiple logistic regression:

$$\ln\left(\frac{1.0}{0.0}\right) = -3.346 + 0.779 \times \text{Highchol} - 0.103 \times \text{Smoker} + \cdots - 0.199 \times \text{Education} \tag{1}$$

$$\ln\left(\frac{2.0}{0.0}\right) = -1.825 + 0.858 \times \text{Highchol} - 0.067 \times \text{Smoker} + \cdots - 0.169 \times \text{Education} \tag{2}$$

The multivariate logistic regression model allows people to predict diabetes and to detect the risk of diabetes in the population. This paper can also analyze the influencing factors, relative to 0.0, the regression coefficient of HighChol is 0.779 with a p-value of less than 0.01 at 1.0, implying that HighChol has a significant positive effect on Diabetes, with an OR of 2.179, implying that the magnitude of the change in HighChol when increased by one unit of HighChol was 2.179 times (change from 0.0 to 1.0).Similar significant positive relationships with p less than 0.01 are also seen in MentlHlth and PhysHlth and Age. In contrast, Income and Education are included in the significant negative effects with p-values less than 0.01. As opposed to 0.0, in the context of 2.0. In the other group, as opposed to 0.0, in the case of 2.0. Except for Smoke and MentHlth, all other influences are significant.

Multiple logistic regression analysis also revealed that HighChole had a significant positive effect on diabetes. The relatively small effect of smoking suggests that the effect of smoking on diabetes is more complex and may be moderated by other factors. HeartDiseaseAttack significantly increased the risk of diabetes, emphasizing the importance of cardiovascular health in the management of diabetes. Physical activity had a large effect on diabetes, and regular physical activity is an important measure to prevent diabetes. Daily fruit and vegetable consumption habits had negative regression coefficients in both models, suggesting that good dietary habits can help reduce the risk of diabetes. Mental health and physical health had smaller regression coefficients in both models, suggesting that they have a smaller direct effect on diabetes, but further research is needed to investigate their potential indirect effects. Sex and age showed a significant positive relationship in both models, suggesting that males and those who are older have a higher risk of developing diabetes. Income and education had negative regression coefficients in both models, suggesting that higher income and education levels are effective in reducing the incidence of diabetes.

Table 2. Model results.

1.0	Cofficients	p	OR
Highchol	0.779	0.000	2.179
Somker	-0.103	0.141	0.902
HeartDiseasorAttack	0.030	0.786	1.030
PhysActivity	-0.163	0.039	0.850
Fruits	0.052	0.476	0.949
Veggies	-0.105	0.226	0.900
MentlHlth	0.013	0.002	1.013
PhysHlth	0.017	0.000	1.017
Sex	0.093	0.192	1.097
Age	0.086	0.000	1.090
Income	-0.064	0.000	0.938
Education	-0.199	0.000	0.820
Constant	-3.346	0.000	0.035
2.0	Cofficients	p	OR
Highchol	0.858	0.000	2.358
Somker	-0.067	0.028	0.935
HeartDiseasorAttack	0.524	0.000	1.689
PhysActivity	-0.323	0.000	0.724
Fruits	-0.091	0.004	0.913
Veggies	-0.137	0.000	0.872
MentlHlth	0.003	0.091	1.003
PhysHlth	0.023	0.000	1.024
Sex	0.244	0.000	1.276
Age	0.132	0.000	1.141
Income	-0.109	0.000	0.896
Education	-0.169	0.000	0.844
Constant	-1.825	0.000	0.161

4. Conclusion

This study provides an insight into the factors influencing diabetes mellitus through multivariate logistic regression analysis. It was found that high cholesterol, heart disease, smoking, less physical activity, unhealthy dietary habits, and lower income and education levels were significant diabetes risk factors. Conversely, regular physical activity, good dietary habits, and higher levels of income and education helped prevent the development of diabetes. These findings provide an important reference for the development of targeted diabetes prevention and management strategies, which can help to reduce the enormous burden of the disease on individuals, families and society, and optimize the efficient use of healthcare resources. Future studies can explore the indirect effects of mental health, social support and

DOI: 10.54254/2753-8818/52/2024CH0102

other factors on diabetes to improve the overall effectiveness of disease management and patients' quality of life.

References

[1] Yang W, et al. 2010 Prevalence of diabetes among men and women in China. New England Journal of Medicine, 362(12), 1090-1101.
[2] Musleh S, et al. 2020 Identification of potential risk factors of diabetes for the Qatari population. In 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies, IEEE, 243-246.
[3] Oyebode O and Orji R 2019 Detecting factors responsible for diabetes prevalence in Nigeria using social media and machine learning. In 2019 15th International Conference on Network and Service Management (CNSM).
[4] Sun S N, et al. 2011 The current status and influential factors of self-management in diabetic patients. Chinese Journal of Nursing, 3, 229-233.
[5] Li J, et al. 2014 Psychosocial factors influencing self-management behaviors in patients with diabetes: A review. Chinese Journal of Nursing, 2, 207-211.
[6] Jia W, et al. 2019 The current situation and influential factors of frailty in elderly patients with diabetes. Chinese Journal of Nursing, 2, 188-193.
[7] Guo X, Liu Z and Li H 2002 Type 2 diabetes mellitus induced by diets and its features of renal involvement in rat. Chinese Journal of Diabetes, 5, 35-39.
[8] Wang J H., et al. 2014 The effects of follow-up by telephone with SMS on lifestyle and behavior habits of patients with type 2 diabetes. Journal of Nursing Administration, 10, 701-703.
[9] Wang Z 2017 The relationship between sedentary behavior, alcohol consumption and diabetes mellitus in the prevalence of middle-aged and elderly people. Journal of Nursing Administration.
[10] Wu X, et al. 2015 Effect of dietary and exercise intervention on the incidence of GDM in women with pre-pregnancy obesity. Modern Clinical Nursing, 9, 24-27.

DOI: 10.54254/2753-8818/52/2024CH0105

Second-hand car price prediction based on multiple linear regression and random forest

Jiaying Gao

School of Mathematics and Artificial Intelligence, Chongqing University of Arts and Sciences, Chongqing, 400000, China

pengyuan@ldy.edu.rs

Abstract. The second-hand car market is a hot topic. Buying a second-hand car has advantages in price and many other aspects. Therefore, it is important to establish a good price prediction model. This paper will explore the factors that affect the price of second-hand cars. After analyzing and learning many kinds of literature, this paper establishes a multiple linear regression model and a random forest model and makes a comparative analysis of the model effect. The sum of the square error and R-square value of the random forest are better than the multiple linear regression model. Among the factors affecting the price of second-hand cars, the year of production has the greatest impact on the price, which shows that the age of the year is an important factor in determining the price of second-hand cars. The next most important factor is the number of kilometers traveled, followed by fuel type and transmission type-finally, engine displacement, number of transfers and number of seats. The random forest model established in this paper has better application value to price prediction.

Keywords: Second-hand car price, influence factor, multiple linear regression, random forest.

1. Introduction

With the development of the economy, people’s living standards have been improving gradually, and their happiness index has been increasing continuously. To facilitate travel, cars have become an essential part of people’s daily lives [1]. However, high purchase costs for new cars and rapid depreciation have always troubled consumers. Consequently, more and more consumers are turning to the second-hand car market, seeking more cost-effective car buying options. In the United States, for example, second-hand cars are traded at roughly three times the rate of new cars [2]. Against this backdrop, the trading volume of the second-hand car market is gradually increasing, and the importance of predicting second-hand car prices is becoming increasingly prominent.

The volatility of the second-hand car market brings significant uncertainty to both sellers and buyers, with various factors influencing second-hand car prices to varying degrees. Wang researched the different degree effects of vehicle age, power, mileage, model and other factors on the price [3]. In reality, buyers might fear purchasing vehicles with low value at high prices, while sellers worry that their vehicles may not fetch reasonable prices [4]. This uncertainty not only affects the market’s trading efficiency but also constrains its healthy development. Therefore, establishing a scientific price prediction model for second-hand cars is crucial in reducing market uncertainty and enhancing trading efficiency.

DOI: 10.54254/2753-8818/52/2024CH0105

Research on old car price prediction can enhance market efficiency and promote its healthy development [5]. By predicting second-hand car prices, consumers and car dealers can grasp market dynamics more accurately, thus making more rational decisions. Moreover, second-hand car prediction is conducive to stabilize market prices, prevent excessive price fluctuations, and maintains a level playing field in the market. Finally, predicting second-hand car prices is beneficial to avoid negotiations and disputes caused by unreasonable prices between consumers and car dealers. This reduces transaction costs and improves the overall efficiency of the market.

There are multiple methods and models to predict the price of second-hand cars. This paper tries to put the influencing factors of second-hand car prices under the test of objective data. Linear regression is a classical method in statistics. With the advent of the era of big data, machine learning algorithms have been studied and applied by more people. Zhu used several machine learning models and neural network models to predict the price of second-hand cars [6]. Zheng and others analyzed and compared the multiple linear regression model and the neural network model [7]. Zhu and Zheng concluded that neural network predictions have more accurate results for respective research problems. Jia built a model based on the Light Gradient Boosting Machine (GMB) algorithm and analyzed and compared it with other algorithms in machine learning to get the advantages of Light GMB [8]. Chetna and other researches compared the random forest algorithm with the e Xtreme gradient boosting algorithm through the integration technique. Finally, it is concluded that the e Xtreme boost algorithm is better [9]. Chen and his companions studied the different advantages of random forest and linear regression models in different situations, and random forest is better for dealing with complex models [10]. This paper mainly uses the linear regression method and random forest model to solve this problem and test the model effect. Preprocessing the data and performing descriptive statistics. This article will explore the relationship between second-hand price and vehicle condition. Because there are categorical variables in the independent, this article chooses to use dummy variable regression. In addition, this paper will also study random forest models to comprehensively analysis to solve this regression problem.

2. Methodology

2.1. Data source and description

The data for this article comes from the Kaggle website and has been updated by Milan Vaddoriya. There are a total of 5513 samples, 8 influencing factors and 1 dependent variable--car price.

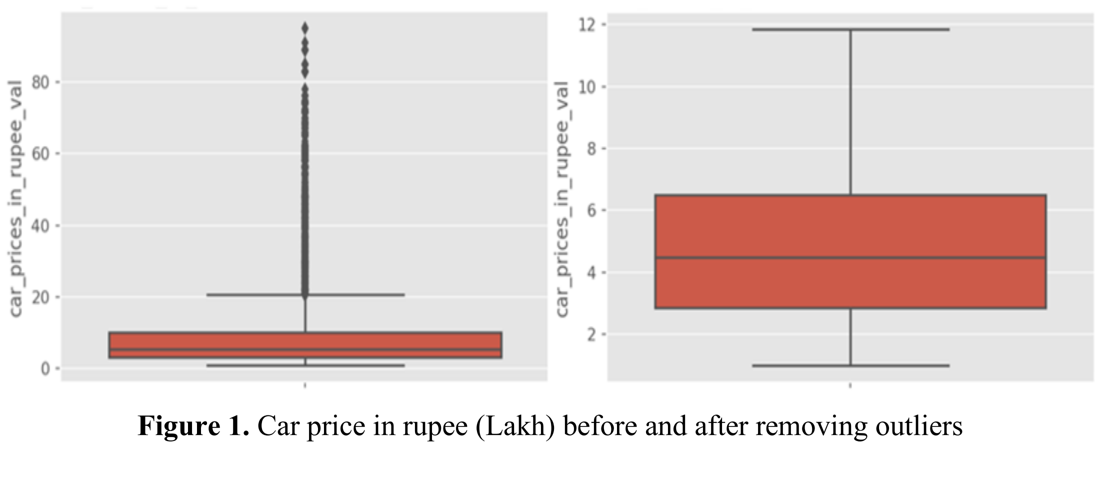


Figure 1. Car price in rupee (Lakh) before and after removing outliers

DOI: 10.54254/2753-8818/52/2024CH0105

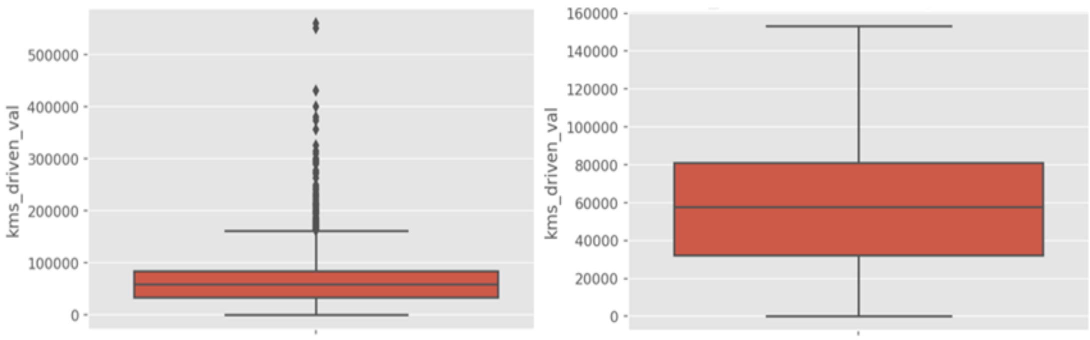


Figure 2. Kilometers driven before and after removing outliers

It can be seen from the box plot (Figure 1, 2) that there are too large values in the car price and kilometers traveled in the data, which not be conducive to the accuracy of the model. Therefore, these outliers are deleted. Now there are 4032 samples.

2.2. Index selection and description

Independent variables include fuel type, transmission type, ownership, seats, kilometers driven, manufacture year, engine capacity and car name. Ownership means the number of transfers, and the number of seats represents the size of the car. According to the objective phenomenon, the second-hand car brand is positively correlated with the price. In other words, the better the brand, the higher the price. However, car brands are used as categorical data, which cannot be assigned a specific value. Therefore, this variable will not be analyzed and further discussed for the following models.

To make the data better for modeling, it is necessary to analyze the data further. There are categorical variables and numeric variables in the influencing factors of second-hand car prices (Table 1). Although the year of manufacturing and the number of transfers is usually treated as discrete variables, to describe the effect of their changes on dependent variables, they are quantified as continuous variables and applied in mathematical methods. The significance of this is to describe their trends, not their numbers themselves.

Table 1. Variable description.

variable	symbol	type	range
car prices in rupee	x_1	numeric	1-11.85
kms driven	x_2	numeric	900-153367
cc _engine	x_3	numeric	0-5461
car name	x_4	-	-
fuel type	x_5	category	1=diesel, 2=petrol, 3=electric, 4=CNG, 5=LPG
ownership	x_6	numeric	1-5
seats	x_7	category	2,4,5,6,7,8
manufacture	x_8	numeric	1995-2022
transmission	x_9	category	1=automatic 2=manual

2.3. Method introduction

2.3.1. Multiple linear regression

Because there are categorical variables (transmission type, fuel type, seats) in the independent variables, the multiple linear regression analysis cannot be carried out directly. Therefore, it is necessary to use dummy variables and then carry out general linear regression. Dummy variables are created using petrol, automatic and ownership=5 as reference groups.

DOI: 10.54254/2753-8818/52/2024CH0105

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i + \varepsilon_i \quad (1)$$

2.3.2. *Random forest*

Random forest is an integrated learning algorithm based on decision trees, which can be used to solve regression problems. Random forest has good accuracy and stability. In the study of the factors affecting the price of second-hand cars, 90% of the data is used as a training set, and 10% is used as a test. Data normalization is used to train the data, and parameters are adjusted to find the optimal model.

3. Results and discussion

3.1. *Descriptive statistics*

There are five fuel types, mainly diesel and petrol, which far outnumber the other three. The number of second-hand petrol cars is almost 2500, and the number for diesel is about 1500. The number of second-hand cars of CNG, electric and LPG types is less than 100 (Figure 3).

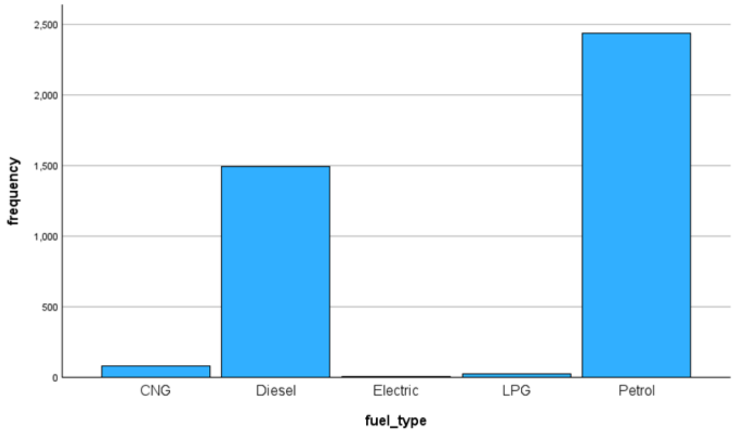


Figure 3. Bar Plot of Fuel Types

It can be seen that most of the cars in this data set are manual transmission cars. There are about seven times as many manual cars as automatic cars. The number of manual cars is about 3500, and the number of automatic cars is about 500 (Figure 4).

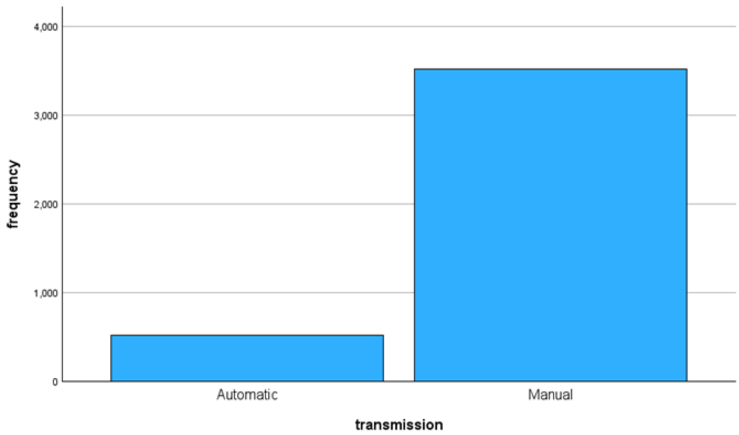


Figure 4. Bar Plot of Transmission Type

The manufacturing years of the cars in the data are mainly concentrated between 2010 and 2020. The most cars were produced in 2017, with nearly 500. Few cars were produced in the years before 2009, less than 100 cars a year (Figure 5).

DOI: 10.54254/2753-8818/52/2024CH0105

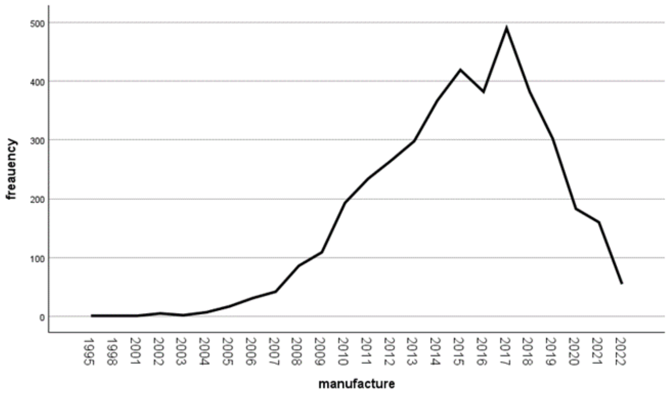


Figure 5. Line Plot of Car Counts by Manufacture Year

As can be seen from the figure, fuel type has a certain impact on second-hand car prices. The average price of second-hand cars of diesel type is the highest, and that of LPG is the lowest (Figure 6).

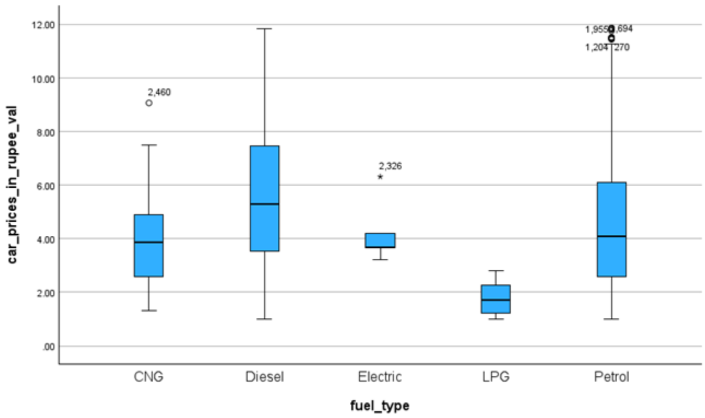


Figure 6. Box Plot of Fuel Type

The price of second-hand cars with automatic transmission is generally higher than manual transmission second-hand cars. The average price of a second-hand car with an automatic transmission is about 3 Lakh more expensive than a manual transmission (Figure 7).

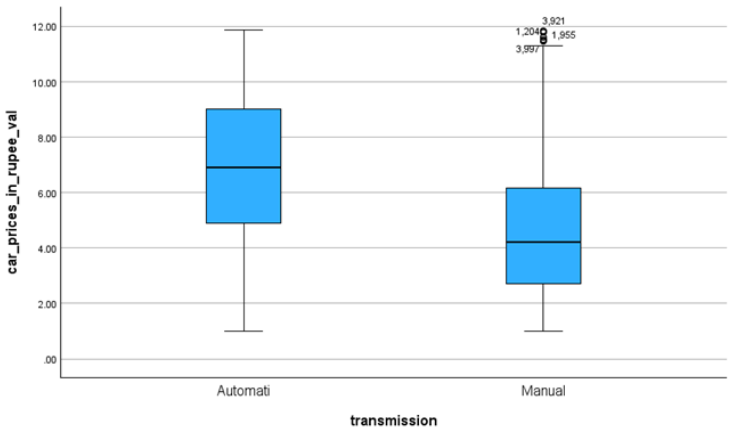


Figure 7. Box Plot of Transmission Type

DOI: 10.54254/2753-8818/52/2024CH0105

The number of transfers is negatively correlated with the mean value of the car price. Within three transfers, the price decreases significantly with the increase in the number of transfers (Figure 8).

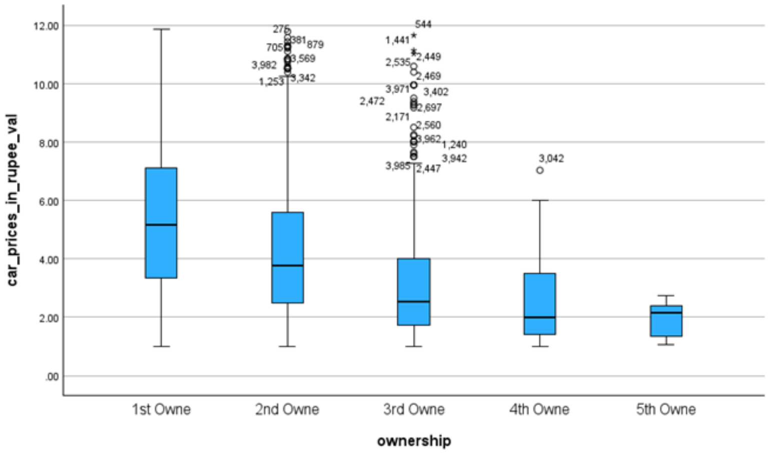


Figure 8. Box Plot of Ownership

There are 85.66% of second-hand cars are five-seater, accounting for the largest proportion. Followed by seven seats, and the proportion of other seats is seldom. The smallest number is the two-seater, with only four cars (Figure 9).

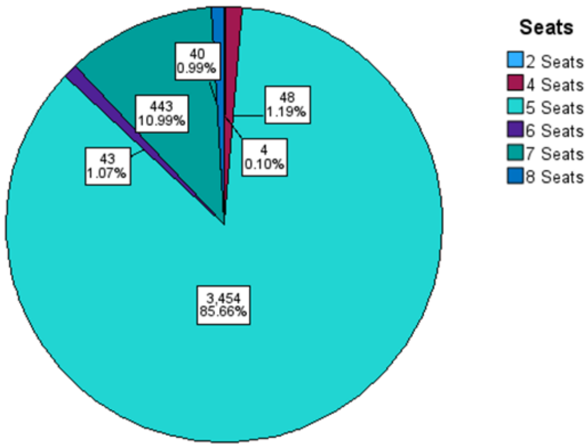


Figure 9. Pie Chart of Seats

3.2. Correlation analysis

This can be seen from Figure 10, the car price is positively correlated with the year of manufacture and is not significantly positively correlated with engine displacement and number of seats, negatively correlated with the number of kilometers driven and ownership. The number of kilometers traveled is positively correlated with the year of production and the number of transfers. There is a negative correlation between the year of production and the number of transfers. In addition, there is no obvious linear correlation between the other variables.

DOI: 10.54254/2753-8818/52/2024CH0105

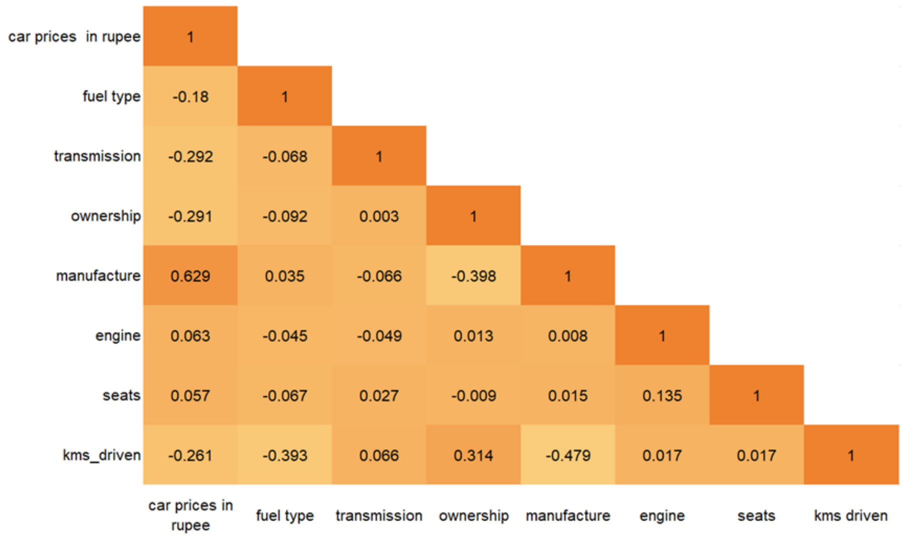


Figure 10. Pearson correlation

3.3. Multiple linear regression model results

From Table 2, note that VIF is much less than 5, and there is no multicollinearity between the independent variables. The model can be carried out accurately, and the parameter estimation is more reliable and accurate. As can be seen from Table 3, kms driven, transmission, fuel type, manufacture, and ownership have a significant impact on the dependent variable second-hand car price.

Table 2. Linear Model Coefficient

model	Unstandardized coef,		standardized	significance	VIF
	B	SE	Beta		
(constant)	-801.804	18.604		<0.001	
kms driven	-4.498E-6	0.000	-0.058	<0.001	1.63
transmission=manual	-2.041	0.084	-0.265	<0.001	1.027
fuel type=CNG	-0.692	0.201	-0.038	<0.001	1.03
fuel type=Diesel	1.467	0.064	0.275	<0.001	1.266
fuel type=Electric	-3.637	0.789	-0.05	<0.001	1.013
fuel type=LPG	-0.452	0.386	-0.013	0.242	1.013
seats=2 seats	-0.758	0.877	-0.009	0.388	1.001
seats=4 seats	-0.601	0.256	-0.025	0.019	1.011
seats=6 seats	-0.246	0.269	-0.01	0.361	1.006
seats=7 seats	0.253	0.09	0.031	0.005	1.032
seats=8 seats	0.078	0.28	0.003	0.779	1.008
manufacture	0.401	0.009	0.571	<0.001	1.493
engine	0.000	0.000	0.028	0.01	1.04
ownership	-0.258	0.043	-0.071	<0.001	1.223

The prediction value of second-hand:

$$car(x_1) = -801.804 - (4.498E - 6)x_2 - \cdots + 2.041transmission2manual + \varepsilon \quad (2)$$

However, there are some non-significant variables here that require further simplification of the model. A new linear regression model is obtained without considering the effect of seat number on price (Table 3).

DOI: 10.54254/2753-8818/52/2024CH0105

Table 3. Simplified model parameters

model	Unstandardized coefficient		standardized	significance
	B	SE	Beta	
(constant)	-803.901	18.613		<0.001
kms driven	-4.493E-6	0.000	-0.058	<0.001
transmission= manual	-2.040	0.084	-0.265	<0.001
fuel type=CNG	-0.682	0.201	-0.037	<0.001
fuel type=Diesel	1.467	0.064	0.275	<0.001
fuel type=Electric	-3.598	0.790	-0.49	<0.001
fuel type=LPG	-0.452	0.386	-0.013	0.242
manufacture	0.402	0.009	0.572	<0.001
engine	0.000	0.000	0.030	0.005
ownership	-0.258	0.043	-0.072	<0.001

The regression equation after the simplified model is:

$$car(x_1) = -803.901 - (4.493E-6)x_2 - 2.040transmission2manual - \cdots + \varepsilon \tag{3}$$

Table 4. Model summary

R	R square	adjusted R^2	SE	DW
0.732 a	0.536	0.535	1.75450	1.784

$$SE = \sqrt{\frac{SSE}{n-k-1}} = 1.7545 \tag{4}$$

The smaller the SE, the better the representation of the regression line to each observation point (Table 4). R^2 measures the degree of fitting of the model to the data.

$$R^2 = \frac{SSR}{SST} \tag{5}$$

The closer the model is to 1, the better the fitting effect is. R^2 is equal to 0.536. Avoid overfitting the adjusted value $R_a^2 = 0.535$. The model has a certain interpretability, but it does not have a high degree of fit.

3.4. Random forest model results

Of these factors, the most influential is manufacture year (58.79%), followed by kilometers driven (10.63%), the type of fuel and the transmission type, and the engine displacement, and finally the number of transfers and the number of seats (Table 5, 6).

Table 5. Characteristic weight value

variable	weighted value
manufacture	0.588
kms driven	0.106
fuel type diesel	0.092
transmission manual	0.090
engine	0.068
transmission automatic	0.029
ownership	0.015
seats	0.007
fuel type petrol	0.003

Table 5. (continued).

Fuel type CNG	0.001
fuel type LPG	0.000
fuel type electric	0.000

Table 6. Model evaluation

index	training set	test set
R-squared	0.701	0.602
MAE	1.059	1.182
MSE	1.983	2.595
RMSE	1.408	1.611
MAD	0.807	0.894
MAPE	9.382	1.135
EVS	0.701	0.602
MSLE	0.055	0.067

For test set,

$$SSE = MSE * n = 2.595 * 4032 = 10463.04 \tag{6}$$

The sum of squares of error between the predicted value and the observed value is 10463.04. $R^2 = 0.602$. This means that the independent variables can explain 60.2% dependent variables.

3.5. Discussion

In the multiple linear regression model, $SSE = 13080.80$. SSE is directly related to the model quality, the smaller its value, the higher the model fitting accuracy. The SSE in the multiple linear regression model is 1.19 times that of the test data in the random forest model. And the R-square in the random forest is greater than the R-square value in the linear regression model. The random forest model has a better fitting effect and accuracy than the regression model.

4. Conclusion

Compared with the multiple linear regression model, the random forest model can better predict second-hand car prices, the fitting effect of this model is better, and the error is relatively small. Among these factors, the year of production of a car has the greatest impact on the price of second-hand cars. For multiple linear regression models, some independent variables may not have linear effects on dependent variables, or the linear relationship between them is not obvious. These independent variables may include but are not limited to fuel type, and transmission type. This situation may lead to increased errors, which is a limitation of a linear regression model. Due to the excessive complexity of car names, this variable is not introduced into the model, and other factors not mentioned in the data affect second-hand car prices, which also leads to a decrease in the accuracy of the fit. The number of car accidents, car color and geographical location factors have a certain impact on the price of second-hand cars. If more factors are taken into account, the model will work better. This can be seen in the model results, the year of production is an important factor in the price of a second-hand car, while the number of seats and most fuel types have little effect on the price.

References

[1] Alhakamy A A, et al. 2023 Are used cars more sustainable? Price prediction based on linear regression. *Sustainability*, 15(2), 911.
[2] Gavazza A, Lizzeri A and Roketskiy N 2014 A quantitative analysis of the used-car market. *American Economic Review*, 104(11), 3668-3700.

DOI: 10.54254/2753-8818/52/2024CH0105

[3] Wang J 2019 Master of second-hand car valuation model based on random forest algorithm. *Beijing Jiaotong University.*
[4] Gajera P, Gondaliya A and Kavathiya J 2021 Old car price prediction with machine learning. *Int. Res. J. Mod. Eng. Technol. Sci*, 3, 284-290.
[5] Andrews T and Benzing C 2007 The determinants of price in internet auctions of used cars. *Atlantic Economic Journal*, 35, 43-57.
[6] Zhu X 2023 Master’s degree in price prediction of used cars based on deep learning. *Shandong Normal University.*
[7] Zheng A, Li B and Guo C 2023 Predictive analysis of used car transaction price based on linear regression and neural network model. *Intelligent Computer and Applications*, 9, 103-110.
[8] Jia P 2021 Master of second-hand car price prediction based on LightGBM. *Shandong Normal University.*
[9] Longani C, Prasad Potharaju S and Deore S 2021 Price prediction for pre-owned cars using ensemble machine learning techniques. *In Recent Trends in Intensive Computing*, 178-187.
[10] Chen C, Hao L and Xu C 2017 Comparative analysis of used car price evaluation models. *In AIP Conference Proceedings*.

DOI: 10.54254/2753-8818/52/2024CH0109

Research on gold price prediction based on ARIMA model

Yi Bai

College of Science, Civil Aviation University of China, Tianjin, 300300, China

210240001@cauc.edu.cn

Abstract. Gold is one of the most prevalent currencies in the world and its price has a very strong influence in the global financial markets. Gold has safe-haven properties, which can have a significant impact on its demand and price, especially in times of social unrest or financial crisis. Now, the demand for gold by investors has increased dramatically. Therefore, being able to accurately predict the direction of the gold price can help investors to effectively develop investment strategies and risk management measures. The overall objective of this study is to forecast the price of gold futures for the next six months. In this study, the Kaggle website was searched to find the price of gold from 2020 to 2024 and finally the CLOSE price was chosen as the final predicted price. This paper uses the ARIMA model for gold price forecasting. By comparing the RSME size of each model, ARIMA (1, 1, 2) is finally chosen. From the prediction results the price of gold remains stable in the first half of the year and then increases significantly. From the results of the residual test, there is no autocorrelation, and then it is white noise.

Keywords: Gold, ARIMA, autocorrelation, prediction.

1. Introduction

Gold as a special non-ferrous metal, both commodity properties and monetary properties, its good ductility, stable chemical properties determine it to become the most suitable for human commodity society as a currency commodity. At the same time, gold also has a good investment, the value of the storage function [1]. Gold tends to follow inflation, so when inflation is very high, the price of gold will also be high. At the same time, when inflation falls, the price of gold will fall. When the price of gold rises, investors will be more willing to invest in gold than in stocks [2]. Gold has been one of the most sought-after commodities for nearly a century and continues to hold a place in asset allocation. In March this year, the price of gold repeated record highs, has a very high potential for growth, the international price of gold has exceeded 2200 USD/ounce. This news has caused a high degree of concern of investors in the market.

With the development of the global economy and the increasing complexity of financial markets, the gold futures market has gradually become the focus of investors' attention. According to the World Gold Council, total positions in the global gold futures market reached more than $100 billion by the end of 2021, and trading volume in the global gold futures market grew by about 50% from 2010 to 2021. The status of gold futures market is increasing and gradually becoming a financial investment market as important as stock market, futures market, bond market and so on [3].

China has always been a major producer and consumer of gold, and in recent years China's economy is growing rapidly, China's investment demand for gold is increasing. Therefore, China has continuously strengthened the establishment and improvement of the gold market. The People's Bank of China in April 2001 announced the lifting of the "unified purchase and distribution". And then, the Shanghai gold exchange in October 2002 formally opened, mainly engaged in spot on-site gold trading, on the establishment of the gold hallmarks of China's entry into the era of market-oriented reform of the gold management system [4].

As an investment tool, the price of gold has stabilized and is always on an upward trend. When the financial market systemic risk, gold will become a hedge tool. 2008 after the financial crisis, the world economy is in the doldrums, but the price of gold is still strong, and since then the investment value of gold by a wide range of concerns. 2020 public health events superimposed on geopolitical conflicts, the financial markets of various countries have caused a serious impact on the many times triggered by the meltdown of the global stock market, the financial market turmoil, during the period of the price of gold has risen against the trend, and the hedge value has been revealed [5]. Therefore, accurately predicting the future direction of the price of gold can effectively avoid certain market risks, help investors to fully understand the market situation, but also can promote the development of the futures market.

Price forecasts for gold futures and the factors that influence the price of gold have been quite controversial in academia. Tully and Lucey investigated the relationship between the price of gold and the price of the USD using the Asymmetric Power generalized autoregressive conditional heteroscedasticity (APGARCH) model and found that the USD is an influential factor in the price of gold futures [6]. Zeng et al. provided a very accurate prediction of gold price based on an improved model of BP neural network with projection seeking optimization [7]. Singh et al. suggested an innovative periodic extreme learning machine for forecasting gold price data. It was found that there are now important links between the various factors [8]. Rapach et al. found that fusion models can fully utilize information from multiple variables and reduce volatility substantially [9].

Therefore, accurately predicting the future direction of the price of gold can effectively avoid certain market risks, help investors to fully understand the market situation, but also can promote the development of the futures market. In addition, gold futures are inextricably linked to the spot, and the price of gold futures is of great significance to the country's macroeconomy [10].

In summary, predicting the price of gold is of great importance. For gold futures price data, the use of machine learning methods for prediction works well, so this paper decides to use the ARIMA model to predict the settlement price of gold futures.

2. Methodology

2.1. Date source

The study used data form the website Kaggle website. This data set provides a comprehensive record of daily gold prices from January 1, 2020, to January 22, 2024. The data included key financial metrics for each trading day (Table 1).

Table 1. Descriptive statistics

Indicator	mean	standard error	standard deviation	Variance
Close	1498.726	5.9634	298.8248	89296.267
Volume	185970.7	1947.735	97600.769	9525910184
Open	1498.726	5.9692	299.1182	89471.69
High	1508.451	6.012	301.2622	90758.94
Low	1488.87	5.9154	296.4177	87863.455

As the settlement price of gold futures has the characteristics of non-linear and unstable fluctuation, and is closely related to many indicators, this paper refers to the research results of related scholars, based on which five variables are selected as the influencing factors of the settlement price of gold futures. It includes the closing price, turnover, opening price, maximum price and minimum price. These indicators are the most used basic indicators in futures price forecasting studies (Table 1).

DOI: 10.54254/2753-8818/52/2024CH0109

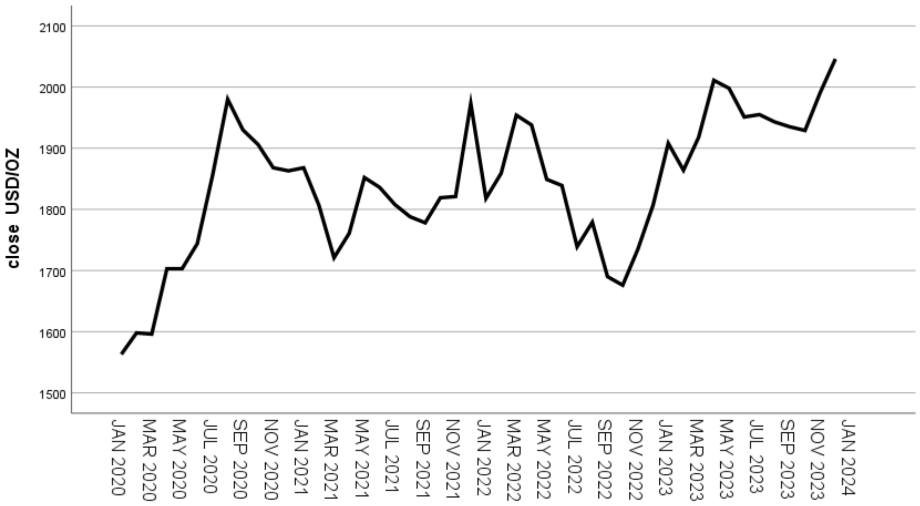


Figure 1. Monthly gold settlement price chart for the period 2020-2024

To see the changes more intuitively, the raw time series plot was chosen to be the average of the settlement price of each month from 2020 to 2024. As can be seen in Figure 1 between 2020 and 2024 the settlement price of gold futures is more volatile, with an overall range of \$1,484 to \$2,073 per ounce. However, there are two time periods with significant price increases and significantly higher price volatility: one in early 2020 and one in 2022. Since 2020, the global economy has faced multiple challenges. The rapid spread of the new Crown Pneumonia epidemic around the world resulted in a severe shock to the global economy. Governments have adopted restrictive measures that have led to lower corporate earnings, supply chain disruptions and poor logistics and transportation. Trade frictions between the United States and countries such as China and the European Union have also exacerbated uncertainty and turmoil in the global economy. And gold, as a relatively stable and safe asset, has naturally become the first choice of investors. Therefore, the deterioration of the global economic situation is an important factor driving up the price of gold.

2.2. Method introduction

Time series model is a statistical model used to analyze and forecast data over time. These data may be trending, seasonal or cyclical, so time series modeling can help people understand the structure of the data and make forecasts.

One of them is Autoregressive Integral Sliding Average (ARIMA) Model which is a classical time series analysis method for forecasting future time series data. ARIMA model is based on the autocorrelation and trendiness of the time series and models the intrinsic structure of the data through the three components of Autoregression (AR), Integration (I) and Moving Average (MA).

The modeling process of the ARIMA model requires the selection of appropriate values of p, d and q, where p is the order of the AR term, d is the number of differences which is required to make the time series stationery and q is the order of the MA term. These values are determined by looking at the autocorrelation (ACF) and partial autocorrelation (PACF) plots. Finally, an ARIMA model was fitted using historical data, residual analysis was performed, and future projections were made.

3. Results and discussion

3.1. Stationarity tests

The ARIMA model requires that the time series must be smooth, so the original data needs to be tested for smoothness. The Figure 1 shows that the series has a long term upward trend. Therefore, this paper needs to perform a first order difference on the data and the first order difference result is shown in Figure 2.

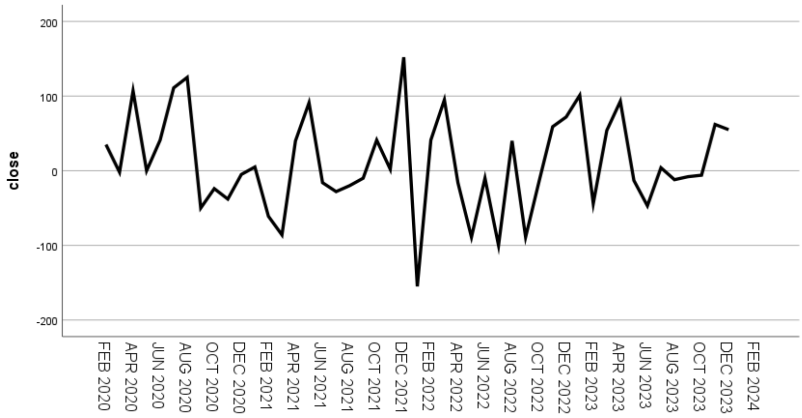


Figure 2. 1st order difference sequence

3.2. Determining ARIMA model order

The first order difference results shows that the data has smoothed out. Then the ACF and PACF plots are examined to observe the tailing, and the results are shown in Figure 3 and 4.

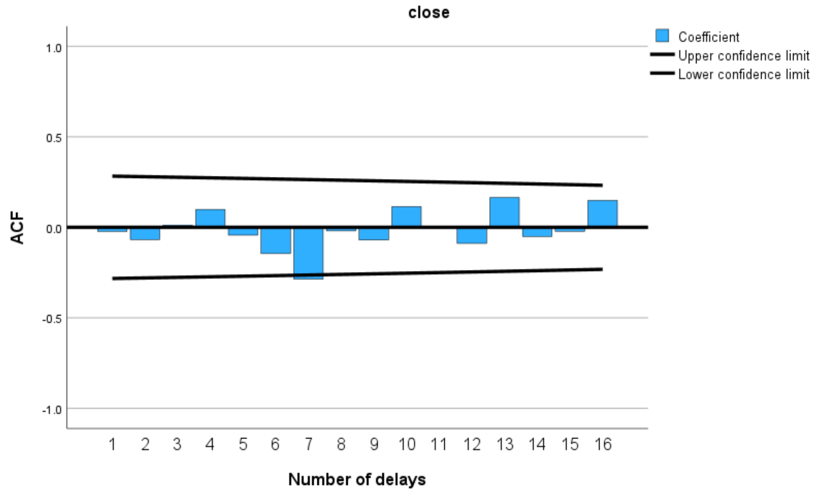


Figure 3. 1st order difference sequence ACF graph

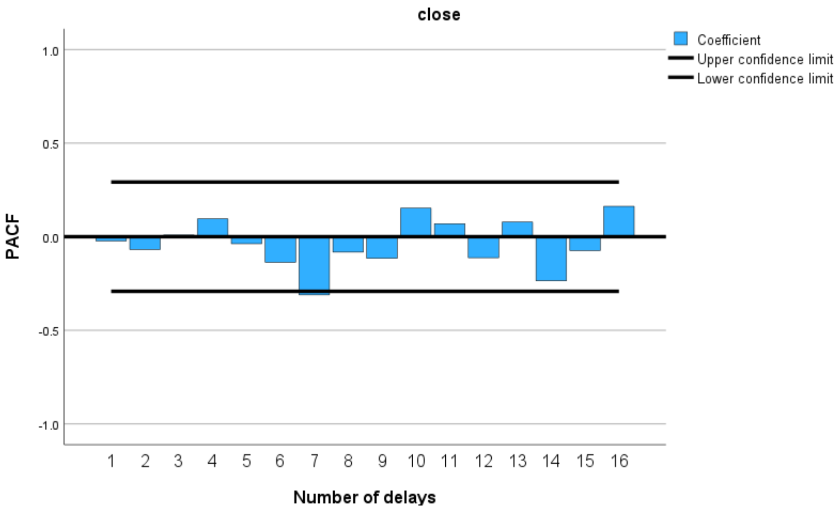


Figure 4. 1st order difference sequence PACF graph

DOI: 10.54254/2753-8818/52/2024CH0109

As can be seen from Figures 3 and 4, the ACF and PACF plots of the original sequence after first-order differencing are all of order 0 trailing, so the sequence after first-order differencing is a smooth sequence, and both p and q are equal to 0.

Table 2. Comparison of models

	AIC	RMSE
ARIMA (1,1,0)	530.616	64.2048
ARIMA (2,1,0)	532.397	64.0499
ARIMA (0,1,1)	530.612	64.2019
ARIMA (0,1,2)	532.432	64.0749
ARIMA (1,1,1)	530.026	62.5005
ARIMA (2,1,1)	534.395	64.0485
ARIMA (1,1,2)	531.856	62.3537

Comparing the fitting of the different ARIMA models. In the table 2, the AIC and RSME of ARIMA models for different values of p and q are shown. It can be found that the RSME of ARIMA (1, 1, 2) is the lowest among all the models. It has the second lowest AIC among all the models. Preference is given to RSME, so ARIMA (1, 1, 2) is finally chosen.

3.3. Model estimation

The most appropriate ARIMA model was selected by comparing the RSME under different models using the ARIMA auto-ranking method. The final model selected was ARIMA (1, 1, 2). Modelling was carried out using SPSS software. The fitting results are shown in Table 2, Table 3 and Figure 5.

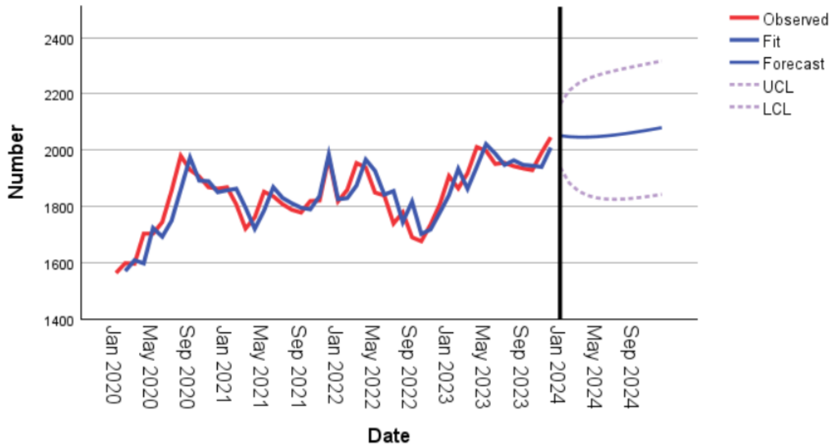


Figure 5. Fitting effect of modeling

The Figure 5 shows the gold settlement price for each month from January to December 2024. The measured and fitted values almost overlap and are very similar, which shows that the ARIMA model makes a better fit and prediction. The predicted values are also within the confidence interval, and from the predicted values.

Table 3. Data predication

	ARIMA (1,1,2)		
Month	Forecast	UCL	LCL
Jan-24	2052	2164	1939
Feb-24	2048	2213	1883

DOI: 10.54254/2753-8818/52/2024CH0109

Table 3. (continued).

Mar-24	2046	2238	1854
Apr-24	2046	2254	1839
May-24	2048	2266	1830
Jun-24	2051	2275	1826
Jul-24	2054	2283	1825
Aug-24	2059	2291	1826
Sep-24	2063	2298	1829
Oct-24	2069	2304	1833
Nov-24	2074	2311	1838
Dec-24	2080	2317	1843

The exact price of the projected gold settlement price for each month of 2024 is shown in Table 3. From the exact prices, there is a slight decrease in the price of gold in the first half of 2024, which generally remains unchanged. This is followed by an upward trend in the second half of 2024.

Table 4. Model statistics

Model	Model fit statistics		Ljung-Box			Outliers
	Stable R^2	R^2	Statistics	DF	Sig.	
ARIMA(1,1,2)	0.313	0.748	12.286	15	0.657	1

The residual PACF coefficients also approximate 0. The sig. term is greater than 0.05 and the residual series is free of autocorrelation in Table 4.

3.4. Residual test

It is also necessary to perform a white noise test on the residual terms If autocorrelation exists in the residuals, one should consider adding an autoregressive or sliding average interpretation, re-modeling and evaluating the model, and then performing a white noise test on the residuals of the new model, and so on, until the residuals are determined to be white noise.

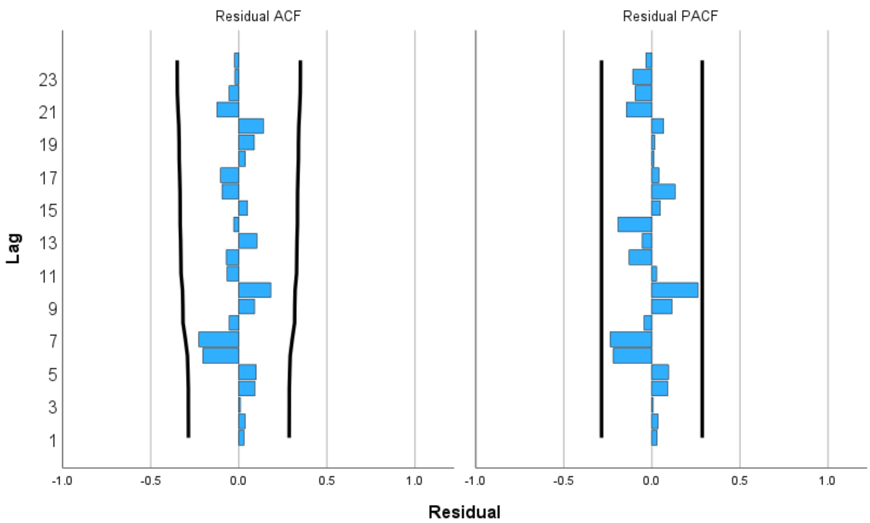


Figure 6. Final model fitting residual

DOI: 10.54254/2753-8818/52/2024CH0109

As can be seen in Figure 6, all autocorrelation coefficients fall within the confidence intervals after the 1st order differencing of the original time series. The residual terms were tested for normality. Q-Q plots were plotted in SPSS and the results are shown in Figure 7.

Most of the residual data from the model fit fall around the diagonal, indicating that the distribution of the data does not differ much from the theoretical one, and thus the residuals are considered to conform to the normal distribution. It can be determined that the residual term after model fitting is a white noise sequence.

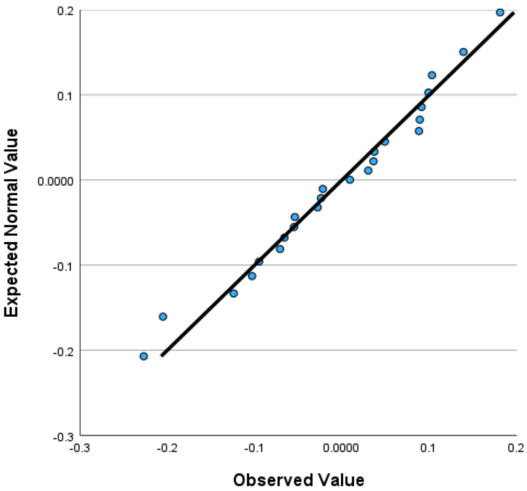


Figure 7. Q-Q plots of model fitting residuals

3.5. Discussion

Although the above model passes the test and shows some accuracy, some problems can still be found in the prediction graphs. The fitted values are slightly backward from the measured values and do not match exactly. With the overall upward trend in the predicted values, it can be assumed that the price of gold will steadily increase in the coming months and will break through the all-time highs. The main reason for this was the continued weakening of the United States dollar. The price of gold is settled in dollars, and when the dollar exchange rate becomes low, which means that the dollar depreciates, then the price of gold goes up. There is also the international political turmoil is now very big, some countries between the conflict is strong, triggering a reduction in the production of gold, the market supply of gold is reduced, prompting the price of gold rose.

4. Conclusion

This study used the AIRMA model to systematically forecast gold futures prices. According to the research, with today's social unrest and challenges such as ongoing geopolitical tensions may provide some support for gold prices. This paper argues that gold still has investment value in 2024, because the United States and other developed economies may face a shift in monetary policy, global instability, gold as a safe-haven asset is expected to increase in value. But the price of gold may be constrained by economic recovery and rising interest rates. Therefore, a comprehensive assessment of the suitability of buying gold in 2024 needs to be made based on individual circumstances and needs. Investors can follow the trend of gold prices, understand the market volatility and risks, and choose reputable gold brands for investment. The final decision should be made based on one's financial situation and investment objectives to get the best return on investment.

References

[1] Yang C 2023 International gold futures forecast based on VMD-LASSO-LSTM. Wuhan University of Science and Technology.

DOI: 10.54254/2753-8818/52/2024CH0109

[2] Abu D I, et al. 2023 Enhancing multilayer perceptron neural network using archive-based harris hawks optimizer to predict gold prices. J. King Saud Univ.-Comput. Inf. Sci.
[3] Liu C Y 2023 Research on Forecasting of Gold Futures Price Based on LASSO-QRNN. Central South University.
[4] Yang T T 2023 Research on price discovery and volatility spillover effect in China's gold market. East China University of Political Science and Law.
[5] Yang X Y 2023 Gold futures price prediction based on CNN-BiGRU model. Henan University.
[6] Tully E and Lucey B M 2007 A power GARCH examination of the gold market. Research in International Business & Finance, 21(2), 316-325.
[7] Lian Z, et al. 2010 Improved gold price prediction based on BP neural network. Computer Integrated Manufacturing Systems, 27(9), 200-203.
[8] Singh G, Tripathy B and Singh J 2023 An Evaluation of Extreme Learning Machine Algorithm to Forecasting the Gold Price. 2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 523-529.
[9] Rapach D E, Strauss J K and Zhou G 2010 Out-of-Sample equity premium prediction: Combination forecasts and links to the real economy. Review of Financial Studies, 23(2), 821–862.
[10] Jin W 2022 Gold futures price prediction research based on CEEMDAN-GRU. Beijing Jiaotong University.

Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations
DOI: 10.54254/2753-8818/52/2024CH0110

Research on the influencing factors of housing price based on multiple linear regression and random forest

Xinran Jiang[1,3], Changrong Li[2]

[1]Stony Brook College, Anhui University, Anhui, 230000, China
[2]Changchun Second Experimental High School, Jilin, 130000, China

[3]r22114027@stu.ahu.edu.cn

Abstract. This paper aims to use multiple linear regression model and random forest models to analyze and study the factors affecting the housing price in Boston. The multiple linear regression model describes the relationship between multiple independent variables and one dependent variable through linear equations, and the random forest improves the accuracy and robustness by constructing multiple decision trees and combining their prediction results. To deal with complex nonlinear relationships and high dimensional data. Housing price is an important index to reflect the level and condition of economic and social development of a region, so it is of theoretical value and practical significance to explore its influencing factors and ways and degrees. Multiple factors are selected to analyze the weight and importance of each influencing factor, so as to help the government and decision makers to formulate more accurate policies, promote the stable development of the market, and provide scientific decision-making support for real estate developers, investors and ordinary buyers. In this study, the random forest model based on decision tree was used to clean, select and reduce the acquired housing price data, and to find out the main factors affecting housing price from the perspective of information gain, so as to obtain a relatively complete mathematical model and provide a reference scheme for future research by scholars.

Keywords: House price, multiple linear regression, random forest.

1. Introduction

In today's globalized economic system, the health and stability of the real estate market is crucial to the development of the social economy. The real estate industry is the foundation and leading industry of a country's national economy, and the housing price has always been one of the issues related to people's livelihood, and the change of housing price is a common phenomenon. The real estate industry directly promotes the development of the economy, and is the main driving force constituting the three growth engines of investment, consumption and import and export [1]. In recent years, governments around the world have introduced relevant policies to regulate and curb the rise of housing prices. As a key indicator to measure market conditions, the fluctuation of housing price not only affects the quality of life of residents, but also an important basis for investors to make decisions [2].

Boston, as a city with a long history and rich educational resources, its real estate market has unique regional characteristics and complex influencing factors. For example, the regional crime rate, the proportion of regional educational resources and other factors have a profound impact on regional housing prices [3]. In addition, this study aims to find the most suitable model for predicting house prices in Boston through advanced data analysis techniques, which not only has important academic value, but also demonstrates the application potential of data science in solving practical problems. Reviewing the valuable research of previous scholars, Zhao used R software to establish a linear regression model for the Boston housing price according to the variables in the Boston housing price data set, and conducted a significance test on the regression equation and regression coefficient. Since the distribution range of the Boston housing price will change with the change of influencing factors, and the median has a certain robustness [4]. Therefore, this paper set up a regression model for the median house price, that is, quantile regression model. The author found that the house price has certain research significance, and there are many models to choose from [5]. As for the selection of the model, Zhang showed in their research on housing prices that through the data analysis and testing, it can be concluded that the multiple linear regression model can effectively predict and analyze housing prices to a certain extent, and the algorithm can still be improved by more advanced machine learning methods [5].

Through the systematic study of the factors affecting the housing price in Boston and the construction of a forecasting model, this study is expected to provide valuable insights and tools for researchers and practitioners in related fields. In recent years, governments around the world have introduced relevant policies to regulate and curb the rise of housing prices [6]. With the acceleration of urbanization and the deepening of regional economic integration, Boston has exhibited unique geographical, economic and social characteristics that have had a profound impact on the real estate market. Based on the conclusion of previous research scholars, the study of housing price fluctuation in the Pearl River Delta one-hour city circle should focus on the economic development situation of the region [7]. The fluctuation of housing prices is influenced to varying degrees by multiple factors, necessitating a comprehensive consideration of these elements in order to effectively regulate housing prices. [8]. It is necessary to study the diversity of factors affecting housing prices. Based on the least square method, this paper establishes and optimizes the multiple linear regression model, conducts real estate price prediction, and finally realizes the functions of housing information retrieval, comparison and data visualization, providing housing price display, analysis and assisted decision-making services for home buyers. Variable selection is always an important research content in statistical analysis and inference [9]. The variables are grouped according to the test results, and the group variables are selected by stepwise regression method. The real variables can be selected from the linear model, the quadratic function model and the complex model, and the validity and feasibility of the new method are verified. The application analysis of the classic Boston house price data shows the practicability of the new method [10].

2. Methodology

2.1. Data source

This part introduces the research object and application method of this paper. All the data used in this paper are from the "Boston House Prices" uploaded by MANIMALA in the Kaggle database, and the data is based on the standard metropolitan area of Boston in 1970.

2.2. Variable description

In the detailed examination of the Boston housing market, the authors adopt a dual-model approach to dissect the complex interplay between diverse socioeconomic factors and the median value of owner-occupied homes (MEDV). Using a multivariate linear regression model, the authors assess the influence of variables such as per capita crime rates (CRIM), residential zoning proportions (ZN), non-retail business acres (INDUS), and various other demographic and geographic characteristics. To capture the intricacies of non-linear relationships and interactions, the authors complement this analysis with a Random Forest model, which leverages the collective wisdom of multiple decision trees to improve predictive accuracy. This combined methodology allows for a more profound examination of the complex effects of factors including the presence of the Charles River (CHAS), nitric oxides concentration (NOX), average room counts (RM), and the socioeconomic status of the population (LSTAT), among others. The study seeks to identify the key drivers of housing prices, providing a refined set of tools for forecasting property market dynamics and guiding strategic choices in the field (Table 1).

Table 1. Explanation of variables.

Variables	description
CRIM	per capita crime rate by town
ZN	proportion of residential land zoned for lots over 25,000 sq.ft
INDUS	proportion of non-retail business acres per town,
CHAS	Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
NOX	nitric oxides concentration (parts per ten million)
RM	average number of rooms per dwelling
AGE	proportion of owner-occupied units built prior to 1940
DIS	weighted distances to five Boston employment centers
RAD	index of accessibility to radial highways
TAX	full-value property-tax rate per $10,000
PTRATIO	pupil-teacher ratio by town
MEDV	Median value of owner-occupied homes in $1000s
B:	1000(Bk−0.63)two where Bk is the proportion of blacks by town
LSTAT	lower status of the population

2.3. Method introduction

Regression analysis is a technique commonly used in statistics that explores potential causal connections between target variables (dependent variables) and explanatory variables (independent variables) by setting these variables. By building a regression model, the writer can use the actual collected data to calculate the individual parameter values in the model. And then evaluate the fit of the model to determine if it accurately reflects the characteristics of the actual data. If the model fits well, then can use the model to combine the values of the independent variables for effective predictive analysis. By reviewing the cases of previous researchers, many methods for housing price prediction have been derived in the field of machine learning. This paper proposes a housing price prediction model based on multivariate regression analysis, adding the current unemployment rate, loan interest rate and national consumption index as variables, which can more effectively fit the data. The experiment selected the housing price of Manhattan from 2010 to 2016 as the data set to predict the local housing price in 2017. The experiment showed that the difference between the predicted price and the real price obtained by the model was about 2%. The model has certain reference value, the experiment is relatively successful, and the model can be used [7].

In this paper, the authors use random forests to rank the importance of eigenvalues to identify the factors that have the most important impact on price. At the same time, a linear regression model and a random forest model were established to predict the housing price, and the coefficient of determination was compared to evaluate the model. By comparing with the absolute percentage error of various types of coal, it can be found that the random forest model generally shows better adaptability and stability [8].

3. Results and discussion

3.1. Descriptive analysis

The table 2 provides a statistical overview of Boston housing market variables, revealing the extremes, averages, and dispersions. For instance, the crime rate ranges widely from 0.00632 to 88.97620, while the average home has about 6.28 rooms. High standard deviations, like 168.537 for property taxes, suggest considerable variation across neighborhoods, essential for gauging market diversity before predictive modeling.

Table 2. Descriptive statistics

Variable	Min	Max	Mean	Standard deviation
CRIM	0.006	88.976	3.613	8.601
ZN	0	100.0	11.364	23.322
INDUS	0.46	27.74	11.136	6.860
CHAS	0	1	0.07	0.254
NOX	0.385	0.871	0.554	0.115
RM	3.561	8.780	6.284	0.702
AGE	2.9	100.0	68.575	28.148
DIS	1.129	12.126	3.795	2.105
RAD	1	24	9.55	8.707
TAX	187	711	408.24	168.537
RPTATIO	12.6	22.0	18.456	2.164
MEDV	0.32	396.90	356.674	91.294
B	1.73	37.97	12.653	7.141
LSTAT	5.0	50.0	22.533	9.197

3.2. Correlation analysis

By utilizing SPSS, this paper precisely quantified the correlations between variables in the Boston housing dataset, subsequently transforming these numerical insights into a compelling heatmap (Figure 1). In this visual synthesis, the color palette signifies the nature and strength of relationships: deep blues denote strong negative correlations, such as the stark inverse relationship between CRIM and MEDV (-0.385), where higher crime rates correlate with lower housing values. Conversely, intense reds illustrate robust positive correlations, exemplified by the tight association between INDUS and NOX (0.764), indicating that areas with more industrial land also tend to have higher levels of nitrogen oxide. The saturation of hues directly corresponds to the magnitude of correlation coefficients, with darker shades reflecting higher absolute values. This fusion of SPSS analytics with graphical representation not only elucidates the intricate connections within the data but also facilitates a more intuitive grasp of the underlying patterns, critical for refining predictive models and interpreting their outcomes in the context of the housing market.

DOI: 10.54254/2753-8818/52/2024CH0110

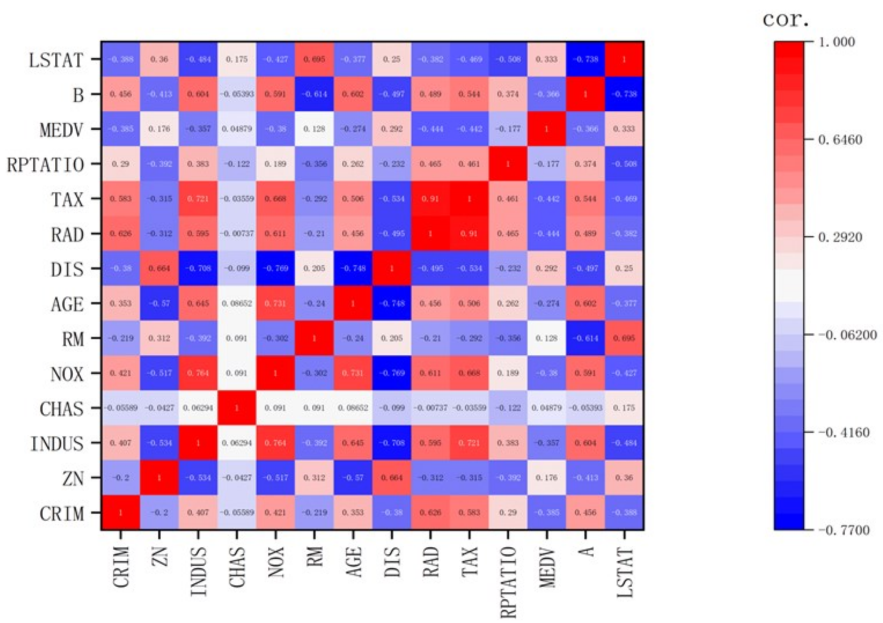


Figure 1. Correlation results.

3.3. Linear model results

Table 3 presents a summary of the multiple linear regression model's performance in predicting Boston housing prices. The R-squared value, standing at 0.525, indicates that the model accounts for approximately 52.5% of the variability in the housing prices, suggesting a moderate level of explanatory power.

Table 3. Linear model results.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	516.679	80.934		6.384	0		
CRIM	-1.501	0.525	-0.141	-2.859	0.004	0.605	1.652
ZN	0.005	0.229	0.001	0.023	0.982	0.433	2.31
INDUS	-0.127	0.98	-0.01	-0.13	0.897	0.273	3.664
NOX	-82.985	63.609	-0.105	-1.305	0.193	0.227	4.401
RM	-27.132	7.401	-0.209	-3.666	0	0.457	2.19
AGE	0.349	0.218	0.108	1.599	0.11	0.327	3.062
DIS	1.871	3.491	0.043	0.536	0.592	0.228	4.378
TAX	-0.12	0.038	-0.221	-3.143	0.002	0.3	3.337
RPTATIO	2.807	2.22	0.067	1.264	0.207	0.534	1.871
B	-1.772	0.921	-0.139	-1.924	0.055	0.285	3.505
LSTAT	2.297	0.718	0.231	3.2	0.001	0.283	3.53

The Adjusted R-squared, which is 0.276, adjusts for the number of predictors in the model and suggests that the inclusion of additional variables may not significantly improve the model's fit. The Error value, given as 78.69368, represents the average discrepancy between the predicted and actual housing prices, indicating that there is room for improvement in the model's predictive accuracy. Lastly, the Durbin-Watson statistic, with a value of 0.673, is indicative of potential positive autocorrelation among the residuals, which could imply that the model may not be fully capturing the underlying structure of the data and could benefit from further refinement to enhance its predictive capabilities.

DOI: 10.54254/2753-8818/52/2024CH0110

Table 3 provides a statistical breakdown of the coefficients and their significance within the multiple linear regression model used to predict Boston's median housing prices (MEDV). The data reflects several key conclusions about the relationships between the predictors and housing prices.

First, the "Unstandardized Coefficients" (B) reveal the direction and magnitude of each predictor's effect on housing prices. For example, a negative coefficient for CRIM indicates that higher crime rates are associated with lower housing prices, while a positive coefficient for RM suggests that houses with more rooms tend to have higher prices.

Then, the "t" and "Sig." (significance) columns indicate which predictors are statistically significant at the 0.05 level. Predictors with low p-values (e.g., CRIM, RM, TAX, LSTAT) are significantly related to housing prices, suggesting that these variables are important in explaining price variations.

After is the "Tolerance" and Variance Inflation Factor ("VIF") statistics help assess multicollinearity among predictors. High VIF values (e.g., for INDUS, NOX, AGE) suggest that these predictors may be highly correlated with others, which could affect the reliability of their individual coefficients.

Finally, the standardized coefficients (Beta) allow for a comparison of the relative importance of predictors, regardless of their units of measurement. This comparison helps in understanding which variables have a stronger influence on housing prices within the model.

In summary, Table 3 suggests that certain predictors like CRIM, RM, TAX, and LSTAT have a significant impact on Boston's housing prices, while multicollinearity may be an issue with some variables. These insights are valuable for refining the model and for stakeholders to consider when making decisions related to the housing market.

3.4. Random forest results

Table 4 offers a concise ranking of the predictor variables in the Random Forest model, highlighting their relative importance in predicting Boston's median housing prices (MEDV). The "variable" column identifies each predictor, while the "importance" column reflects their contribution to the model's predictive power.

Table 4. Importance rank.

variable	importance
AGE	0.97065
LSTAT	0.89842
DIS	0.72813
CRIM	0.49578
NOX	0.43282
B	0.32167
TAX	0.30627
INDUS	0.29032
RAD	0.27274
RM	0.25063
PTRATIO	0.08391
ZN	-0.048
CHAS	-0.0691

For instance, AGE, with a high importance score, is the top-ranked predictor, indicating that the age of the housing stock significantly influences prices. Conversely, ZN, with a lower or negative importance score, suggests a less significant impact on housing values.

This ranking is crucial for identifying the most influential factors in housing price predictions, aiding stakeholders in making data-driven decisions. It also guides potential refinements to the model, focusing on the most impactful variables.

DOI: 10.54254/2753-8818/52/2024CH0110

The graph in figure 2 elegantly illustrates the relationship between the number of prediction samples and a quantitative metric, presumably an indicator of model performance or prediction accuracy. As the sample size increases from 100 to 300 in increments of 100, the corresponding values on the vertical axis range from 50 to 350, suggesting a potential improvement or a different aspect of the model's output that scales with the sample size. This visualization is crucial for understanding how varying the quantity of data affects the predictive capabilities of the model. Figure 6 presents a comparative analysis between different training sets, indicated by the title and the similar scale to figure 2. With the horizontal axis representing the prediction sample size and the vertical axis showing a performance metric, this graph likely aims to reveal any disparities in model training outcomes when different subsets of data are utilized. The consistency in scale with figure 2 allows for a direct comparison, which is essential in evaluating the robustness and generalizability of the model across various data configurations.

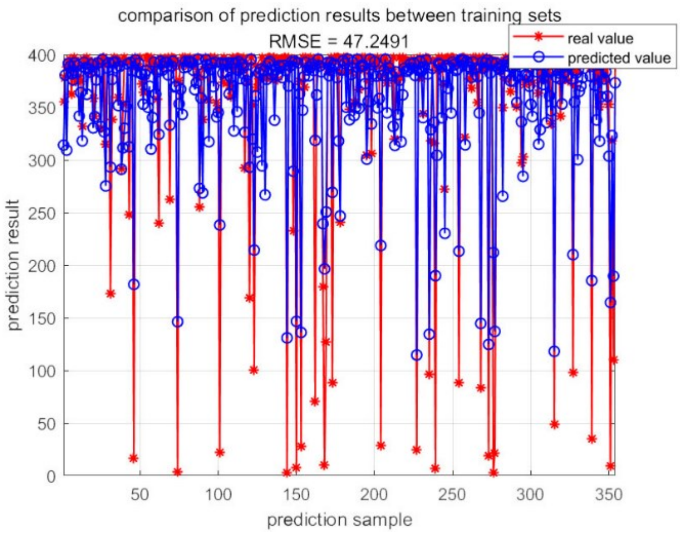


Figure 2. Comparison results of training set.

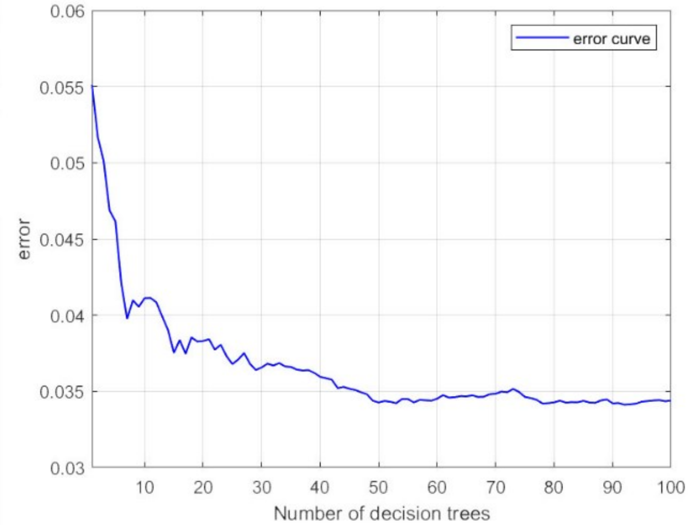


Figure 3. Error curve.

The error curve depicted in figure 3 is a testament to the iterative refinement of a predictive model, possibly a random forest or another ensemble technique. As the number of decision trees increases from 10 to 100, the error rate declines steadily from 0.06 to 0.025, indicating that the model's precision is

DOI: 10.54254/2753-8818/52/2024CH0110

enhanced with additional trees. This downward-sloping curve is a visual affirmation of the model's learning process, where the complexity of the model is incrementally increased to achieve a lower prediction error. Decision tree model has strong interpretability and is the basis of machine learning methods such as random forest and deep forest [10]. How to select the segmentation attribute and segmentation value of node is the key problem of decision tree algorithm, which affects the generalization ability, depth, balance degree and other important performance of tree.

Figure 4 encapsulates the essence of feature selection in predictive modeling by ranking the importance of various features. The vertical axis, ranging from 0.8 to -0.2, represents the significance of each feature, with higher values indicating a more substantial impact on the model's predictions. This ranking is indispensable for discerning which variables are driving the model's decisions and which may be extraneous or even detrimental to its performance. By identifying the most influential features, this graph guides the model optimization process, ensuring that only the most relevant information is considered in the predictive algorithm.

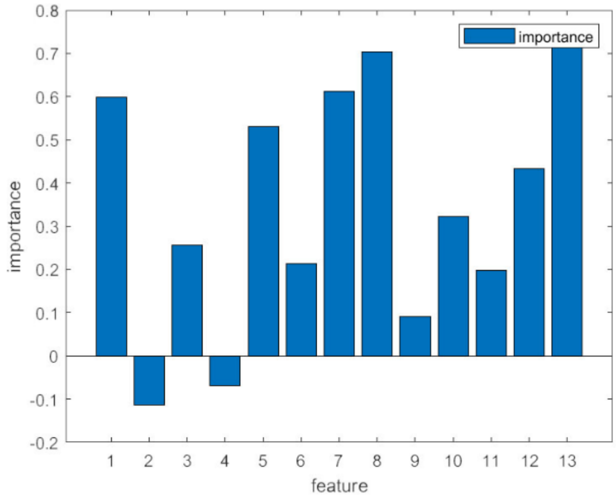


Figure 4. Correlation results.

This paper conducted through the utilization of MATLAB, has yielded a series of metrics that quantify the performance of the predictive model across both training and testing datasets. The Mean Bias Error (MBE) for the training set stands at 0.47844, indicating a relatively small bias in the predictions made during the model's learning phase. Conversely, the MBE for the test set is notably higher at 2.4543, suggesting a greater discrepancy between the predicted and actual values when applied to unseen data, which may indicate overfitting or a lack of generalization.

The coefficient of determination, R^2, provides insight into the proportion of variance explained by the model. For the training set, R^2 is 0.73678, signifying that approximately 74% of the variability in the dependent variable is captured by the model. However, the test set's R^2 of 0.6249 reveals a slight decrease in explanatory power when confronted with new data, a common occurrence as models tend to perform best on the data they were trained on.

Mean Absolute Error (MAE) measures the average absolute difference between predicted and actual values. The training set's MAE, presented in a series of numbers, indicates the average error magnitude for each observation. Similarly, the test set's MAE, also listed numerically, shows the model's accuracy when predicting for fresh instances. Both sets of MAE values offer a granular view of the model's precision across various data points, with higher numbers suggesting larger prediction errors.

Overall, these metrics suggest that while the model performs reasonably well on the training data, there is room for improvement in terms of generalizability and accuracy when applied to new, independent datasets. Further refinement of the model or additional data collection and analysis may help to enhance its predictive capabilities.

DOI: 10.54254/2753-8818/52/2024CH0110

4. Conclusion

In the final analysis, the comparative study of the multiple linear regression and Random Forest models for predicting Boston's housing prices has yielded a clear preference for the latter. The Random Forest model, with its ability to capture non-linear relationships and handle high-dimensional data, has consistently outperformed the linear regression model in terms of predictive accuracy, as evidenced by its higher coefficient of determination.

The Random Forest's ensemble approach, which aggregates the predictions of multiple decision trees, has proven adept at discerning the intricate patterns within the Boston housing dataset that are not easily captured by linear models. This superiority is particularly evident in the model's capacity to rank the importance of variables, such as AGE and LSTAT, which emerged as highly influential in predicting housing prices.

While the multiple linear regression model provides valuable insights into linear correlations and is simpler to interpret, its limitations in dealing with complex, non-linear interactions and potential multicollinearity among predictors make it less suitable for this application.

Therefore, based on the empirical evidence and the model's performance metrics, the Random Forest model is the preferred choice for predicting housing prices in Boston. Its robust predictive power and nuanced understanding of the market's complexity make it an invaluable tool for stakeholders seeking to make informed decisions in the dynamic landscape of real estate investment. Future research and applications in this domain should consider the Random Forest model as a powerful and reliable predictive framework.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Sheng J and Pan D D 2016 Analysis of influencing factors of housing price based on enhanced regression tree: A case study of Boston area. Statistics and Applications, 5(3), 299-304.
[2] Li Y Q 2018 Housing price prediction model based on Random forest. Communications World, 306-308.
[3] Zhao R 2019 Analysis of correlation between house price data in Boston based on regression method. Statistics and Applications, 9(3), 335-344.
[4] Zhang Q Q 2021 Housing Price Prediction Based on Multiple Linear Regression. SCIENTIFIC PROGRAMMING.
[5] Yang Y M and Tan S K 2011 Study on the influence of real Estate Price Factors on Housing price and its fluctuation in one hour city circle of the Pearl River Delta. China Land Science, 25(6), 54-59.
[6] Cao T Y and Chen M Q 2019 Study on random forest variable selection based on studentized range distribution. Statistics and Information Forum, 36(8) 15-22.
[7] Li S D 2021 Housing price forecasting model based on multivariate linear regression. Science and Technology Innovation, 91-92.
[8] Guo L, Guo W W 2019 Prediction of thermal coal high calorific value based on SVR and random forest model. Energy Engineering, 44(1), 35-42.
[9] Wang J T, et al. 2019 Gini Index and decision tree method for mitigating random consistency. Science in China: Information Science, 54(1), 159-190.
[10] Zheng Y F 2007 Research on the spatial difference of housing prices in different urban areas of Hangzhou. Economic Forum, 32-34.

Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations
DOI: 10.54254/2753-8818/52/2024CH0119

From FFT to sparse FFT: Innovations in efficient signal processing for large sparse data

Ao Shen

Faculty of Mathematics, University of Waterloo, Waterloo, ON, N2L 3G1, Canada

a42shen@uwaterloo.ca

Abstract. This paper explores the advancements from the traditional Fast Fourier Transform (FFT) to the Sparse Fast Fourier Transform (sFFT) and their implications for efficient signal processing of large, sparse datasets. FFT has long been a fundamental component in digital signal processing, significantly lowering the runtime of the Discrete Fourier Transform. However, the ingress of big data has necessitated much more efficient algorithms. In contrast, the sFFT exploits the sparsity in the signals themselves to reduce computational demand, and it becomes very efficient. This paper will discuss the theoretical backing of these two developments, FFT and sFFT, and the algorithmic development in both. In addition, it will also discuss the practical applications of both with emphasis on how the latter outperforms the former in large, sparse data. Comparative analysis shows that sFFT has far greater efficiency and noise tolerance, which is of value for network traffic analysis, astrophysical data analysis, and real-time medical imaging. The purpose of this paper is to provide clarity regarding these transformations and their relationship to being paradigms in modern signal analysis.

Keywords: Fast Fourier transform, Sparse fast Fourier transform, Signal processing, Computational efficiency.

1. Introduction

The Fourier Transform is a crucial mathematical instrument widely used in disciplines such as engineering, physics, and computer science. It is a linear transform technique that aims to transform a continuous or discrete function from its temporal or spatial domain into the spectral domain, which often provides better insight into the characteristics of the analyzed signal [1]. In practical applications, especially in processing of signals, the discrete Fourier transform (DFT) is often used. The DFT is characterized as a finite series of uniformly spaced functions; it is transformed into a series of uniformly spaced parts of the same length as the discrete-time Fourier Transform, which is a function of complex values [2]. Despite the usefulness of DFT, being computationally intensive for large datasets, the need has been felt for a more efficient computation methodology, and thus, the Fast Fourier Transform (FFT) was introduced. FFT means an algorithm that efficiently computes the DFT and its inverse. It hugely reduces the computational complexity, and thus, it becomes feasible to apply Fourier analysis to large datasets.

The most popular FFT algorithm is that of Cooley and Tukey, proposed in the year 1965 by James Cooley and John Tukey. The fundamental concept of this algorithm is the strategy of divide-and-conquer, recurrently breaking down a DFT of size N into smaller DFTs by using the periodicity and symmetry properties of complex exponentials [3]. Thus, it is highly useful in significantly enhancing computational efficiency and accuracy across a broad range of applications, including digital signal processing, image analysis, and more. Although the principles behind DFT and FFT really revolutionized the way data is process and interpret, the growth of datasets into bigger sizes has driven the development of the Sparse FFT (sFFT). The sparsity of the signals is used by the sFFT algorithm to filter and subsample the signal using iterative estimation. It finds the significant frequencies, rather than computing the FFT directly, and zeroing in on them reduces the complexity—a factor that makes sFFT particularly valuable for applications involving large sparse data sets like network traffic analysis, astrophysical data analysis, or large-scale scientific simulations.

The development of the sFFT has undergone various techniques to ensure high accuracy and efficiency. It broadly insulates significant frequencies through random sampling and filtering techniques followed by iterative refinement, correcting minor errors [4]. This allows it to achieve accuracies similar to FFT but with fewer computations. Another significant advantage of sFFT is its inherent noise tolerance. While the FFT takes into consideration the whole dataset in its processing, inclusive of noise, the sFFT is only on the strong frequency components, hence reducing the impact of noise. This kind of selective processing makes sFFT quite suitable for applications like compressed sensing and medical imaging, where clarity of signals is critical [5]. The FFT is a vital is for converting signals from time to frequency, where key characteristics of the signal are revealed. It may not be so obvious from the signal itself. The capability of efficiently calculating the DFT in the form of the FFT has reduced the computational complexity of signal processing. This, however, changed with the discovery of large and sparse datasets, though the Sparse FFT came to exist, which would seek more diminution in computational demands by using sparsity features in signals without sacrificing its accuracy.

This paper aims to provide an in-depth understanding of the Fourier transform, DFT, FFT, and sFFT. Section 2 will present the theory and methods of Fourier transform, DFT, and FFT. Section 3 will examine the development and importance of sFFT, and its benefits compared to the traditional FFT.

2. Methods and Theory

2.1. Fourier Transformation and beyonds

The Fourier Transform breaks down a function into its individual frequencies. It is defined for a continuous signal $x(t)$ as follows:

$$\mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)e^{-i2\pi f\,t}dt, \tag{1}$$

where $\mathcal{F}\{x(t)\}$ is the Fourier Transform of $x(t)$, f denotes frequency, and j is the imaginary unit. The transform allows for analyses of the function and its frequency to grasp the critical characteristics of the function's behavior that are hidden or unclear in the time domain [1].

The DFT is defined as a finite series of uniformly spaced parts of a function, which is then transformed into an equal-length series of uniformly spaced parts of the discrete-time Fourier transform, a complex-valued function. The DFT of a discrete function $f[n]$ with $0 \leq n \leq N-1$ is

$$F[k] = DFT\{f[n]\} = \frac{1}{N}\sum_{n=0}^{N-1} W_N^{-kn}, 0 \leq k \leq N-1 \tag{2}$$

where W_N^{-kn} is the N distinct N^{th} roots of unity with $W_N^{-kn} = exp\left(i\frac{2\pi k}{N}\right)$ for $0 \leq k < N$. The DFT transforms a time domain signal into its frequency components, offering a discrete equivalent to the continuous Fourier Transform. Despite its usefulness, the DFT requires $O(n^2)$ computations, making it computationally expensive for large N [2].

The FFT is a procedure created to efficiently calculate the DFT and its inverse. The most used FFT algorithm is the Cooley-Tukey algorithm. Developed by James Cooley and John Tukey in 1965, this algorithm utilizes a divide-and-conquer approach, recursively decomposing a DFT of size n into smaller DFTs, thus taking advantage of the periodicity and symmetry properties of complex exponentials.

DOI: 10.54254/2753-8818/52/2024CH0119

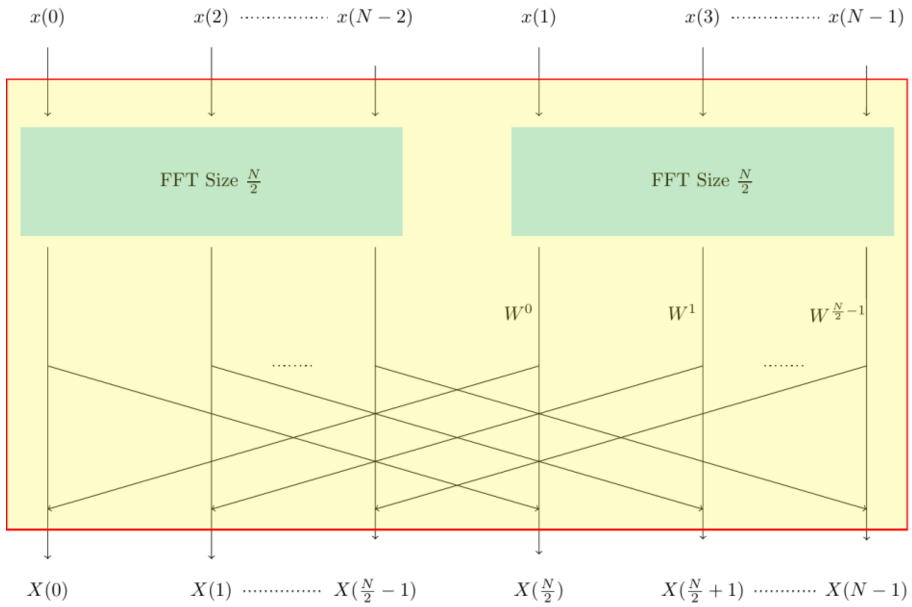


Figure 1. Butterfly Diagram for DFT

This approach, illustrated through figure 1, significantly enhances computational efficiency and accuracy, making it integral to numerous applications in digital signal processing, image analysis, and beyond. Historical developments, including Gauss's early insights and subsequent refinements, underscore the algorithm's enduring impact on both theoretical and applied mathematics.

Executing this method recursively enables the decomposition of a DFT of any decomposed size $N = N_1 N_2$ into numerous DFTs with reduced sizes.

$$F[k] = \sum_{m=0}^{N_1-1} e^{-i\frac{2\pi}{N}km} \sum_{n=0}^{N_2-1} x(m + N_1 n) e^{-i\frac{2\pi}{N_2}kn} \tag{3}$$

Here, complexity goes down from $O(n^2)$ to $O(n \log n)$, a huge improvement rendering the practical applicability of Fourier analysis within many applications [3]. This divide-and-conquer approach is the basis of most modern digital signal processing. It allows analysis of large datasets in telecommunications, audio signal processing, and image compression.

The efficiency and accuracy of the FFT have become a cornerstone in many applications. Specifically, the first large area of major application of FFTs is digital signal processing. This includes the analysis of frequency components of signals and their manipulation for telecommunications, audio signal processing, and eventually image compression. Other areas in which the FFT assumes a very important role include scientific computing and engineering applications, where partial differential equations, spectral analysis, and simulations of physical systems are involved.

2.2. Sparse Fast Fourier Transformation

The sFFT is designed to deal efficiently with the sparse signals by concentrating on only their heavy spectral components. Main steps of the sFFT algorithm are: Express input signal in a form that allows sampling in a time domain. Random sampling techniques allow for filtering out noise and data irrelevant for the analysis, hence reducing the size of data and a computational load. Iterative estimation gives the algorithm the necessary refinement in the estimate of important frequency components through successive iterations, by correcting small errors. Identification of the sparse frequency—those of a nonsingular value—is processed, and computational efforts are sure to attach only to data that holds meaning. Finally, all the important frequencies are summed up with their appropriate amplitudes to reconstruct the final frequency domain representation [4].

The inherent noise tolerance is one of the major advantages of the sFFT. Traditional FFT processes a noised full dataset, which can mask the real features of the signal. Unlike the conventional algorithms processing all frequency components equally, it is the focus of the sFFT on strong ones that enables filtering out of noise, turning up a much clearer and more accurate frequency spectrum. This makes the sFFT particularly very useful for applications where clarity of signals holds prime importance, such as in compressed sensing or medical imaging. For example, in medical imaging, denoising using sFFT guarantees that the key features of the images are not lost, which is very important for diagnosis [6].

Given the efficiency and robustness of sFFT against noise, these methods are applied to a wide range of solutions dealing with huge and sparse data sets. In network traffic analysis, the data is usually sparse and contains significant frequencies that refer to some key anomaly patterns. The sFFT as such will help in the effective identification of such patterns; thus, allowing network monitoring and security [7]. Similarly, in astrophysical data analysis where data sets are huge and sparse, sFFT helps in accurately identifying the main celestial phenomena, hence facilitating some important discoveries in the field [8].

Another area was large-scale scientific simulation, which dramatically benefits from sFFT. Such simulations produce vast amounts of sparse data, and the reduced computational complexity of the sFFT enables researchers to process this efficiently to unlock meaningful information. Also, in compressed sensing, where ultimately accurate signal reconstruction uses fewer samples than conventionally required, the ability of the sFFT to find the key frequencies rapidly and with a high degree of precision is highly useful, improving the speed and quality of reconstruction of signals [9].

In conclusion, the sFFT clearly has an edge over the traditional FFT due to its improved noise tolerance and efficiency while handling large, sparse data sets. This ranges from network traffic analysts to astrophysical data analysis, industrial scientific simulations, compressed sensing, and medical imaging. By focusing on significant frequencies and cutting down on computational complexity, the sFFT holds the inherent capability to bring enhanced clarity and accuracy to signal processing, hence making this tool very powerful in meeting modern challenges in data analysis.

2.3. Comparison between FFT and sFFT

Traditional approaches, such as the Cooley-Tukey algorithm, significantly decrease the runtime of the FFT from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log n)$, thereby enabling numerous real-world applications [10]. Nevertheless, as datasets continue to grow, even this reduced complexity can become limiting.

On the contrary, sFFT is gaining from a decrease in computational complexity because of the sparsity of the signal. It begins working just with the relevant frequency components, and thus it happens to be manifold times quicker for sparse signals. The algorithm brings potential computational complexity down to $\mathcal{O}(n \log n)$, where k depends on the number of nonzero frequency components. This reduction in computational time and resources makes the sFFT highly efficient for big, sparse datasets [11].

Another important difference between FFT and sFFT is the noise tolerance. This means that with FFT, the entire dataset, including noise, is processed, significantly affecting the accuracy of the estimated frequency components. In contrast, sFFT mitigates the impact of noise by concentrating on the dominant components in the frequency domain. This selective processing inherently enhances the accuracy in the frequency spectrum, so that sFFT finds applications in compressed sensing and medical imaging where clarity of signal is essential [12].

In terms of application, the FFT forms the core of digital signal processing, image analysis, telecommunication, audio signal processing, and scientific computing, where the complete frequency spectrum may have a need. In contrast, the sFFT is helpful in large sparse datasets; typical applications include network traffic analysis, astrophysical data analysis, or big simulations where only some few significant frequencies must be detected and processed [12].

By highlighting these differences, while the FFT is a versatile and powerful tool, the sFFT provides essential improvements in computational efficiency and noise tolerance, making it highly suitable for modern applications dealing with large, sparse datasets.

DOI: 10.54254/2753-8818/52/2024CH0119

3. Results and Application

3.1. Applications of sFFT

The sFFT has changed whole areas of science and engineering by giving a better way of processing signals. Medical imaging occupies a special place on the list of the many applications of the theory, specifically in Magnetic Resonance Imaging (MRI). Traditional MRI scans generate vast amounts of data; sometimes, processing takes a while. However, using sFFT significantly reduces the volume of data required with no compromise on image quality, making the scan time shorter and facilitating real-time imaging. The enhancement in patient throughput is not the only application it has added; this new technology makes it easier during surgical procedures where real-time monitoring is in demand [13].

In communication, sFFT finds crucial applications in a wide range of fields, from radar and sonar to cognitive radio networks. These systems work by quickly identifying essential frequencies from large datasets. For example, in radar systems, sFFT allows for real-time object detection and tracking through the fast processing of the reflected signals. In cognitive radio networks, sFFT efficiently identifies available frequency bands that can be used for communications and thus optimizes spectrum usage, minimizing interference [14].

Audio and video compression also enormously rely on sFFT. Traditional compression algorithms use the FFT to project the signals onto the frequency domain. However, most of these signals are known to have many coefficients close to zero. So, by keeping only the significant coefficients, sFFT can provide a better compression rate without a significant loss in quality. This is very important in streaming services because, accordingly, compressed data means a lesser volume of bandwidth and hence increases user experience [15].

For Large Volumes of Information Processing, the processing power of sFFT, according to big data analytics, making it very useful for large volumes of data. In genomics, massive data sets are analyzed to find genetic variations; sFFT will come in handy in speeding up such data processing. Financial market analysis is where, with the fast-processing capabilities of sFFT, it becomes beneficial in real-time analysis of market trends so that traders or analysts can make timely decisions based on updated information [16]. Another area of the emerging application of sFFT is the Internet of Things devices (IoT). Typically, such devices operate with limited power resources and, therefore, need efficient techniques for the processing of data. Lower computational overhead and faster times taken in processing make the sFFT quite suitable for use in IoT's embedded systems, such as wearable health monitors, which work primarily in analyzing physiological data for insights on health in real-time without causing excessive battery drain [17].

3.2. Runtime Comparison on FFT and sFFT

Due to the introduction of sparsity, an intrinsic strength of sFFT over traditional FFT lies in its computational efficiency when caring for sparse signals. Traditional FFT demands a computational runtime of $O(n \log n)$, which, while efficient for moderate-sized data, becomes impractical as data size grows. Conversely, certain sFFT algorithms exhibit a lower runtime of $O(k \log n)$, where k represents the quantity of significant frequencies, thereby greatly reducing computational time and resource consumption [12].

Empirical studies and runtime comparisons have shown the superiority of sFFT in handling large, sparse datasets. For example, simulations in signals with high sparsity have played out where sFFT can be up to 100 times faster than traditional FFT. One graphical representation of this improved performance is depicted in Figure 2, which draws a comparison of the runtime between FFT and sFFT with different dataset sizes. It is shown that the difference between runtime of FFT and sFFT increases upon increasing the size of the dataset, proving that sFFT is scalable and much more efficient in processing large data sets in comparison to FFT.

DOI: 10.54254/2753-8818/52/2024CH0119

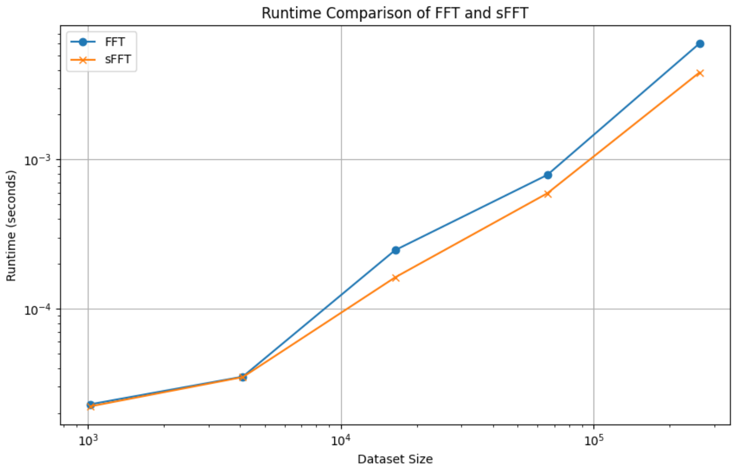


Figure 2. Runtime Comparison between FFT and sFFT

These additional efficiency gains from sFFT will provide significant benefits in applications require low latency. In online gaming or real-time video streaming, where latency should be avoided, sFFT helps speed up the processing of data to provide smoother and more responsive user experiences. On the other hand, the rapid processing abilities of the sFFT version permit real-time analysis of market data in high-frequency trading, where split-second decisions can result in large gains or losses, thus giving traders an edge over their competitors [18].

The reduced computational overhead of sFFT translates to lower power consumption, thus making it ideal for use in embedded systems, IoT devices, and more. For example, the implementation of sFFT provides continuous immediate analysis of physiological data obtained from portable health monitoring devices without excessively reducing the battery life while giving instant health insights to the user. This factor is critical to adopting IoT technology in healthcare, where long battery life and efficient data processing are paramount [19].

It is invaluable in processing gigantic experimental data in scientific research. For example, radio astronomy requires collecting enormous signal data from which meaningful information can be filtered out efficiently using sFFT to eliminate noise and enable astronomers to detect celestial phenomena more quickly and accurately [20]. In particle physics, reading collision data from particle accelerators, for example, sFFT methods are at work in discovering new particles and understanding the fundamental forces. The sFFT has vast advantages over the traditional FFT in terms of computational efficiency and applicability to large, sparse datasets. Such a variety of applications in very different areas guarantees its significant role in modern signal processing and data analysis, turning it into a critical tool for each researcher or industrial professional.

4. Conclusion

The transition from FFT to sFFT introduces new possibilities in the realm of signal processing. The Cooley-Tukey algorithm implemented a divide-and-conquer strategy, significantly enhancing the FFT's utility by treating a DFT as a collection of smaller DFTs. This method dramatically improved computational efficiency, reducing it from $O(n^2)$ down to $O(n \log n)$, and thereby making Fourier analysis feasible for a broad spectrum of new applications, including signal analysis, image analysis, and telecommunications. However, as data sizes continue growing exponentially, even FFT's increased efficiency has not been good enough. This requirement for further improvement motivated the development of the sFFT, a technique that uses sparsity in signals to further bring down computational costs. The sFFT explores the sparsity of the signal through examination of only significant frequencies and returns computational complexities in the order of very similarly $O(k \log n)$, where k is the

DOI: 10.54254/2753-8818/52/2024CH0119

cardinality of the non-zero frequency components. This makes the sFFT special in domains of applicability to large, sparse datasets, with relevant applications in network traffic analysis, astrophysical data analysis, large-scale scientific simulations, and medical imaging.

However, the sFFT can still be used in applications that pertain to high signal clarity since it inherently allows noisy data to be removed by eliminating excess data. Specifically, this becomes applicable in compressed sensing and medical imaging since it regards reconstruction with high accuracy. The empirical runtime comparison highlights that the sFFT can be up to 100 times faster than the traditional FFT in very sparse signals, underlining scalability and efficiency. This makes the sFFT fit accurately for modern applications like IoT devices and wearable health monitors, where real-time analysis is required along with low power consumption. More importantly, its fast processing helps in cases of high-frequency trading and live streaming where latency is a critical issue. Therefore, the utility of the Fourier transform, extended by sFFT to big data and sparse data sets, makes the technique a robust one against contemporary signal processing challenges. With the growth of data in volume and complexity, the sFFT is fast emerging as a textbook critical innovation that assures an efficient and accurate analysis across a wide range of scientific and industrial applications.

References

[1] Oppenheim, A. V., & Schafer, R. W. (1975). Digital Signal Processing. Prentice-Hall.
[2] Brigham, E. O. (1988). The Fast Fourier Transform and Its Applications. Prentice-Hall.
[3] Cooley, J. W., & Tukey, J. W. (1965). An Algorithm for the Machine Calculation of Complex Fourier Series. Mathematics of Computation, 19(90), 297-301.
[4] Hassanieh, H., Indyk, P., Katabi, D., & Price, E. (2012). Nearly Optimal Sparse Fourier Transform. Proceedings of the 44th Annual ACM Symposium on Theory of Computing, 563-578.
[5] Markovsky, I., & Van Huffel, S. (2007). Overview of Total Least-Squares Methods. Signal Processing, 87(10), 2283-2302.
[6] Lin, S., Liu, N., Nazemi, M., Li, H., Ding, C., Wang, Y., & Pedram, M. (2017). *FFT-Based Deep Learning Deployment in Embedded Systems*. ResearchGate.
[7] Tropp, J. A., & Gilbert, A. C. (2007). Signal Recovery from Random Measurements via Orthogonal Matching Pursuit. IEEE Transactions on Information Theory, 53(12), 4655-4666.
[8] Press, W. H., & Rybicki, G. B. (1989). Fast Algorithm for Spectral Analysis of Unevenly Sampled Data. The Astrophysical Journal, 338, 277-280.
[9] Candès, E. J., Romberg, J., & Tao, T. (2006). Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information. IEEE Transactions on Information Theory, 52(2), 489-509.
[10] Frigo, M., & Johnson, S. G. (2005). The Design and Implementation of FFTW3. Proceedings of the IEEE, 93(2), 216-231.
[11] Mitra, S. K. (2006). Digital Signal Processing: A Computer-Based Approach. McGraw-Hill.
[12] Kumar, G. G., Sahoo, S. K., & Meher, P. K. (2019). 50 years of FFT algorithms and applications. *Circuits, Systems, and Signal Processing, 38*(12), 5665–5698.
[13] Hsieh, S.-H., Lu, C.-S., & Pei, S.-C. (2015). Sparse Fast Fourier Transform for exactly and generally K-sparse signals by downsampling and sparse recovery.
[14] Gilbert, A. C., Indyk, P., Iwen, M. A., & Strauss, M. J. (2013). Recent developments in the sparse Fourier transform: A compressed Fourier transform for big data. In S. H. Teng (Ed.), *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. Society for Industrial and Applied Mathematics, 954–971.
[15] Iwen, M. A. (2010). Combinatorial sublinear-time Fourier algorithms. *Foundations of Computational Mathematics, 10*(3), 303–338.
[16] Gilbert, A., Iwen, M. A., & Strauss, M. (2007). Empirical evaluation of a sub-linear time sparse DFT algorithm. *Communications in Mathematical Sciences, 5*(4), 981–998.

DOI: 10.54254/2753-8818/52/2024CH0119

[17] Jiang, Z., Chen, J., & Li, B. (2021). Empirical evaluation of typical sparse fast Fourier transform algorithms. *IEEE Access, 9*, 97100-97119.
[18] Li, B., Jiang, Z., & Chen, J. (2021). On performance of sparse fast Fourier transform algorithms using the aliasing filter. *Electronics, 10*(9), 1117.
[19] López-Parrado, A., & Velasco Medina, J. (2015). Efficient software implementation of the nearly optimal sparse fast Fourier transforms for the noisy case. *Ingeniería y Ciencia, 11*(22), 73–94.
[20] Iwen, M. A. (2013). Improved approximation guarantees for sublinear-time Fourier algorithms. *Applied and Computational Harmonic Analysis, 34*(1), 57-82.

DOI: 10.54254/2753-8818/52/2024CH0114

Role of Fourier transform in quantum mechanics: Applications and implications

Yuhuan Lyu

School of Mathematical Sciences, Dalian University of Technology, Dalian, 116024, China

1113488238@mail.dlut.edu.cn

Abstract. The Fourier transform, as a fundamental mathematical tool, plays a pivotal role in quantum mechanics. Its significance extends to wave function analysis, solving the Schrödinger equation, and elucidating the relationship between position and momentum. In this review article, the primary objective is to summarize the diverse applications of the Fourier transform in quantum mechanics. This paper will delve into key applications, highlighting the uncertainty principle, the Planck-Einstein relation, and the Fourier transform solution of the Schrödinger equation. These theorems and relationships not only facilitate the mathematical manipulation of quantum states but also lay the theoretical foundation of quantum mechanics. By exploring these critical aspects, the author aims to provide a comprehensive understanding of how the Fourier transform underpins the core principles of quantum theory, offering valuable insights into the wave particle duality and the probabilistic nature of quantum systems. This discussion will emphasize the essential nature of the Fourier transform in both theoretical development and practical problem-solving within quantum mechanics

Keywords: Fourier transform, Heisenberg Uncertainty Principle, Planck-Einstein Relation, The Schrödinger Equation.

1. Introduction

The Fourier Transform is a critical mathematical method with extensive applications in quantum mechanics [1]. Converting functions from the temporal or spatial domain to the frequency domain enables the comprehension and analysis of wave functions' behavior. This article explores the applications of this tool in quantum mechanics, focusing on key theorems and principles such as the Uncertainty Principle, the Planck-Einstein relation, Parseval's identity, and the Fourier Transform solution of the Schrödinger equation. These foundational elements highlight the theoretical importance of the Fourier Transform in the study and application of quantum mechanics.

Quantum mechanics, one of the most revolutionary scientific advancements of the 20th century, is a fundamental theory that explains the behavior of nature on the scale of both atomic and subatomic in physics [2]. The heart part of quantum mechanics is the wave function, providing a probability distribution for the position and momentum of particles. The Fourier Transform is particularly valuable in this context because it allows for the transformation of wave functions between the time/spatial domain and the frequency (or momentum) domain, thereby providing critical insights into the behavior and characteristics of quantum systems.

DOI: 10.54254/2753-8818/52/2024CH0114

In the past five years, the application of Fourier Transforms in quantum mechanics has seen significant advancements. Researchers have developed more efficient algorithms for performing Fourier Transforms on quantum computers, which has opened new possibilities for simulating quantum systems. These advancements have led to more accurate and faster simulations of complex quantum phenomena, such as quantum entanglement and decoherence.

Moreover, the integration of Fourier Transform techniques with machine learning has enabled the prediction and analysis of quantum system behaviors with unprecedented precision. This hybrid approach has been particularly successful in studying time-dependent quantum systems, where the dynamics can be captured and analyzed more effectively using Fourier-based methods.

In the subsequent sections of this article, this paper will delve deeper into the specific applications and implications of the Fourier Transform in quantum mechanics. This paper will start by discussing the theoretical underpinnings and mathematical properties of the Fourier Transform, including its linearity, shift, and scaling properties. Following this, this paper will explore the Plancherel theorem and Parseval's identity, highlighting their importance in preserving energy and inner products in quantum systems.

Next, this paper will examine the Heisenberg Uncertainty Principle in detail, demonstrating how the Fourier Transform provides a mathematical framework for understanding the trade-offs between position and momentum uncertainties. This paper will also discuss the Planck-Einstein relation and its relevance to the energy distribution of quantum states, using Fourier analysis to elucidate these relationships.

Finally, this paper will present case studies and examples of how the Fourier Transform is utilized to solve the Schrödinger equation in different quantum mechanical scenarios. These examples will illustrate the practical applications of Fourier analysis in simplifying and solving complex quantum problems, reinforcing the transformative impact of this mathematical tool on the field of quantum mechanics.

2. Methods and Theory

2.1. Fourier Transform

This article centers on the use of Fourier transforms to prove or solve theorems or problems within quantum mechanics. The main method of this article is Fourier transform [3]

$$\hat{f}(t) = \int_R f(x)e^{-2\pi ix}\mathrm{dx}, \quad (1)$$

where $\hat{f}(t)$ is called the Fourier function of $f(x)$. The first linear property means that if there are two signals $f(t)$ and $g(t)$, and their respective Fourier transforms are $\mathfrak{F}(\omega)$ and $\mathfrak{G}(\omega)$, then the Fourier transforms satisfy the following relationship for any two invariants m and n:

$$\mathrm{m}\mathfrak{F}(\omega) + \mathrm{n}\mathfrak{G}(\omega) = \mathfrak{F}[mf(t) + ng(t)], \mathrm{m}\mathfrak{F}(\omega) + \mathrm{n}\mathfrak{G}(\omega) = \mathfrak{F}[mf(t) + ng(t)] \quad (2)$$

This means that the Fourier transform can be performed on $f(t)$ and $g(t)$ separately, and then the results can be summed up to give the same result as if the Fourier transform were performed directly on $mf(t) + ng(t)$.

If the Fourier transform of $f(t)$ is $\mathfrak{F}(\omega)$, then the Fourier transform of $f(t - t_0)$ (i.e., $f(t)$ is shifted to the right by t_0 in the time domain) is: $e^{-j\omega t_0}\mathfrak{F}(\omega)$). This indicates that a time domain shift leads to a phase factor in the frequency domain. If the Fourier transform of $f(t)$ is $\mathfrak{F}(\omega)$, then the Fourier transform of $f(at)$(i.e., the scaling of $f(t)$ by a factor of a in the time domain) is

$$F = \frac{1}{|a|}\mathfrak{F}\left(\frac{\omega}{a}\right) \quad (3)$$

which demonstrates that scaling in the time domain results in alterations to both scaling and magnitude in the frequency domain.

The symmetry properties of the Fourier transform include the symmetry of the real and imaginary parts. If $f(t)$ is a real function, then the real part of its Fourier transform $\mathcal{F}(\omega)$ is an even function, and the imaginary part is an odd function:

DOI: 10.54254/2753-8818/52/2024CH0114

$$Re[\mathfrak{F}(-\omega)] = Re[\mathfrak{F}(\omega)] \tag{4}$$
$$Im[\mathfrak{F}(-\omega)] = -Im[\mathfrak{F}(\omega)] \tag{5}$$

Furthermore, if $f(t)$ is a real function, then the Fourier transform $\mathcal{F}(\omega)$ of $f(t)$ satisfies the following relation:

$$\mathfrak{F}(-\omega) = \mathfrak{F} * (\omega) \tag{6}$$

where $\mathfrak{F} * (\omega)$ is the complex conjugate of $\mathfrak{F}$ (ω), which indicates that the Fourier transform of $f(t)$ is conjugate symmetric.

2.2. Plancherel Theorem and Parseval's Identity

Plancherel Theorem and Parseval's identity are fundamental results in Fourier analysis that are crucial for energy and inner product preservation in quantum mechanics [4].

The Plancherel theorem states that for any function $f \in L^2(R^n)$, its Fourier transform $\hat{f} \in L^2(R^n)$, and the following equality holds:

$$\int_{R^n} |f(x)|^2 \, dx = \int_{R^n} \left|\hat{f}(k)\right|^2 \, dk \tag{7}$$

Parseval's identity states that for any functions $(f, g \in L^2(R^n))$, the inner product is preserved under the Fourier transform:

$$\int_{R^n} f(x)\overline{g(x)} \, dx = \int_{R^n} \hat{f}(k)\overline{\hat{g}(k)} \, dk \tag{8}$$

This shows that the Fourier transform preserves the inner product in the L^2 space, demonstrating its importance in quantum mechanics. These theorems underline the fundamental properties of the Fourier transform, ensuring energy conservation and inner product preservation, which are crucial for quantum mechanical systems.

2.3. Heisenberg Uncertainty Principle

Heisenberg's uncertainty principle explains the relationship between the uncertainty in a particle's position and momentum. [5]:

$$\Delta x \cdot \Delta p \geq \frac{\hbar}{2} \tag{9}$$

where Δx represents the uncertainty of the position, Δp denotes the uncertainty in the momentum, and $\hbar$ is the approximate Planck constant.

In quantum mechanics, a particle's state is described by the wave function $\psi(x)$, This wave function represents the probability amplitude of locating the particle at a specific position x. Essentially, $\psi(x\ )$ provides a mathematical description of the likelihood of finding the particle at various positions in space. [6].

The Fourier function $\phi(p)$ of wave function $\varphi(x)$ describes the probability amplitude of the particle's momentum space

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_R \varphi(x) e^{-\frac{ipx}{\hbar}} \mathrm{dx} \tag{10}$$

inverse transform as

$$\varphi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_R \phi(p) e^{\frac{ipx}{\hbar}} \mathrm{dp}. \tag{11}$$

Uncertainty of position Δx and momentum Δp is defined as the standard deviation of wave function $\psi(x)$ and $\phi(p)$:

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 \tag{12}$$
$$(\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2 \tag{13}$$

According to the nature of the Fourier transform, the wave function has a certain duality between the time and frequency domains (momentum space). Specifically, the width of the wave function is inversely related to the width of its Fourier transform. Suppose that $\varphi(x)$ and $\phi(p)$ satisfied the normalizing conditions:

$$\int_{-\infty}^{\infty} |\varphi(x)|^2 dx = 1, \int_{-\infty}^{\infty} |\phi(p)|^2 dx = 1 \qquad (14)$$

Applying the Cauchy-Schwarz inequality [7], and by conducting norm analysis on ε (x) and ε (p), the following relationship is obtained:

$$\Delta x \cdot \Delta p \geq \frac{\hbar}{2} \qquad (9)$$

The Fourier transform played the role of a bridge in proving Heisenberg's inequality, through which two interrelated physical quantities, position, and momentum, could be mathematically linked, thus revealing the fundamental uncertainty principle in quantum mechanics.

Using Parseval's identity, this paper can demonstrate the preservation of inner products in position and momentum space, which directly relates to the uncertainty principle. Parseval's identity states that for any wave functions $\psi(x)$ and $\phi(x)$, their Fourier transforms ($\hat{\psi}(p)$) and ($\hat{\phi}(p)$) satisfy:

$$\int_R \psi(x)\overline{\phi(x)}\ dx = \int_R \hat{\psi}(p)\overline{\hat{\phi}(p)}\ dp \qquad (15)$$

This implies that the inner product of wave functions is preserved under Fourier transform, ensuring that the probability distributions in position and momentum space are related by:

$$\int_R |\psi(x)|^2\ dx = \int_R |\hat{\psi}(p)|^2\ dp \qquad (16)$$

This equality supports the Heisenberg uncertainty principle by linking the spread (or uncertainty) in position space to the spread in momentum space.

2.4. Planck-Einstein Relation

The Planck Einstein relationship mainly describes the relationship between the energy of photons and their frequency [8]:

$$E = h\upsilon \qquad (17)$$

Among them, E is the energy of the photon, h is the Planck constant, and υ is the frequency of the photon.

The state of particles (including photons) is defined by the wave function $\psi(x,t)$)In quantum mechanics. [9]. This wave function provides a complete description of the quantum state, encapsulating all the information about the particle's position and momentum probabilities at any given time. The Fourier transform enables the conversion of the wave function from the time domain to the frequency domain, thus revealing its frequency components. Assuming that the wave function $\psi(x,t)$ can be represented by Fourier transform as:

$$\psi(x,t) = \int_R \tilde{\psi}(k,\omega)e^{i(kx-\omega t)}\,\mathrm{dx} \qquad (18)$$

Among them, $\tilde{\psi}(k,\omega)$ is the representation of the wave function in the frequency domain (or wavenumber domain), k is the wavenumber, and ω is the angular frequency.

According to the Planck Einstein relationship, the relationship between energy E and frequency υ can be written as [10]:

$$E = h\upsilon = \hbar\omega \qquad (19)$$

where $\omega = 2\pi\upsilon$. The Fourier transform is a powerful tool that can help to understand the distribution of wave functions at different frequencies (or energies). For example, suppose the representation of the wave function in the time domain is $\psi(t)$. By applying the Fourier transform to $\psi(t)$, then obtain $\tilde{\psi}(\omega)$, which represents the distribution of the wave function in the frequency domain (or energy domain):

$$\tilde{\psi}(\omega) = \frac{1}{\sqrt{2\pi}}\int_R \psi(t)e^{-i\omega t}\,\mathrm{dt} \qquad (20)$$

DOI: 10.54254/2753-8818/52/2024CH0114

By analyzing $\tilde{\psi}(\omega)$, the contribution of the wave function at different energies (or frequencies) can be obtained.

The Fourier transform can help to understand the relationship between the wave function in the time and frequency (or energy) domains, thus revealing the energy distribution and frequency characteristics of particles (including photons)

2.5. Solution of The Schrödinger Equation

The Fourier transform is crucial for solving the Schrödinger equation, which characterizes the evolution of a particle's wave function in quantum mechanics over time and space. This fundamental equation describes how the quantum state of a physical system changes over time, providing a comprehensive framework for predicting the behavior of particles.

By applying the Fourier transform to the Schrödinger equation, the problem is converted from the spatiotemporal domain to the frequency (or wavenumber) domain. This transformation simplifies the solving process by breaking down complex wave functions into their constituent frequency components. In the frequency domain, differential equations often become algebraic equations, which are generally easier to solve. This method allows for more straightforward analysis and interpretation of the wave function's behavior.

The Schrödinger equation for a non relativistic single particle is typically written as:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t}=-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}+V(x)\psi(x,t) \qquad (21)$$

where, $\psi(x,t)$ is wave function; i is an imaginary unit; $\hbar$ is the reduced Planck constant; m is the mass of the particle; $V(x)$ is the potential energy function. After performing a Fourier transform on the wave function $\psi(x,t)$, the spatial derivative term $\frac{\partial^2\psi(x,t)}{\partial x^2}$ is transformed into:

$$\frac{\partial^2\psi(x,t)}{\partial x^2}=\frac{1}{\sqrt{2\pi}}\int_R -k^{2\tilde{\psi}(k,t)}e^{-ikx}\,\mathrm{dx} \qquad (22)$$

In this way, the spatial derivative term in the Schrödinger equation is transformed into a form in the wavenumber domain.

Perform Fourier transform on the entire Schrödinger equation to obtain the Schrödinger equation in the wavenumber domain:

$$\mathrm{i}\hbar\frac{\partial\tilde{\psi}(k,t)}{\partial t}=\frac{\hbar^2k^2}{2m}\tilde{\psi}(k,t)+\frac{1}{\sqrt{2\pi}}\int_R V(x)\psi(x,t)e^{-ikx}\,\mathrm{dx} \qquad (23)$$

For different forms of potential energy $V(x)$, the Fourier transformed Schrödinger equation may be easier to solve. For example, for a free particle (i.e. $V(x)=0$), the Fourier transformed Schrödinger equation becomes:

$$i\hbar\frac{\partial\tilde{\psi}(k,t)}{\partial t}=\frac{\hbar^2k^2}{2m}\tilde{\psi}(k,t) \qquad (24)$$

This equation is a simple linear differential equation that can be solved using standard methods:

$$\tilde{\psi}(k,t)=\tilde{\psi}(k,0)e^{-\frac{\hbar^2k^2}{2m}t} \qquad (25)$$

By solving the Schrödinger equation after Fourier transform to obtain k, and then returning to the spacetime domain through inverse Fourier transform to obtain the wave function $\psi(x,t)$

$$\psi(x,t)=\frac{1}{\sqrt{2\pi}}\int_R\tilde{\psi}(k,t)e^{ikx}\,\mathrm{dk} \qquad (26)$$

This transformation makes the spatial derivative term easier to handle, particularly for specific forms of potential energy functions, and can significantly simplify the solution process. The use of the Fourier transform and its inverse enables flexible conversion between different domains, thereby allowing for more effective analysis and problem-solving in quantum mechanics.

DOI: 10.54254/2753-8818/52/2024CH0114

3. Results and Application

3.1. Application of Uncertainty Principle:

Consider a Gaussian wave packet defined in position space as [11]:

$$\psi(x) = \left(\frac{1}{2\pi\sigma_x^2}\right)^{1/4} e^{-\frac{x^2}{4\sigma_x^2}} \tag{27}$$

The Fourier transform of this wave packet gives its representation in momentum space:

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathrm{R}} \psi(x) e^{-\frac{ipx}{\hbar}} \, d\mathrm{x} \tag{28}$$

For a Gaussian wave packet, the Fourier transform is also a Gaussian:

$$\phi(p) = \left(\frac{\sigma_x^2}{\hbar^2}\right)^{1/4} e^{-\frac{p^2\sigma_x^2}{\hbar^2}} \tag{29}$$

This shows that a Gaussian wave packet in position space remains Gaussian in momentum space, but with a different width parameter.

For the Gaussian wave packet $\psi(x)$, the position uncertainty Δx is given by the standard deviation: $\Delta x = \sigma_x$. For the Gaussian wave packet's Fourier transform $\phi(p)$, the momentum uncertainty Δp is: $\Delta p = \frac{\hbar}{2\sigma_x}$ .To verify the Heisenberg uncertainty principle, this paper compute the product of these uncertainties:

$$\Delta x \cdot \Delta p = \sigma_x \cdot \frac{\hbar}{2\sigma_x} = \frac{\hbar}{2} \tag{30}$$

This confirms the Heisenberg uncertainty principle: $\Delta x \cdot \Delta p \geq \frac{\hbar}{2}$.

The author also uses numerical methods to simulate the evolution of a Gaussian wave packet over time [12]. The time-dependent Schrödinger equation for a free particle is a fundamental equation in quantum mechanics that describes how the wave function of a particle evolves over time. For a free particle, which is not subjected to any external forces or potential fields, the equation simplifies to:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} \tag{31}$$

The solution can be expressed using the Fourier transform:

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_R \phi(p) e^{i\left(px - \frac{p^2}{2m\hbar}t\right)} \, d\mathrm{p} \tag{32}$$

For a Gaussian wave packet, the dispersion over time can be observed numerically [13]. As the wave packet evolves, its width in position space increases due to the spread in momentum space, which can be shown by:

$$\psi(x,t) = \left(\frac{1}{2\pi\sigma_x^2\left(1 + \frac{i\hbar t}{m\sigma_x^2}\right)}\right)^{1/4} e^{-\frac{x^2}{4\sigma_x^2\left(1+\frac{i\hbar t}{m\sigma_x^2}\right)}} \tag{33}$$

Using numerical simulations, this paper can plot the probability density$|\psi(x,t)|^2$ at various time intervals to visually demonstrate the dispersion of the wave packet. By examining these plots, the results clearly illustrate that the wave packet spreads out over time. This spreading effect is a direct consequence of the Heisenberg uncertainty principle, which dictates that the product of the uncertainties in position and momentum must remain constant. As time progresses, the uncertainty in position increases, leading to a broader wave packet, while the corresponding uncertainty in momentum adjusts accordingly to maintain this fundamental quantum relationship. This analysis provides a compelling visual and quantitative confirmation of the uncertainty principle in action, enhancing people's understanding of the dynamic behavior of quantum wave packets.

The author performs numerical simulations using Python to visualize the dispersion of the Gaussian wave packet. The initial wave packet and its evolution over time are plotted to demonstrate the increase in position uncertainty. Below are the plots showing the Gaussian wave packet at initial time $t = 0$ and at later times t_1, t_2, t_3: as shown in Figure 1. These plots illustrate the impact of the uncertainty principle on wave packet evolution, highlighting the relationship between position and momentum uncertainties.

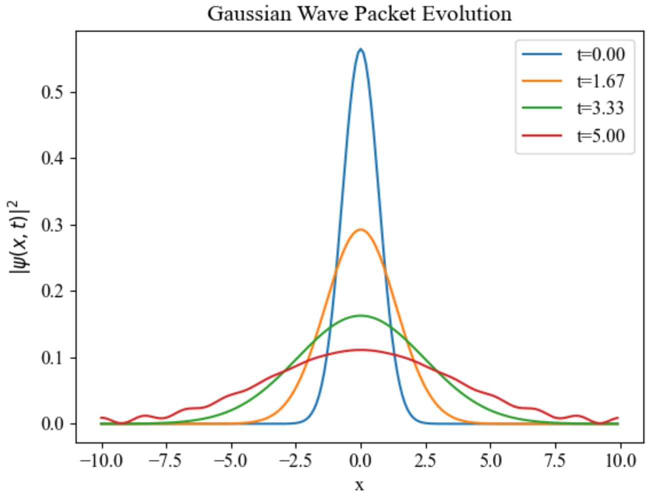


Figure 1. Evolution of a Gaussian wave packet over time, showing the probability density $|\psi(x,t)\,|^2$ at different time intervals. At ($t = 0.00$), the initial Gaussian wave packet. At ($t_1 = 1.67$), the wave packet begins to spread. At ($t_2 = 3.33$), the spread continues, showing increased uncertainty in position. At ($t_3 = 5.00$), the wave packet is significantly dispersed.

3.2. Quantum Measurements:

Quantum measurement experiments frequently involve the precise determination of a particle's position or momentum [14], leading to inherent uncertainties described by the Heisenberg uncertainty principle. These uncertainties highlight the fundamental limits of measurement precision in quantum mechanics. In these experiments, the Fourier transform emerges as a powerful analytical tool, enabling researchers to analyze and interpret the resulting data more effectively. Consider a typical quantum measurement experiment that a particle with measured position [15]. The resulting wave function $\psi(x)$ obtained from this measurement represents the probability amplitude of locating the particle at a specific position x. To gain insights into the momentum distribution of the particle, it is necessary to transform this wave function from the position space to the momentum space. This is achieved by applying the Fourier transform to $\psi(x)$, resulting in $\phi(p)$, the wave function in momentum space. The wave function $\phi(p)$ provides a comprehensive understanding of the particle's momentum characteristics, complementing the positional information initially obtained.

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_R \psi(x) e^{-\frac{ipx}{\hbar}}\, dx \tag{34}$$

By examining $\phi(p)$, then can gain insights into the particle's momentum distribution.

The relationship between position and momentum uncertainties during a quantum measurement can be thoroughly analyzed using the Fourier transform. This mathematical technique allows people to convert the wave function from its original position representation to its momentum representation, providing a detailed insight into the distribution and interplay of these uncertainties. When measuring a particle's position with high precision, the wave function $\psi(x)$ becomes sharply peaked, leading to a broad spread in $\phi(p)$. This illustrates the trade-off between position and momentum uncertainties, as described by the Heisenberg uncertainty principle. To demonstrate this, consider experimental data where the position of a particle is measured multiple times. The resulting probability distribution $P(x)$

can be Fourier transformed to obtain the momentum distribution $P(p)$. Using the experimental position data x_i,the author can construct the wave function:

$$\psi(x) = \sum_i \delta(x - x_i) \tag{35}$$

Applying the Fourier transform to $\psi(x)$, it is easy to get

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_R \sum_i \delta(x - x_i) e^{-\frac{ipx}{\hbar}} dx = \frac{1}{\sqrt{2\pi\hbar}} \sum_i e^{-\frac{ipx_i}{\hbar}} \tag{36}$$

By analyzing $\phi(p)$, the author can visualize the momentum distribution corresponding to the measured positions.

To illustrate the application of the Fourier transform in analyzing quantum measurements, this paper can plot the frequency spectrum of the measurement results, see Figure 2. Consider an experiment where the position of a particle is measured, resulting in a set of position data points x_i. The Fourier transform of the position distribution $P(x)$ gives the frequency spectrum $P(p)$. This spectrum reveals the contributions of different momentum components to the particle's wave function. Using the position data x_i from the experiment, the author can calculate the Fourier transform and plot the frequency spectrum. The plots show the position distribution $P(x)$ and the corresponding momentum distribution $P(p)$ obtained through the Fourier transform. This demonstrates how the Fourier transform can be used to analyze and interpret quantum measurement results, highlighting the relationship between position and momentum uncertainties.

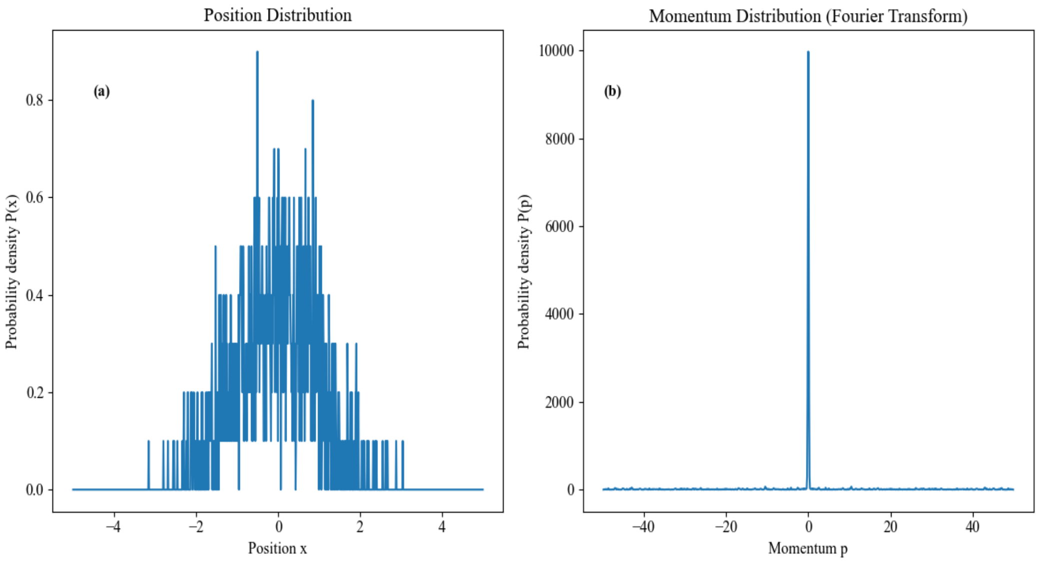


Figure 2. (a) Position distribution $P(x)$ showing the probability density as a function of position x.(b) Momentum distribution $P(p)$ obtained through Fourier transform, showing the probability density as a function of momentum p.

4. Conclusion

This paper systematically explores the significant applications and theoretical implications of the Fourier transform in quantum mechanics. By analyzing the Uncertainty Principle, the Planck-Einstein relation, and utilizing the Fourier transform to solve the Schrödinger equation, this paper demonstrates the extensive use of the Fourier transform as a fundamental tool in quantum mechanics. The Fourier transform provides a robust framework for understanding wave functions and their behavior in both position and momentum space at a theoretical level. Its application greatly simplifies the analysis and solution process of complex quantum systems in practice. Through specific case studies and numerical simulations, this paper showcases the practical applications of the Fourier transform in verifying the

DOI: 10.54254/2753-8818/52/2024CH0114

Heisenberg Uncertainty Principle and quantum measurements, further emphasizing its crucial role in quantum mechanic's research. Additionally, the Fourier transform aids in illustrating the wave-particle duality and the probabilistic nature of quantum states, which are central to quantum theory. Looking ahead, with advancements in computational power and algorithm development, the integration of Fourier transform techniques with quantum computing and machine learning is expected to offer new possibilities for simulating and analyzing quantum systems with unprecedented precision. These developments could lead to breakthroughs into understanding and manipulation of quantum phenomena, reinforcing the indispensable role of the Fourier transform in the ongoing evolution of quantum mechanics.

References

[1] Golse, F., & Paul, T. (2017). Empirical Measures and Quantum Mechanics: Applications to the Mean-Field Limit. *Communications in Mathematical Physics*, 369, 1021-1048.
[2] Jaeger, G. (2015). Measurement and Fundamental Processes in Quantum Mechanics. Foundations of Physics, 45(2), 152-165.
[3] Oppenheim, A. V., & Schafer, R. W. (1975). Digital Signal Processing. Prentice-Hall.
[4] Kouba, O. (2013). A Mixed Parseval-Plancherel Formula. Journal of Classical Analysis, 4(1), 63-76.
[5] Poojary, B. (2015). Origin of Heisenberg's Uncertainty Principle. American Journal of Modern Physics, 4(4), 193-198.
[6] Gao, S. (2011). The Wave Function and Quantum Reality. Journal of Modern Physics, 2, 482-488.
[7] Wang, X. (2009). The Cauchy-Schwarz Inequality and Its Applications. Mathematical Journal, 45(3), 203-217.
[8] Williams, P. E. (2013). Phat photons and phat lasers. Proceedings of SPIE, 8841
[9] Deng, X., & Deng, Z. (2022). The Matter Wave Is Space-Time Wave. Hypothesis and Theory Journal, 2(1), 60-75.
[10] Livine, E. (2023). Evolution of the wave function's shape in a time-dependent harmonic potential. Europhysics Letters, 122(3), 1-12.
[11] Ryan, F. (1988). Acoustic wave propagation employing Gaussian wave packets. The Journal of the Acoustical Society of America, 84(5), 1792-1799.
[12] Li, H., Feng, B., & Wang, H. (2012). Gaussian Packets Modeling and Migration. The Journal of the Acoustical Society of America, 84(5), 1792-1799
[13] Reimers, J., & Heller, E. (1985). The exact thermal rotational spectrum of a two-dimensional rigid rotor obtained using Gaussian wave packet dynamics. The Journal of Chemical Physics, 83(2), 516-523.
[14] Wick, K. (2019). On the Quantum Mechanical Description of the Interaction between Particle and Detector. Physics, 3(4), 61.
[15] Wan, K., & McLean, R. (1991). Are all quantum measurements reducible to local position measurements. Journal of Physics A: Mathematical and General, 24(8), 1681-1691.

Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations
DOI: 10.54254/2753-8818/52/2024CH0116

Explore random music generator based on Short-Time Fourier Transform

Cailin Peng

Peabody Institute, Johns Hopkins University, Baltimore, Maryland, United States

cpeng28@jh.edu

Abstract. This paper delves into the mechanism of how the Short-Time Fourier Transform (STFT) is used for generating random music. A brief review of the history of stochastic music development is at the outset. The paper contains the principle of audio digitization. This starts with how to convert a continuous audio signal into discrete samples. The Nyquist Theorem plays an important role in that process to preserve signal integrity. The Heisenberg uncertainty principle takes effect as the STFT is applied to covert these samples. It states that when converting the audio signal between the time domain and the frequency domain, the audio signal can only have the properties of one or the other. This paper categorized the audio signals into three categories based on their spectral characteristics. The paper also points out the reason why natural sound effects are always chosen to generate random music due to their inherent complexity and randomness. This paper demonstrates the detail of STFT in generating random music, explaining how to specifically manipulate spectrum analysis to generate random music with practical examples. The paper concludes by discussing the possible future role and direction of STFT in the field of stochastic music.

Keywords: Short-time Fourier transform, Stochastic music, Random music, Audio analysis.

1. Introduction

In the early 20th century 1951, a piece called “Music of Changes” was published, and it was written by John Cage. This work demonstrates John Cage's quest for randomness in his music, to remove the influence of personal intentionality from his work, and to pursue the music itself. He used the I Ching to generate a randomized chart of the different parameters of the music to create this piece. The music in it is determined by the random date that appears in the I Ching rather than by the composer's intentions. In 1952, He presented the famous piece “4’33”, which contains nothing but natural sound effects around the audience. Here the randomness in his music is taken to another level. John Cage is well known throughout the world along with its great controversy. Where exactly are the boundaries of musical randomness? In recent years, randomness in music has risen in tandem with the growing modern quest for innovation in music style. Completely randomness does not exist, yet when it comes infinitely close to what the world calls random, can it still be called music? This article will not delve into the huge philosophical contradictions behind it, but an explore of STFT in the use of generate the randomness in stochastic music. Another influential composer in explored indeterminacy in music is called Iannis Xenakis. His piece "Pithoprakta" applies probability theory and mathematical models.

The STFT appears with the needs of analysis for the non-stationary signals—signals whose frequency content changes over time. In 1946 Dennis Gabor, a physicist and Nobel laureate, brings the idea of analyzing both time and frequency domains simultaneously by use of Gaussian functions as window functions in the traditional Fourier Transform (FT). The alternatives are needed for relaxed translation invariance assumptions [1]. Based on this idea, the STFT is then formalized and perfected in subsequent decades. In the audio analysis field, the STFT usually fits better than the FT or Discrete Fourier Transform (DFT). Compared to FT and DFT, the STFT enables the capture of how frequencies evolve over the timeline. It divides the signal into small overlapping time segments and then applies the Fourier Transform to each segment. The invention of STFT allows further analysis of complex sounds and textures. Jason Brown, a mathematician from Canada successfully replicates the chord in '*A Hard Day's Nigh'* by the Beatles using the FT [2]. By transforming and recombining spectral components, STFT allows for the generation of real-time random soundscapes. These technical play an important role in the development of modern random music. The paper will dive deep into the idea of the relationship between STFT and stochastic music.

2. Methods and theory

2.1. The principle of audio digitization

When an audio file is open, it is typically displayed as a waveform in Figure 1. This waveform is not a perfect storage of the original continuous sound signal. If zoomed in, the image will appear as a stepped line as shown, rather than a smooth curve. This discrepancy arises because digital sampling cannot record continuous sound signals. Instead, it records signals by sampling points through a fixed period. This process converts a continuous-time signal into a discrete-time signal.

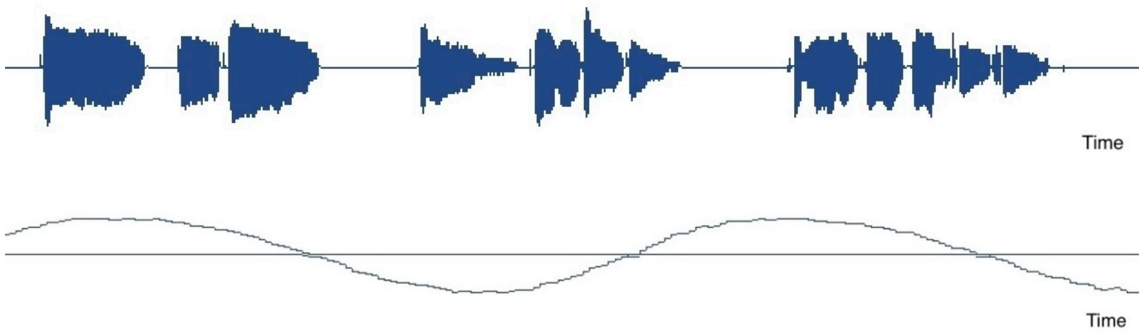


Figure 1. Top: An audio recording of the author's voice in Logic. Bottom: Zoom-in graph.

According to The Nyquist Theorem, to accurately reconstruct a continuous-time signal from its samples, the sampling rate must be at least twice the highest frequency component present in the signal.

The Nyquist-Shannon Sampling Theorem so called the Nyquist Theorem has a vital status in the field of signal processing. As shown in Figure 2, the redline is $y = sin4x$. If considered as audio signal input, it has the frequency $2/\pi$. The Nyquist Theorem states the lowest sampling rate to sample this signal and ensure that the original frequency can be restored from the sampled data is $2(2/\pi) = 4/\pi$. According to this sampling rate, the black spots shown in Figure 2 can restore the original frequency $2/\pi$. If the sampling rate is less than $4/\pi$, which violates The Nyquist Theorem, alias will appear when decomposing the signal. As shown in Figure 2 the red spots, can reconstruct more than one possible frequency based on the sampling data. So, to recover all the frequencies in the original signal, the sampling rate must be twice the maximum frequency. This relationship is expressed as:

$$f_s \geq 2f_{max} \tag{1}$$

Where f_s is the sampling rate, and $f_{max}$is the highest frequency component of the signal.

If the filter can attenuate the above half frequency below the analog-to-digital converter, the aliasing will not appear [3]. It is true that in some special cases, through irregular random sampling, the original signal can be accurately restored when the sampling rate is lower than the Nyquist Theorem states. The upper limit of human hearing is approximately 20kHz. To satisfy the Nyquist criterion, common sampling rates are set at 44.1 kHz, which is more than twice the maximum audible frequency. The higher the sampling rate results the more similar the restored signal to the original.

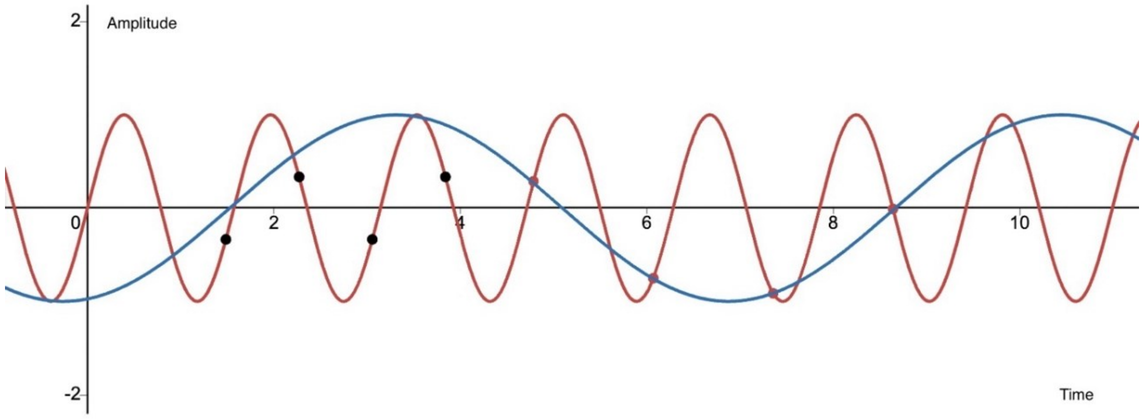


Figure 2. Graph demonstration in Desmos. The redline is $y = sin4x$ with the set of black spots being the sampling points at the same time interval $\pi/4$. The red sampling points have time intervals larger than $\pi/4$, and its set on both redline and blueline.

2.2. Short-Time Fourier Transform (STFT)

As the Fourier Series was built, it allowed the analysis of sound waves produced by strings and columns. Fourier analyses are used to decompose the naturally occurring harmonics [4]. The Fourier Transform, the predecessor of STFT, allows the transformation of the audio signal to the frequency spectrums. STFT overcomes the lack of time information drawback by considering an analysis window that has a specific time-frequency resolution property. In STFT, a window, which is a function being zero-valued outside of some chosen interval, is employed for extracting time information [5]. Each small-time interval STFT produces applied FFT to reveal the frequency-time spectrum switch. With curtain synthesis, these time segments are combined and construct data that contains frequency values as time progresses [6].

To apply the Short-Time Fourier Transform (STFT) to the signal input, the sampling is divided into a separate time each assumed to be periodic. The STFT function is given by:

$$X(t, f) = \int_{-\infty}^{\infty} x(\tau) \cdot w(\tau - t) \cdot e^{-j2\pi f\tau} d\tau, \qquad (2)$$

where w(τ−t) is the window function centered at time t, used to extract a time window of the original audio for analysis.

2.3. Categories of audio signals

The convertible audio signals can, based on their spectral image after the STFT, be divided into three broadly defined categories. The expected result for a pure sine wave input is a single peak lies on the exact Hz on the spectrum. The Sinusoidal formula is expressed as:

$$y(t) = Asin(2\pi ft + \phi) \qquad (3)$$

where A represents the amplitude, f represents the frequency, and t represents the time. ϕ is phase, the initial angle at $t = 0$. Sounds with harmonic structures usually refer to the musical notes played by all kinds of instruments. It contains a fundamental frequency additional with multiple harmonics at integer multiples of the original Hz. Natural sound effects are a superposition of random sources, and the spectral values presented with the FT are non-periodic. It is highly randomized and may contain noise

DOI: 10.54254/2753-8818/52/2024CH0116

with no apparent peaks. To obtain randomized input data, the natural sound effects are always chosen due to their complex and varied spectral characteristics.

2.4. Time-frequency trade-off

Time and frequency domains are two mutually exclusive representations of audio signals. The frequency spectrum represents the average of all times, and the time domain represents the average of all frequency values as time goes on [7]. A segment of audio samples is transformed into a spectrum by the STFT as the data is converted from the time domain to the frequency domain. This transformation highlights the Time-Frequency Trade-off. Heisenberg's Uncertainty Principle states that one cannot simultaneously achieve arbitrary precision in both time and frequency domains. Mathematically, this principle is expressed as:

$$\Delta t \cdot \Delta f \geq \frac{1}{4\pi} \tag{4}$$

To illustrate the concept of random music generation, this paper will not consider the middle tradeoffs situation and only focuses on analyzing signals either entirely in the time or frequency domain.

3. Results and Application

3.1. Manipulating and reconstruct

A way to introduce randomness in the music generation is to modify the values obtained in the spectrum. This can be done by altering the magnitude and phase components of the STFT output. The original phase contains information on the temporal structure of the original sound. Embedding random algorithms to specific frequency bins can result in the adjustment of amplitude. More chaotic effects can be created if both sound magnitude and phase are randomized.

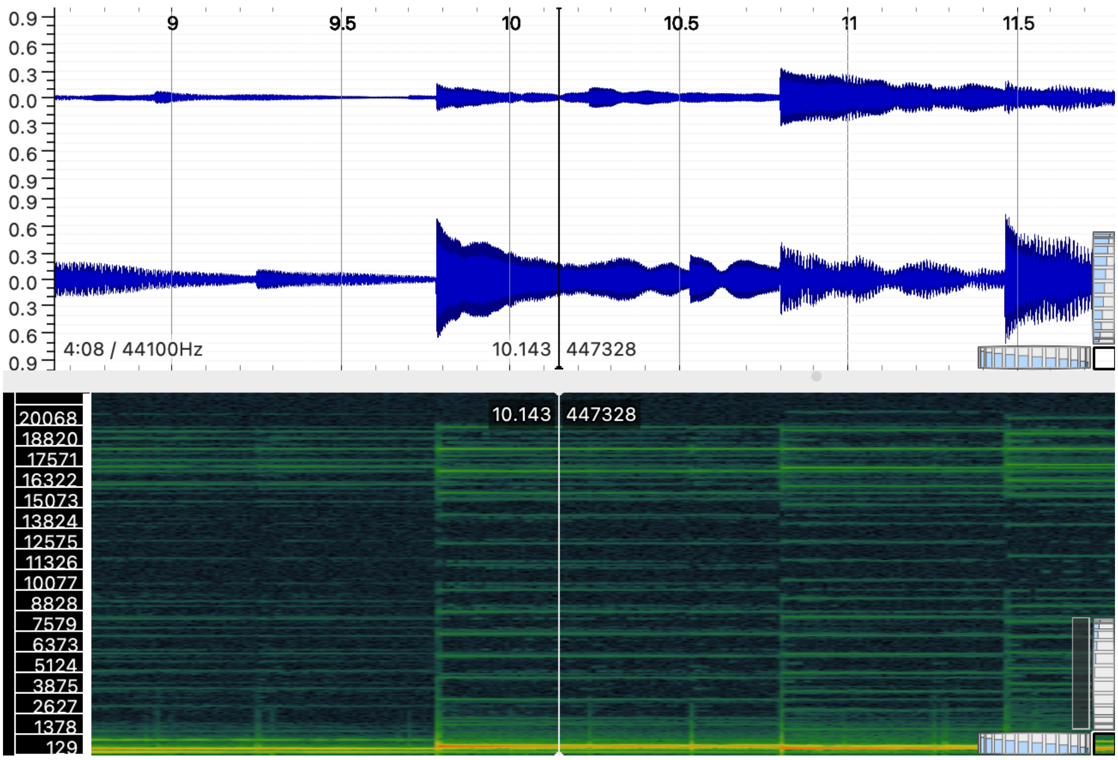


Figure 3. Real-time conversion diagram of a sound source from time domain to frequency domain in Sonic Visualiser.

For example, in Figure 3, modified data of the frequency spectrum values in some will result in fresh audio output. The modified STFT representations are converted back to the time domain using the inverse STFT (ISTFT). The ISTFT reconstructs the time-domain signal by integrating the modified spectra across all frames, windows stitch together seamlessly by overlap and add method. The new audio sounds will retain the temporal dynamics of the original sound but with randomized spectral characteristics.

Randomization of sound spectrum values serves as the foundation for creating random music. By applying different randomization strategies to the spectra and experimenting with various window sizes and overlap ratios, a diverse range of musical textures and effects can be generated. The article "Short Time Fourier Transform Based Music Genre Classification" by Ahmet Elbir, Hamza Osman Ilhan, Gorkem Serbes, and Nizamettin Aydin, presented at the 2017 25th Signal Processing and Communications Applications Conference (SIU), explores the use of Short Time Fourier Transform (STFT) for music genre classification. This paper evaluates the impact of six window types used in the STFT extraction step on music genre classification [8]. Additionally, interactive elements can be incorporated by linking the randomization parameters to external controls, allowing real-time manipulation of the generated music.

3.2. Application in real life

Jean-Claude Risset is a famous computer musician, well known for his innovative works in the field of sound synthesis and spectral analysis. With a simple version Fourier transform technique, he successfully decomposes the audio signal into constituent sine wave components. Risset put a lot of effort into researching the components of sound and how to resynthesize it with the FT techniques. The spectral techniques allowed precise control of analyzing the harmonic sound over the timeline. In 1985, He presented a piece called "Sud". In this piece, he decomposes natural sounds and re-synthesizes them to create a seamless blend of acoustic and electronic textures.

The works of Bell Labs contribute to the whole scientific world including building the foundation for the development of computer music software and techniques. Tenney did research in computer music techniques including FT-related synthesis at Bell Labs in the 1960s. He applied principles related to Fourier analysis to explore the stochastic nature of sound. In 1961, Tenney published a work called "Analog #1: Noise Study", in which he used noise and its spectral properties as the elements of this composition.

The first computer music-generated sound is produced by Max Mathews' MUSIC software, using the temporal envelope modulates Fourier series to define the timbre of instruments [9]. In Max/MSP STFT is used to analyze and resynthesize grains. This technique allows the dynamic manipulation of the pitch of the sound. And adjusts the timbre and duration of the audio signal. Simplifying the instructions software makes STFT and ISTFT easier to operate for their user. Effects such as spectral gating, morphing, and filtering can be implemented by manipulating the frequency components obtained from STFT. Like Max, Pure Data (Pd) is an open-source visual programming language for audio and multimedia processing. The real-time effects such as time-stretching and pitch-shifting can be made through Pd using STFT during the live performance. The Ableton Live, as many producers are familiar with, exists lives' warping features to align beats and harmonies from different tracks, creating seamless transitions and remixes. This includes advanced warping and stretching algorithms that rely on STFT for time-domain manipulation. Sound designers use STFT-based tools in Ableton Live to manipulate audio, creating granular textures and complex rhythmic patterns. In Ref [10], the author discusses the technique of classifying the music genres by input time-frequency spectrum to machine learning.

4. Conclusion

The paper introduces in detail the specific process of audio digitization and highlights the critical role that STFT plays in frequency randomization. The time and frequency tradeoff that is generated by STFT is been used widely in music randomization. By dividing the signal into separate time Windows, STFT can analyze the frequency components that change over time, thus it can be applied to the non-stationary

DOI: 10.54254/2753-8818/52/2024CH0116

audio signal. It provides a powerful tool for the innovation of digital music. The sound generated by the modulation of its frequency spectral is quite random and develops dynamically along the timeline. Achieved randomness and dynamics through the modulation spectrum of STFT have a wide range of application prospects. Not only in the field of music but also in fields such as signal processing, communication encryption, and audio engineering. This paper shows the significant impact of STFT technology on contemporary experimental music, demonstrating its potential to create new sound textures and expand the boundaries of music innovation. Many of the more complex sound syntheses and analyses of modern times are built on STFT. By leveraging the capabilities of STFT, musicians and engineers can explore new avenues for digital sound synthesis, resulting in innovative auditory experiences

References

[1] Torrésani B. (1999). Time-Frequency and Time-Scale Analysis. Academic Press, 55.
[2] Jessop S. (2017). The Historical Connection of Fourier Analysis to Music. The Mathematics Enthusiast, 14(1), 85.
[3] Harvey A. F, Cerna M. (1993). The fundamentals of FFT-Based Signal Analysis and Measurement in LabVIEW and LabWindows. National Instruments, 34055B-01, 7-8.
[4] Hammond J. (2011). Mathematics of Music. UW-L Journal of Undergraduate Research XIV, 4.
[5] Lenssen N & Needell D. (2014). An Introduction to Fourier Analysis with Applications to Music. Journal of Humanistic Mathematics, 4, 72-91.
[6] Smith J O. (2007). Mathematics of the Discrete Fourier Transform (DFT), W3K Publishing.
[7] Boashash B. (2003). Time-Frequency Signal Analysis and Processing: A Comprehensive Reference. Protoavis Productions. 4.
[8] Elbir A, İlhan H. O, Serbes G, & Aydın N. (2018). Short Time Fourier Transform Based Music Genre Classification. *Electric* Electronics*, Computer Science, Biomedical Engineerings' Meeting (EBBT)*, 1-4.
[9] Lostanlen V, Andén J, Lagrange M. (2019). Fourier at the heart of computer music: From harmonic sounds to texture. Comptes Rendus Physique, 20(5), 463.
[10] Toshniwal T, Tandon P, Nithyakani P. (2022). Music Genre Recognition Using Short Time Fourier Transform And CNN. 2022 International Conference on Computer Communication and Informatics (ICCCI), 1-4.

Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations
DOI: 10.54254/2753-8818/52/2024CH0138

Stock prediction based on HMM and LSTM model and model selection using SVM

Moxun Deng

Leicester International Institute, Dalian University of Technology, Dalian, 116085, China

2939467211@mail.dlut.edu.cn

Abstract. In the field of time series analysis, stock prediction has always been a hot research topic. Hidden Markov Model (HMM) and Long Short-Term Memory (LSTM) are two powerful models that are very suitable for the stock prediction, but they will have different performance for different datasets. Therefore, selecting the appropriate model for different datasets has become a challenge. For different datasets, their statistical characteristics are an important means of analyzing datasets and the main feature that distinguishes them. Therefore, by analyzing the statistical features of the dataset and using these features for classification, it is possible to determine which model is most suitable for the dataset. This paper will use HMM and LSTM to predict some stock data, consider the effectiveness of the model through two dimensions of Mean Absolute Percentage Error (MAPE), and extract statistical feature values of these data. Finally, Support Vector Machine (SVM) will be used to classify the test dataset and verify whether the predicted model performs better than another.

Keywords: Stock prices, Hidden Markov Model, Long Short-Term Memory, Support Vector Machine.

1. Introduction

With the unprecedented prosperity of the stock market, researchers have never stopped accurately predicting stock prices. The stock model is almost the most complex and influential model in the real world. The price and return of stocks are only observable states from the outside world, and a large number of implicit factors are the main factors leading to stock price fluctuations. Morck et al. indicated that stock prices are related to the current level of economic development in a country [1]. Wen et al. believed that stock prices are related to changes in consumer emotions [2]. Hong et al. thought that prejudice is also one of the important factors affecting stock prices [3]. Song and Yu believed that monetary policy is also an important factor affecting stock prices [4]. It is not difficult to see that the standards proposed above are not quantifiable, which poses a huge obstacle to accurate stock prediction. But as people delve deeper into stock research, more and more models are being used to predict stock models. These models have their own characteristics in predictive performance, making it difficult to choose the best model suitable for all types of data.

In the previous work of Hassan et al. Hidden Markov Model (HMM) was used for stock research, with a focus on aviation stocks, and was used to predict prices [5]. Nguyet introduced Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) into the HMM algorithm for

DOI: 10.54254/2753-8818/52/2024CH0138

state selection [6]. Su and Deng believed that model SVM can be used as a tool for model selection [7]. Hassan et al. combined HMM, Artificial Neural Network (ANN), and Genetic Algorithm (GA) algorithms to form a new algorithm [8]. Nelson et al. used Long Short-Term Memory (LSTM) to predict stock prices [9]. Jiang and Liu combined Auto-Regressive Moving Average Model (ARMA) with LSTM for stock prediction [10].

In summary, the research on stock prices has attracted numerous scholars. This article will implement HMM and LSTM, and compare the prediction results. After comparison, this article will conduct an innovative study: analyzing why the same model performs differently in different datasets by extracting statistical feature values from the dataset, and selecting models by learning statistical features from previous datasets. In the process of seeking the best model, this article adopts Support Vector Machine (SVM) as the algorithm for training the model.

2. Methodology

2.1. Data source

The data for this literature was collected from the website investing.com, which contains a large amount of stock data that can be used for model training. The dataset includes closing prices, opening prices, the highest and lowest prices traded on the day, and the rate of return calculated based on the average price.

The downloaded dataset includes stock data from January 2014 to June 2024, with approximately 2600 pieces of data. In order to maintain the reliability of the prediction results, the author trimmed some of the data to ensure consistent data volume for all participating datasets (Table 1).

Table 1. Sample Dataset

Date	Close	Open	High	Low	Return
2024/6/28	183.42	185.72	186.58	183.34	-1.84%
……	186.86	185.65	187.49	185.54	0.80%
2014/1/29	185.37	184.19	185.93	184	-0.11%

The dataset requires certain preprocessing when making formal model predictions, which will be explained later in this paper. To train the SVM model, people need to perform further operations on the basic dataset. The statistical indicators including mean, weighted mean (calculated using weights assigned to each value), trimmed mean (calculated after removing extreme values), median (median in sorted datasets), skewness (measure of distribution asymmetry), MAD (mean absolute deviation, providing dispersion around the mean), Q1 (first quartile), Q3 (third quartile), IQR (representing the interquartile range of the middle 50% distribution), mode (most common value), range (difference between maximum and minimum values), variance (measure of data distribution), and kurtosis (distribution shape). Not all of these statistical features are useful, so it is necessary to reduce some to improve model performance.

2.2. Method introduction

This section of this paper will focus on the algorithms used in the prediction and selection processes. This section of the paper will concentrate on the algorithms employed in the prediction and selection processes. The paper will use Hidden Markov Models (HMM) and Long Short-Term Memory (LSTM) for predicting stock data, and capturing their Mean Absolute Percentage Error (MAPE). Subsequently, it will compute the statistical characteristic values of these datasets and compile a dataset. Finally, the Support Vector Machine (SVM) will be employed to classify the test dataset, and the classification outcomes will be compared with the actual results.

2.2.1. The principle of HMM

The hidden Markov model (HMM) is a probabilistic model used for analyzing time series data. Its fundamental component is the hidden Markov chain, a variant of the Markov chain. A standard Markov chain consists of a state transition matrix and a list of observable states. Each state in the list is observable, while the transition matrix records the probabilities of transitioning from one state to another. When using Markov chains for data prediction, it's assumed that the next state depends solely on the current state, known as the Markov hypothesis.

However, practical applications have shown that Markov chains often struggle to accurately predict data. The hidden Markov chain aims to address this limitation by uncovering and studying these hidden variables through observable data, thereby improving predictions based on trends in these hidden factors.

Overall, a hidden Markov chain consists of the following parts. Firstly, it is the observation state, which is a state that people can directly observe during observation, and is usually the state that people ultimately want to obtain results in predicting problems. Next is the hidden state. Hidden states are states that people cannot obtain through observation. Finally, there is a transition matrix between the observed state and the hidden state. In a hidden Markov chain, the observed state is only determined by the hidden state at the current time point, while the current hidden state is only determined by the hidden state at the previous time point. The specific situation is shown in the figure 1:

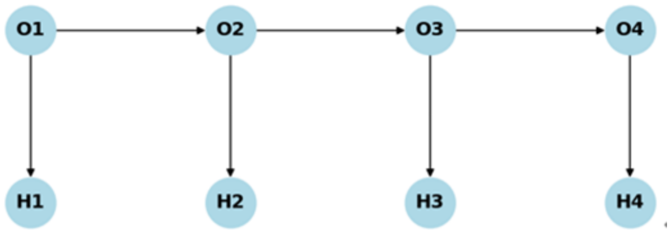


Figure 1. Hidden Markov Chain.

Therefore, although people can calculate the transition matrix of the observed state, in most cases, this transition matrix does not have practical significance.

For Hidden Markov Models, selecting the number of states is another challenge. This paper adopts the Bayesian decision-making method to select the number of states [2]. When selecting a model for prediction, there is a default assumption that there exists an optimal model for any given dataset, which can be found based on known data, and Bayesian Information Criterion is a simple selection criterion. The formula for calculating Bayesian Information Criterion is:

$$BIC = -2\ln(L) + \ln(n) * k \tag{1}$$

Where L is the maximum likelihood under this model, n is the number of data, and k is the number of variables in the model. When selecting the model state, the first step is to set a range for the model state, and then select the optimal situation based on the calculation results of BIC.

2.2.2. The principle of LSTM

Long Short-Term Memory (LSTM) represents an enhancement over the traditional Recurrent Neural Network (RNN) algorithm. It's widely acknowledged that vanilla RNNs struggle with processing long sequences due to issues like gradient explosion and vanishing gradients. Gradient explosion can be managed by capping gradients to prevent NaN outputs during training. Conversely, vanishing gradients are more insidious as they're harder to detect and mitigate.

When vanishing gradients occur, earlier data in the training process gradually lose significance, potentially reducing their contribution to the model's overall training effectiveness or rendering them negligible. Researchers often infer gradient vanishing from poor model training outcomes rather than direct detection.

LSTM addresses the issue of RNN forgetting long-term dependencies by introducing cell state c_t dedicated to retaining long-term information. LSTM typically have three inputs and two outputs. Inputs

DOI: 10.54254/2753-8818/52/2024CH0138

include the previous cell state c_{t-1} and output h_{t-1} from the preceding LSTM, along with the current network input x_t. Outputs consist of the current cell state c_t and output h_t for the subsequent LSTM.

In order to handle the relationship between these three input values, the concept of gate was introduced in LSTM. LSTM uses two gates to control the content of cell state c_t, one is a forget gate, which determines how much of the previous cell state is retained until the current time c_t. Another is the input gate, which determines how much of the network's input x_t is saved to the cell state c_t at the current time. LSTM uses an output gate to control how much the cell state c_t outputs to the current output value h_t of LSTM.

2.2.3. *The principle of SVM*

SVM are generally used to solve classification problems, where a given dataset T, find a hyperplane that can separate the dataset, and this optimal hyperplane should maximize the geometric interval between the support vector and the hyperplane (Figure 2).

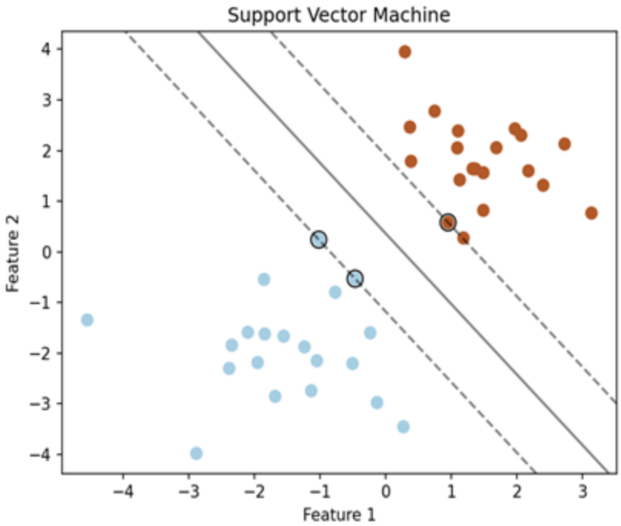


Figure 2. SVM schematic diagram.

In this paper, SVM is used to select the better term from HMM and LSTM in the prediction of the same dataset. As mentioned earlier, when choosing a prediction algorithm, it means that there must be an optimal solution here, and this optimal solution can be found by analyzing known data. For different datasets, algorithms have different performance, and there is no doubt that this change is caused by changes in certain features of the dataset itself. Therefore, analyzing these features can help people find possible algorithms with better performance among various algorithms. The statistical characteristics of the dataset are an important set of features that cannot be ignored. Therefore, SVM can be used for classification learning by combining the performance level of previous predictions on the dataset with the statistical characteristics of the dataset itself. Before making numerical predictions on the new dataset, calculate the statistical characteristics of the dataset, use the already trained SVM to select the algorithm, and use it as the evaluation criterion for algorithm selection.

3. Results and discussion

3.1. *Evaluation criterion*

This paper uses Mean Absolute Percentage Error (MAPE) to evaluate the performance of the model. MAPE is obtained by dividing the absolute difference between the predicted value and the true value by the true value. In specific model calculations, the MAPE of each step is added and divided by the number of model steps to obtain the value of the entire model's MAPE. It is worth noting that once a value of 0

DOI: 10.54254/2753-8818/52/2024CH0138

occurs in the model, MAPE will overflow and become unusable. Therefore, in order to solve this problem, the raw data needs to undergo certain preprocessing.

In model training, any price will not be zero, but the return may not change. Given that the possible range of returns is (-1, +∞), the author add 1 to the return and make its range (0, +∞) to avoid MAPE overflow. It is worth noting that the processed data is still valuable, and return at this point can be more conveniently calculated for the expected return when investors invest x units of cash.

3.2. HMM prediction results

The prediction results of HMM are shown in the following figure 3, 4 and table 2.

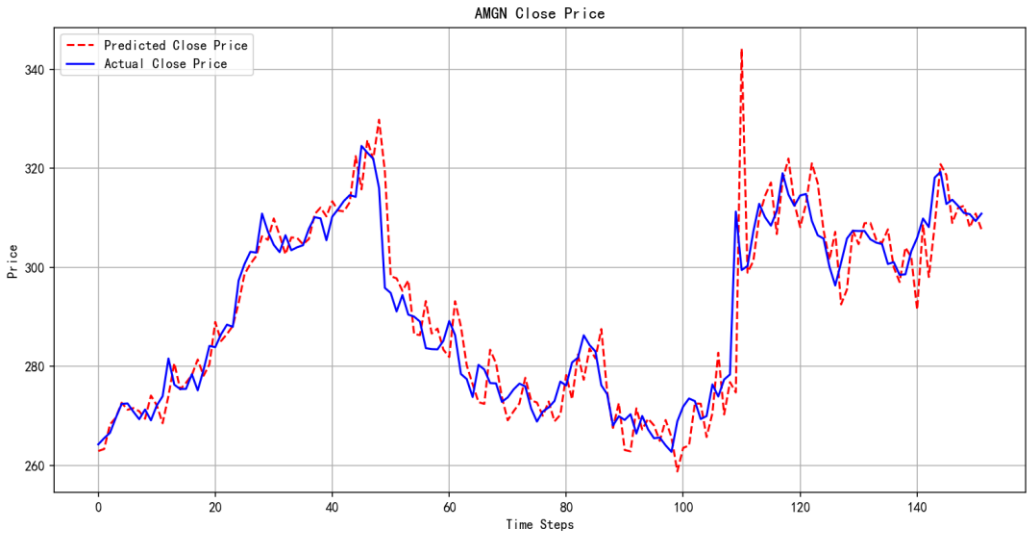


Figure 3. Sample of predictions on stock prices

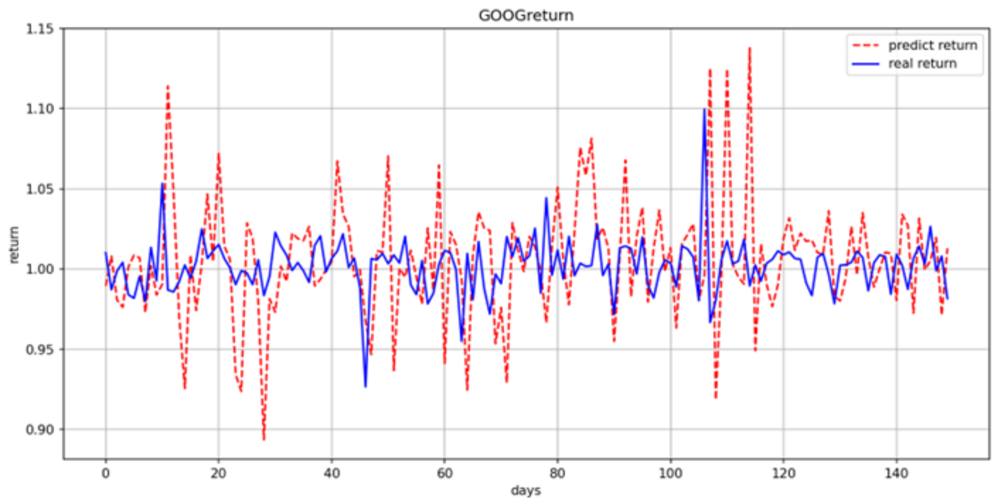


Figure 4. Sample of prediction of returns

It is not difficult to see that HMM performs well in predicting stock values, especially in predicting stock prices. Correspondingly, HMM's prediction of stock returns is not as accurate as price prediction. However, overall, HMM can accurately predict the trend of stock fluctuations, although the prediction is still not accurate to a certain extent, it is within an acceptable range.

Table 2. MAPE of prediction

Stock	MAPE of Close Price	MAPE of Open Price	MAPE of High Price	MAPE of Low Price	MAPE of Return Rate
GOOG	0.0185	0.0171	0.0178	0.0181	0.0343
TSLA	0.0161	0.0160	0.0154	0.0170	0.0486
NKE	0.0186	0.0178	0.0162	0.0176	0.0514
AMZN	0.0181	0.0172	0.0162	0.0176	0.0277
BA	0.0223	0.0213	0.0191	0.0213	0.0366
CVX	0.0130	0.0124	0.0125	0.0117	0.0207
CAT	0.0138	0.0171	0.0143	0.0145	0.0239
INTC	0.0257	0.0241	0.0228	0.0224	0.0415
MSFT	0.0127	0.0136	0.0116	0.0117	0.0216
DIS	0.0158	0.0155	0.0137	0.0121	0.0283
CSCO	0.0101	0.0110	0.0096	0.0085	0.0167
GS	0.0156	0.0124	0.0136	0.0125	0.0208
KO	0.0076	0.0077	0.0068	0.0076	0.0126
MCD	0.0108	0.0120	0.0094	0.0105	0.0167
MRK	0.0111	0.0118	0.0110	0.0104	0.0205
MMM	0.0157	0.0150	0.0159	0.0155	0.0305
AAPL	0.0148	0.0153	0.0140	0.0143	0.0224
AMGN	0.0151	0.0157	0.0157	0.0144	0.0268

The table 2 above provides a more detailed description of the specific situation when the HMM model predicts stocks. It is not difficult to see that the HMM model can adapt well to stock model prediction, and the MAPE is within 10%, which is an acceptable error.

3.3. LSTM prediction results

The prediction results of LSTM are shown in the following figure 5, 6 and table 3.

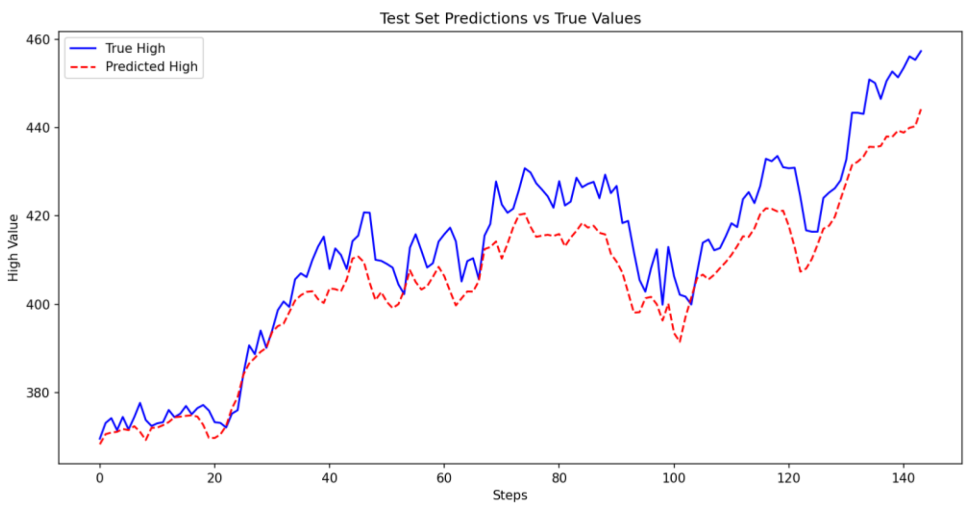


Figure 5. Sample of predictions on stock prices

DOI: 10.54254/2753-8818/52/2024CH0138

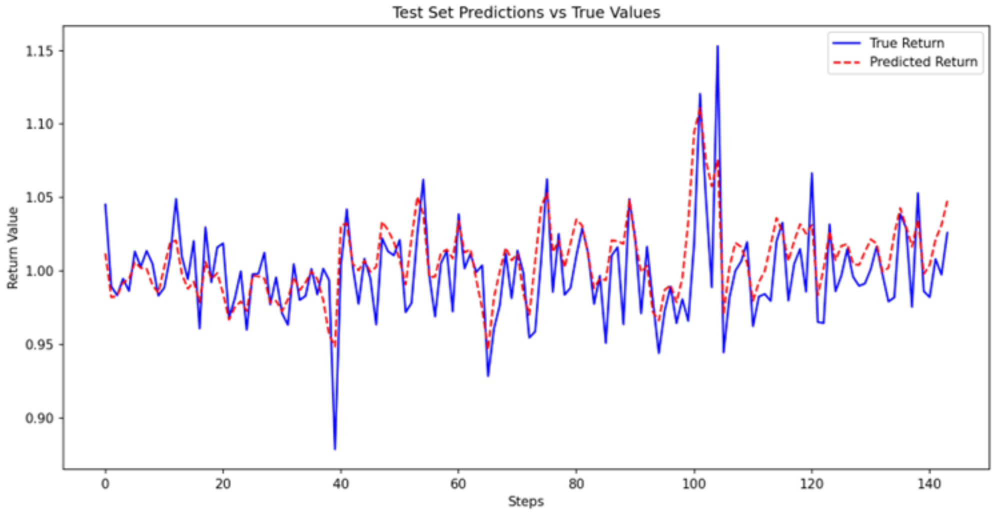


Figure 6. Sample of prediction of returns

It is not difficult to see that LSTM performs well in predicting stock values, especially in predicting stock returns. However, at the same time, LSTM cannot accurately predict stock prices, although the error is within an acceptable range. Compared with HMM, it is not difficult to find that these two different algorithms have different predictive effects on different forms of data. Therefore, it can be considered to find a more suitable algorithm for predicting data by calculating statistical features.

Table 3. MAPE of prediction

Stock	MAPE of Close Price	MAPE of Open Price	MAPE of High Price	MAPE of Low Price	MAPE of Return Rate
GOOG	0.0288	0.0244	0.0272	0.0357	0.0303
TSLA	0.0293	0.0272	0.031	0.0258	0.0387
NKE	0.0198	0.0156	0.0175	0.017	0.0309
AMZN	0.0202	0.0168	0.017	0.0182	0.0331
BA	0.0217	0.0162	0.0193	0.0181	0.0232
CVX	0.0123	0.0092	0.0107	0.0107	0.0337
CAT	0.025	0.023	0.0246	0.0275	0.0285
INTC	0.0242	0.019	0.0222	0.0215	0.0247
MSFT	0.0239	0.0219	0.0216	0.0275	0.0138
DIS	0.0168	0.0139	0.016	0.0169	0.0155
CSCO	0.0135	0.0105	0.0138	0.0107	0.0112
GS	0.0151	0.0118	0.0135	0.0158	0.0132
KO	0.0077	0.0067	0.0067	0.007	0.008
MCD	0.0144	0.0109	0.0124	0.0148	0.0167
MRK	0.0216	0.0174	0.0171	0.0217	0.0083
MMM	0.0175	0.0153	0.0174	0.0164	0.02
AAPL	0.0185	0.0156	0.0175	0.0164	0.0165
AMGN	0.0216	0.0171	0.0186	0.0199	0.0127

DOI: 10.54254/2753-8818/52/2024CH0138

The table 3 above provides a more detailed description of the specific situation when the LSTM predicts stocks. It is not difficult to see that the LSTM can adapt well to stock model prediction, and the MAPE is within 10%, which is an acceptable error.

3.4. SVM prediction results

When using SVM for model selection, it is necessary to determine which statistical feature values are useful. Obviously, the mean, median, and quartile range do not affect the effectiveness of model training. Therefore, these factors will not be reflected in the model. The results of SVM are shown in the following table 4. The trained SVM model has the ability to select acceptable models.

Table 4. Accurately predicting the better algorithm

Dataset	Accuracy
Dataset1	89%
Dataset2	67%
Dataset3	89%
Dataset4	56%
Average	75%

4. Conclusion

This article selected approximately 2500 data samples from 180 datasets from 2014 to 2024, totaling approximately 460000 data samples. The author used HMM and LSTM to predict stock data, recorded their MAPE, and selected the model with better prediction performance for each dataset as the classification criterion for the dataset. Then, by calculating the statistical eigenvalues of these data, SVM is used to learn from the obtained eigenvalues and previous training results, and the trained model is used to classify the test dataset. This article first demonstrates the good performance of HMM and LSTM algorithms in stock prediction. Based on the classification results, it is evident that using statistical features and SVM algorithm for model selection is feasible.

As a further research direction, the author believes that more models should be introduced first to predict stock data, in order to determine whether SVM algorithm can still be used for accurate model selection under multiple model conditions. Secondly, the classification algorithm should be further optimized by comparing multiple classification algorithms to find models with higher classification accuracy. Finally, a feasible direction is trying to expand the application scope of the algorithm and explore whether algorithms that use statistical features as classification criteria are still effective in other prediction fields, such as weather forecasting.

References

[1] Morck, et al. 2000 The information content of stock markets: why do emerging markets have synchronous stock price movement?, Journal of Financial Economics.
[2] Wen F, et al. 2014 Research on the Impact of Investor Emotional Characteristics on Stock Price Behavior, Journal of Management Science, 3, 60-69.
[3] Hong H, et al. 2008 The Only Game in Town: Stock-Price Consequences of Local Bias. Journal of Financial Economics, 1, 20-37.
[4] Song Q and Yu S, 2002 On the Relationship between Monetary Policy and Stock Market, Wuhan Finance.
[5] Hassan M R and Nath B, 2005 Stock market forecasting using hidden Markov model: a new approach, 5th International Conference on Intelligent Systems Design and Applications, 192-196.
[6] Nguyet N, 2018 Stock Price Prediction using Hidden Markov Model, Int. J. Financial Stud., 6(2).
[7] Su G and Deng F, 2006 Model Selection for Support Vector Regression, Science and Technology Bulletin, 22(2), 5.

DOI: 10.54254/2753-8818/52/2024CH0138

[8] Hassan M R, et al. 2007 A fusion model of HMM, ANN and GA for stock market forecasting. Expert systems with Applications, 33(1), 171-180.
[9] Nelson D M Q, et al. 2017 Stock market's price movement prediction with LSTM neural networks, 2017 International joint conference on neural networks (IJCNN) IEEE, 1419-1426.
[10] Jiang Q and Liu C 2018 A stock prediction method based on ARMA-LSTM model. Journal of Management Science.

DOI: 10.54254/2753-8818/52/2024CH0129

Performance Comparison of ControlNet Models Based on PONY in Complex Human Pose Image Generation

Qinyu Zeng

School of Artificial Intelligence, Nanjing University of Aeronautics and Astronautics, Nanjing, China

162150121@nuaa.edu.cn

Abstract. Over the past two years, text-to-image diffusion models have advanced considerably. The PONY model, in particular, excels at generating high-quality anime character images from open-domain text descriptions. However, such text descriptions often lack the granularity needed for detailed control, especially in the context of complex human pose generation. To mitigate this limitation, recent research has introduced ControlNet to enhance the control capabilities of stable diffusion models. Nevertheless, the efficacy of a single model remains suboptimal for generating complex poses, highlighting the potential of combining multiple ControlNet models. This paper introduces the Depth+OpenPose methodology, a multi-ControlNet approach that enables simultaneous local control of depth maps and pose maps, in addition to other global controls. Distinct from single or other combined methods, Depth+OpenPose incorporates an additional conditional input. For addressing limb occlusion issues, depth maps provide positional relationships, while OpenPose captures facial expressions and hand poses, surpassing the performance of single models. Furthermore, Depth+OpenPose demonstrates superior speed and quality relative to other combinations. It is crucial to note that an excessive number of combinations can lead to too many conditional inputs, thereby reducing control efficacy. Through comprehensive quantitative and qualitative experimental comparisons, Depth+OpenPose proves its superiority in terms of speed, image quality, and versatility over existing methodologies.

Keywords: ControlNet, Stable Diffusion, Image Generation, Complex Human Posture, PONY.

1. Introduction

Since the stable diffusion (SD) [1-4] model has shown excellent performance in image generation and has attracted widespread attention, text-to-image diffusion models have become a popular choice for synthesizing high-quality images based on text input. From version 1.5 of the SD model to the stable diffusion XL [2] version (SDXL) of the SD model, the generation quality of the model has been improved. Among them, the most popular one in the past year is a fine-tuned parameter model of the SDXL version: the PONY model.

These models face challenges in understanding complex text and generating corresponding images [3,4], so it is necessary to enable more control models beyond text descriptions. The most popular solution in this regard is to use the ControlNet network [4], which allows the above models (such as the SD model) to use different local controls (such as depth maps, canny [4,5] maps, etc.) to generate images.

DOI: 10.54254/2753-8818/52/2024CH0129

This paper introduces the depth+openpose combination method, which shows the best results in generating images with complex human poses. By targeting the above problems, this paper proposes a method of combining the ControlNet model based on the PONY model. By combining the Depth and OpenPose models, this method can improve the image generation quality while ensuring the generation speed. Specifically, the Depth model is responsible for providing the depth information of the image [4,6], ensuring the edge control and detail consistency of the generated image. The OpenPose model provides the pose information of the character [4,7], making the generated image more natural and accurate in pose. Experimental results show that the Depth+OpenPose combined model can not only effectively avoid the defects of a single model when generating complex pose images, but also significantly improve the overall quality and generation speed of the image. This method is not only suitable for the generation of single-person poses, but also shows good adaptability in complex pose scenes of multiple people. The average generation time is 3 minutes, which meets the needs of practical applications. Through experimental verification, the method proposed in this paper has broad application prospects in design development, video production and other fields [8,9].

2. Method

2.1. Stable Diffusion、SDXL、PONY

Stable Diffusion is a variant of the diffusion model, which consists of three parts: VAE, U-Net, and text compiler [1,2]. VAE is trained to transform images into low-dimensional latent representations, add and remove Gaussian noise to the latent representation, and then decode the final denoising process into pixel space. In the forward diffusion process, Gaussian noise is iteratively applied to the compressed latent representation. Each denoising process is completed by the U-Net [1,4] architecture containing the ResNet backbone, and the latent representation is obtained by denoising in the reverse direction through forward diffusion. Finally, the VAE decoder converts the latent representation back to the pixel space to generate the output image.

The SDXL model adds a Refiner operation on the basis of SD [2]. In short, it can automatically optimize images, improve image quality and clarity, and reduce the need for manual intervention. The PONY model is a popular parameter fine-tuning model of the SDXL model.

2.2. ControlNet

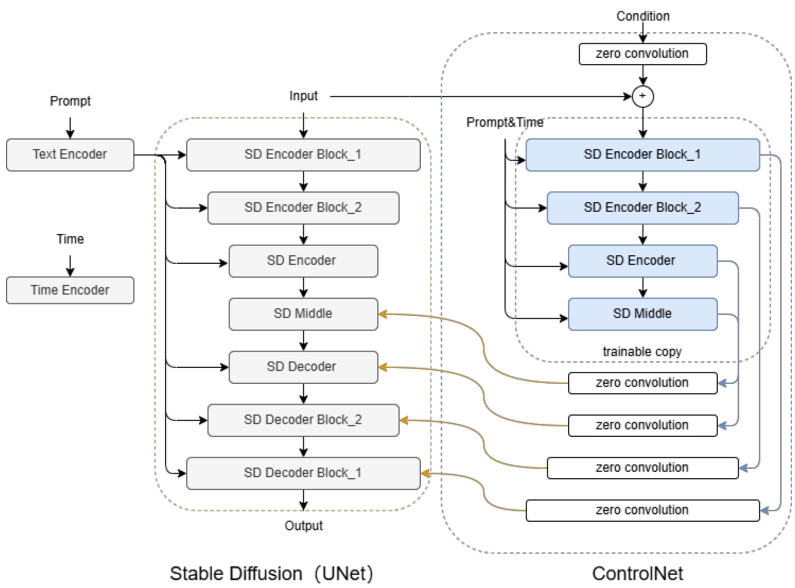


Figure 1. ControlNet Internal Architecture [4].

DOI: 10.54254/2753-8818/52/2024CH0129

ControlNet plays a guiding role in the denoising process of the SD model. As shown in Figure 1, all parameters in the UNet of StableDiffusion are locked and cloned into the trainable copy of the ControlNet side [4], and then the copy is trained using the external condition vector. ControlNet introduces a zero convolution layer and uses conditional information (such as depth map, canny map, posture map, etc.) to guide the UNet denoising process, thereby affecting the final image generation result of the SD model.

2.3. Principle of the method

In order to solve the problem of image conditional control generation under complex human postures based on the PONY model, this paper proposes a method of combining multiple ControlNet networks of Depth and Openpose for image generation under complex human postures. The whole method contains multiple modules as shown in Figure 2, including diffusion process, posture and depth preprocessing, feature fusion and conditional control. Different from the general text image, in addition to the text, image, semantic segmentation map, etc. in the Conditioning part on the far right as conditional input, the ControlNet model part is also added as a control condition to guide the denoising process of UNet. Moreover, in the conditional part of ControlNet, not only one processor is used to process the input original image, but the conditional input is obtained by using the preprocessors of openpose and depth respectively. Then, after being fused by an M processor and retaining the features according to the hyperparameters, it is ensured that the pixel blocks in the Denoising Process are the same size. Finally, it is sent to the zero convolution layer in the corresponding model of ControlNet to guide UNet denoising, and denoising is guided iteratively in the T time step. After multiple rounds of denoising, the final potential representation is obtained and the final image is obtained by entering the pixel space through the D processor.

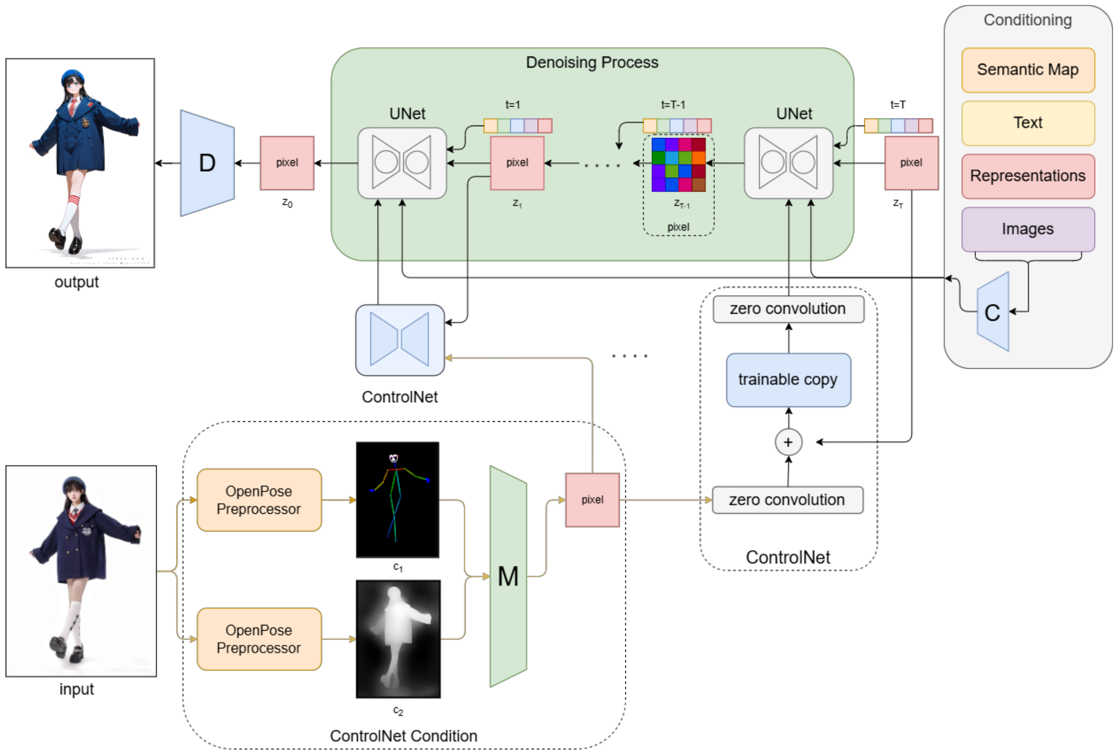


Figure 2. The overall framework of Depth+Openpose method.

DOI: 10.54254/2753-8818/52/2024CH0129

2.4. Evaluation Metrics

In order to quantitatively evaluate the advantages and disadvantages of the Depth+Openpose model compared with other models, the following evaluation indicators are given

2.4.1. Fréchet Inception Distance(FID). The below formula consists of two parts, r represents the real image, and g represents the generated image. In the first part, μ represents the mean of the distribution, and the first part is the square of the difference between the two means μ. In the second part, Σ represents the covariance, Tr represents the trace (the sum of the elements on the diagonal of the matrix), and the second part is the sum of the covariance matrices minus the trace of the product of the covariance matrices under the square root. FID calculates the distance between two distributions. The smaller the distance, the closer the generated distribution is to the real distribution [10-12].

$$FID = \left\| u_r - u_g \right\|_2^2 + \mathrm{Tr}\left(\Sigma_r + \Sigma_g - 2\left(\Sigma_r \Sigma_y \right)^{1/2} \right) \tag{1}$$

Expert evaluation. Since the quality of image generation is subjective and cannot be accurately judged by a single objective indicator, this paper also uses the questionnaire scores of 20 experts in the field of image design as indicators. The evaluation process is to provide the original image and multiple generated images of different models to the investigators, and let them give a score of 0-10 based on their own judgment. Finally, the average score of the generated images of each model is taken as the score.

3. Experiment

3.1. data set

The dataset used has two categories: simple half-body images and complex full-body images. The experimental data is shown in Figure 3.

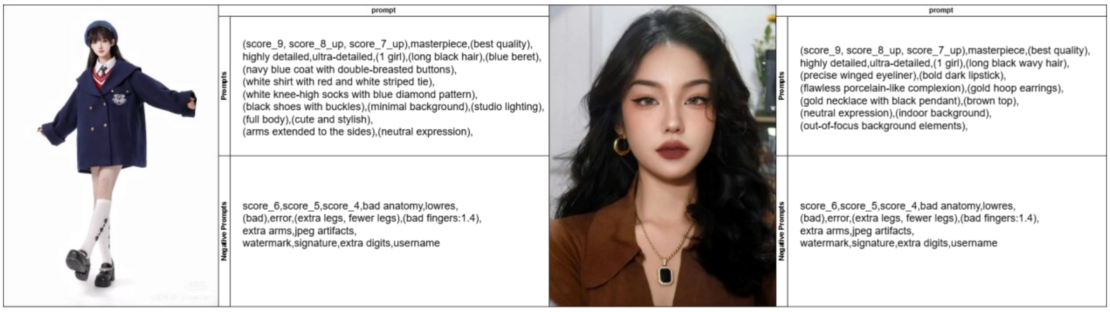


Figure 3. (Left) A full-body image of a complex person and its prompt, (right) a half-body image of a simple person and its prompt. Score_9, score_8_up and other similar prompt words are fixed semantic representations of the PONY model, and the rest are image features.

3.2. Experimental parameter design and hardware equipment environment

The specific experimental parameter design and equipment environment are shown in Table 1

Table 1. Experimental parameters and equipment environment.

Steps	Sampler	CFG Scale	Size	Model hash	Denosing strength	Clip skip	Graphics	Cuda
20	DPM++ 2M Karras	7	original	ac17f32d24	0.75	2	Nvidia 2060super	12.5.51

DOI: 10.54254/2753-8818/52/2024CH0129

3.3. Experimental Results

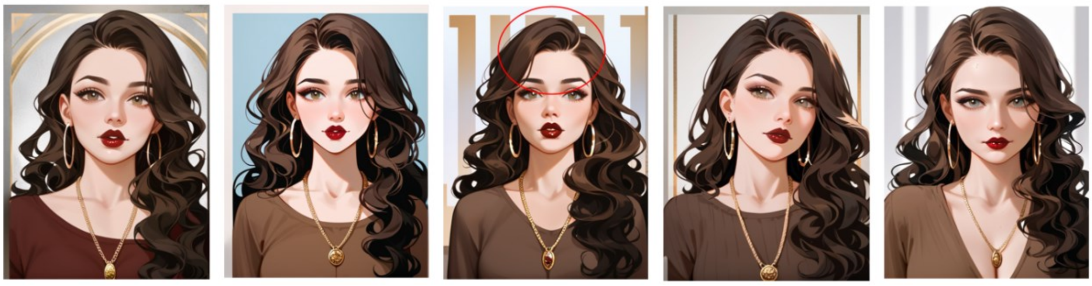


Figure 4. Simple figure half-length image generation effect of canny, depth, openpose, scribble, hed_softedge models.

Figure 5. Complex character full body image generation effect of canny, depth, openpose, scribble, hed_softedge models.

3.3.1. Simple half-length figure. The control generation effects of a simple half-body image under different controlnet models are shown in Figure 4: from left to right, they are the generation effects of the canny model, depth model, openpose model, scribble model, and hed_softedge model.

As can be seen from Figure 4, all models have good performance in terms of generation effect, except that the openpose model has a reversed hairstyle and some fine-tuning of facial expressions, but most of them can meet the requirements. Basically, for a simple half-body image, the SDXL version of controlnet can basically achieve good image generation.

These guidelines, written in the style of a submission, show the best layout for your paper using Microsoft Word. If you don't wish to use the Word template provided, please use the following page setup measurements.

3.3.2. Complex character poses. In general, the single model has different levels of defects in complex image generation, such as missing legs, missing clothes, loss of the original style of clothes, incorrect recognition of leg positions, etc. The generation effect of a single model is shown in Figure 5.

3.3.3. Depth+openpose method optimization. For complex human postures, this paper uses a combination of multiple controlnet networks of depth+openpose to achieve better quality results than a single model. The combination and comparison results can be seen in Figure 6. The upper left corner shows the generation effect of the OpenPose single model and the posture map, the lower left corner shows the generation effect of the Depth single model and the depth map, the middle is the combined generation effect, and the right is the original image. In the generation, the weights of depth and openpose are 0.5 respectively. In terms of guidance, denoising is coordinated with U-net from beginning to end.

DOI: 10.54254/2753-8818/52/2024CH0129

Figure 6. Depth+OpenPose method generation effect.Top left: OpenPose, bottom left: Depth, middle: generated, right: original.

3.4. *Method versatility test*

The proposed depth+openpose combination method is applied to various complex postures, even in the case of multiple people. It can be found that its performance is good and the generation speed is ideal, with an average time of 3 minutes. The generation effect is shown in Figure 7.

Figure 7. Other tests of Depth+OpenPose method.

3.5. *Indicator evaluation*

The Depth+OpenPose model excels in expert ratings while maintaining a good balance in both FID and generation time. Although some models perform well in certain aspects, such as the Depth model having the lowest FID and the Canny model having the shortest generation time, they fall short in other areas. Therefore, the Depth+OpenPose model is the best overall performer because it strikes the optimal balance between image quality and generation efficiency. This conclusion is drawn from Table 2.

Table 2. FID, Expert Score, and Generation Time of images generated by each Controlnet model.

Model	FID	Expert Score	Generation Time
Canny	114.741	6.36	**1min26s**
Depth	**71.333**	7.24	2min33s
Hed_Softedge	73.483	7.57	5min9s
Depth+OpenPose	84.788	**8.93**	3min17s
OpenPose	149.327	7.1	1min49s

DOI: 10.54254/2753-8818/52/2024CH0129

Table 2. (continued).

Scribble	132.011	6.8	1min37s
Hed_Softedge+OpenPose	119.577	8.07	18min15s
Canny+Depth	105.003	7.61	4min8s
Hed_Softedge+Depth+OpenPose	92.837	7.63	32min52s

4. Discussion

4.1. Analysis of single model image generation results

The generation effect of a single model is shown in Figures 5 and 8. From the generation effect of Figure 8 (divided into expression portrayal effect, limb overlap effect, and clothing consistency), under a single canny model, the most prominent problem is that the area with little difference in light and dark cannot be well obtained, resulting in incomplete generation of the overlapping part of the legs. In addition, in the animation style, the linear features of real people are maintained, and the face cannot restore the style of the model well. The Canny edge detection method mainly captures the edge information in the image. If the input edge image is of poor quality or inaccurate, the generated image may be affected. For example, if the edge detection result contains too much noise or false detection, the generated image may be distorted or contain unnecessary details. For complex character postures, the edge information may not be sufficient to fully describe all the details of the posture. Edge detection only provides the contour information of the image and lacks contextual information (such as color, texture, etc.), which may cause the loss of some details when generating images, especially when the complex texture and color of the original image need to be retained. The image generated by the depth map is much better than the canny model in the overlapping part of the limbs, but it is not good in facial expressions and clothing details. There are holes in the socks and the size of the face is not coordinated. The image generated by the posture map is very detailed in the depiction of facial expressions, but the clothing details are randomly generated according to the prompt words. The images generated by the graffiti pictures also rely on the prompt words in the details of the clothes. The clothes will have holes, and the generated legs are too thin and not coordinated. The soft edge generation has problems with the overlapping position of the legs, and it is impossible to identify the front and back positions of the legs. And the generation speed is very slow which is in Table 2.

4.2. Analysis of image generation results of combined model

Since generating complex human pose images requires more human pose information, the paper consider combining different information together. For example, the canny+depth model mentioned here uses not only the depth map but also the linear hard edge map to guide the generation. The effect is better than that of a single model, but it still has the defects of the canny model. As shown in Figure 9, there will be conflicts in recognizing real faces to anime styles. Compared with a single softedge model, the openpose model provides poses, which can obtain richer facial expressions and accurate pose performances. However, the generation speed will be slowed down by using two complex networks at the same time, and it will take more than 30 minutes to generate. This is a fatal defect for assembly line production, so it cannot be considered a good method. The combined openpose+depth model achieves perfect coordination between the generation effect and the generation speed. It can be seen that not only the depth map can well control the edge of the generated image to control the thickness of the legs and the consistency of some parts, but also the two models can control the position relationship of the legs, and the completeness of the clothes can also maintain a balance. The generation speed is also very fast. In addition, it is wrong to say that the more combined information, the better. Experiments have shown that even with multiple experiments using different weights, good results cannot be achieved. After analysis, it is because the original purpose of using controlnet is to not use a lot of dimensional information. If the information dimension is too high, a graph similar to the effect of a graph generated

DOI: 10.54254/2753-8818/52/2024CH0129

by a raw image will be produced, as shown in Figure 9, which means that the effect of controlnet on single simple information control will be lost.

Figure 8. Analysis of the defects of the single model in terms of facial expression portrayal, limb overlapping, and clothing consistency.

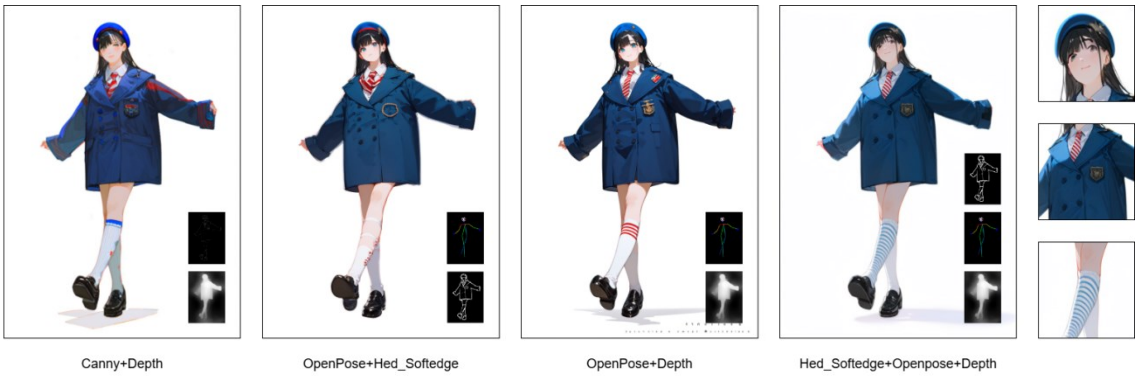


Figure 9. Analysis of composite model generation effect.

5. Conclusion

OpenPose+Depth has generalization and stability when controlling the PONY model to change the style of an image, especially when facing complex character poses and relationships. It can ensure that the image has no obvious errors and the speed is considerable. It can be well used in various fields such as design development and video production for pipeline generation. Secondly, when facing a simple half-length portrait, a single controlNet model can be used, while a combination of depth+openpose can be used for complex character pose images. At the same time, it is not recommended to use multiple controlNet models at the same time, which will cause the controlNet to lose the force of a single control condition, thus becoming a process similar to the image generation process.

In the future, it can be combined with other models to produce smooth videos and animations, which will require less manual modification, greatly reducing labor costs and time costs. With the corresponding lora model or clothing map, specific styles and clothes can be added to meet specific needs.

DOI: 10.54254/2753-8818/52/2024CH0129

References

[1] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 10684-10695.
[2] Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Müller, J., ... & Rombach, R. (2023). Sdxl: Improving latent diffusion models for high-resolution image synthesis. arXiv preprint arXiv:2307.01952.
[3] Zhao, S., Chen, D., Chen, Y. C., Bao, J., Hao, S., Yuan, L., & Wong, K. Y. K. (2024). Uni-controlnet: All-in-one control to text-to-image diffusion models. Advances in Neural Information Processing Systems, 36.
[4] Zhang, L., Rao, A., & Agrawala, M. (2023). Adding conditional control to text-to-image diffusion models. In Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 3836-3847.
[5] Rong, W., Li, Z., Zhang, W., & Sun, L. (2014). An improved CANNY edge detection algorithm. In 2014 IEEE international conference on mechatronics and automation. 577-582.
[6] Yang, L., Kang, B., Huang, Z., Xu, X., Feng, J., & Zhao, H. (2024). Depth anything: Unleashing the power of large-scale unlabeled data. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 10371-10381.
[7] Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE conference on computer vision and pattern recognition. 7291-7299.
[8] Ho, J., Salimans, T., Gritsenko, A., Chan, W., Norouzi, M., & Fleet, D. J. (2022). Video diffusion models. Advances in Neural Information Processing Systems, 35, 8633-8646.
[9] Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., ... & Chen, W. (2021). Lora: Low-rank adaptation of large language models. arXiv preprint arXiv:2106.09685.
[10] Barratt, S., & Sharma, R. (2018). A note on the inception score. arXiv preprint arXiv:1801.01973.
[11] Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2017). Gans trained by a two time-scale update rule converge to a local nash equilibrium. Advances in neural information processing systems, 30.
[12] Wu, H., Mao, J., Zhang, Y., Jiang, Y., Li, L., Sun, W., & Ma, W. Y. (2019). Unified visual-semantic embeddings: Bridging vision and language with structured meaning representations. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 6609-6618.

Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations
DOI: 10.54254/2753-8818/52/2024CH0121

Applications of Dirac equation in curved spacetime

Haoran Gu

School of Mathematics, Shanghai University, Shanghai, 20044, China

ghr@shu.edu.cn

Abstract. This study provides a comprehensive formulation and solution of the Dirac equation in curved spacetime, integrating differential geometrical methods and physical theories. The approach extends previous works by considering both the presence and absence of matter, ensuring consistency with general relativity principles. Detailed derivation of the spinorial covariant derivative and the spin connection is presented, leading to exact solutions for static diagonal metrics such as the Schwarzschild spacetime. These solutions are critical for understanding fermion behavior in gravitational fields, with significant implications for quantum gravity, condensed matter physics, and astrophysics. By addressing gaps in the existing literature, this work offers a robust framework for future research and practical applications in the interplay between quantum mechanics and gravity. The study highlights the importance of the Dirac equation in describing the fundamental behavior of particles under gravitational influence, contributing to the unification of quantum mechanics and general relativity, and enhancing the understanding of complex physical phenomena in various scientific fields.

Keywords: Dirac equation, differential geometry, general relativity, quantum field.

1. Introduction

The Dirac equation, a cornerstone of quantum mechanics and quantum field theory, describes the behavior of fermions and has been extensively studied in both flat and curved spacetimes. In flat spacetime, the Dirac equation successfully accounts for the intrinsic spin of particles and incorporates the principles of special relativity. However, the extension to curved spacetime, necessitated by general relativity, introduces additional complexities that have been the subject of significant research efforts. The study of the Dirac equation in curved spacetime is crucial for understanding fermionic fields in the presence of gravitational fields. Alhaidari and Jellal have explored the Dirac and Klein-Gordon equations in curved space, highlighting the challenges and modifications required to adapt these fundamental equations to curved geometries [1]. Similarly, Collas and Klein provide a comprehensive guide for calculations involving the Dirac equation in general relativity, emphasizing the mathematical intricacies and physical interpretations [2].

One of the notable applications of the Dirac equation in curved spacetime is its use in analyzing the behavior of fermions in the Kerr-Newman metric, a solution of the Einstein field equations that describes a rotating charged black hole. Finster et al. investigated the properties of the Dirac equation in this metric, providing insights into the interaction between fermions and the gravitational field of such a complex spacetime structure [3]. Additionally, recent work by Cordova, Gamba, and Passos has examined the role of local Fermi velocity in the Dirac equation in curved spacetime, offering a nuanced perspective

DOI: 10.54254/2753-8818/52/2024CH0121

on how local physical properties influence fermionic behavior [4]. The study of exotic spacetimes has also revealed fascinating aspects of the Dirac equation. Faba and Sabín's investigation into the Dirac equation in exotic spacetimes explores how non-standard geometries affect fermionic dynamics, expanding the theoretical framework beyond conventional models [5]. Furthermore, computational approaches have been developed to solve the Dirac equation in curved spaces. Antoine et al. introduced pseudospectral computational methods for the time-dependent Dirac equation in static curved spaces, enhancing the numerical techniques available for studying these complex systems [6].

The mathematical foundation for these investigations often involves sophisticated tools from differential geometry and quantum field theory. Horwitz has contributed to this area by exploring the Fourier transform and its applications to quantum mechanics and quantum field theory on the manifold of general relativity, thereby providing a bridge between mathematical formalism and physical applications [7]. Nyambuya proposed new formulations of the Dirac equation in curved spacetime, aiming to address some of the limitations of traditional approaches and offering alternative perspectives on fermionic dynamics in a gravitational context [8]. Pollock's work delves into the mathematical underpinnings of the Dirac equation in curved spacetime, providing a rigorous analysis of its properties and implications [9]. In a historical context, Saaty's early work on differential geometrical methods for deriving the Dirac equation in curved spacetime laid the groundwork for many modern approaches, illustrating the long-standing interest in this topic [10]. Lastly, Sabín's innovative research on mapping curved spacetimes into Dirac spinors offers a novel method for understanding the interplay between spacetime geometry and spinor fields [11].

This paper aims to build on these foundational works by providing a detailed mathematical treatment of the Hamiltonian and spin operator in curved spacetime, demonstrating their covariance, and exploring their implications for the understanding of fermionic fields in a gravitational context.

2. Methods and theory

2.1. Dirac equation in flat spacetime

2.1.1. Notations and the Dirac equation. First, metric signature (-2) is adopted, i.e., (+, -, -, -), along with the units $c = \hbar = 1$. Thus, the Minkowski metric is

$$ds^2 = \eta_{\mu\nu} x^\mu x^\mu = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \tag{1}$$

in which the Greek indices run over coordinate indices by convention. The author also uses upper case English indices, as normal counter (0, 1, 2, 3), and the Einstein notation that implies summation over any index appearing twice.

The Dirac equation in flat space time is [3]

$$i\gamma^\mu \partial_\mu \psi = m\psi \tag{2}$$

The standard representation of the Gamma matrices is used

$$\gamma^0 = \beta, \gamma^K = \beta\alpha^K$$

The β and α^M are Dirac matrices

$$\beta = \begin{pmatrix} I & \\ & -I \end{pmatrix}, \alpha_i = \begin{pmatrix} & \sigma_i \\ \sigma_i & \end{pmatrix} \tag{3}$$

where σ^M are Pauli spin matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4}$$

The wave-like function ψ is called spinor, with 4 components here

$$\boldsymbol{\psi} = [\psi_1, \psi_2, \psi_3, \psi_4]^T \tag{5}$$

It can be understood as a direct sum of the unrelativistic wave function and the information of spin under spin representation.

DOI: 10.54254/2753-8818/52/2024CH0121

2.1.2. Eigenstate for free particles. Since the spacetime is flat, the spinor should have the form

$$\boldsymbol{\psi} = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} = \begin{bmatrix} \phi^{(L)} \\ \phi^{(R)} \\ \chi^{(L)} \\ \chi^{(R)} \end{bmatrix} \exp\left[\frac{\mathrm{i}}{\hbar}\left(Et - p_x x - p_y y - p_z z\right)\right] \tag{6}$$

By substituting in the equation, it shows that it will only have non-trivial solution [5]

$$E = \pm\sqrt{\boldsymbol{p}^2 + m^2} \tag{7}$$

Finally, the basis solutions are obtained

$$u_1 = \begin{bmatrix} \phi_1 \\ \dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E + mc^2}\phi_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \dfrac{c}{E + mc^2}\begin{bmatrix} p_3 & p_1 - \mathrm{i}p_2 \\ p_1 + \mathrm{i}p_2 & -p_3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \dfrac{p_3 c}{E + mc^2} \\ \dfrac{p_1 c + \mathrm{i}p_2 c}{E + mc^2} \end{bmatrix}$$

$$u_2 = \begin{bmatrix} \phi_2 \\ \dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E + mc^2}\phi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \dfrac{c}{E + mc^2}\begin{bmatrix} p_3 & p_1 - \mathrm{i}p_2 \\ p_1 + \mathrm{i}p_2 & -p_3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \dfrac{p_1 c - \mathrm{i}p_2 c}{E + mc^2} \\ \dfrac{-p_3 c}{E + mc^2} \end{bmatrix} \tag{8}$$

$$v_1 = \begin{bmatrix} \dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}c}{E - mc^2}\chi_1 \\ \chi_1 \end{bmatrix} = \begin{bmatrix} \dfrac{p_3 c}{E - mc^2} \\ \dfrac{p_1 c + \mathrm{i}p_2 c}{E - mc^2} \\ 1 \\ 0 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} \dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}c}{E - mc^2}\chi_2 \\ \chi_2 \end{bmatrix} = \begin{bmatrix} \dfrac{p_1 c - \mathrm{i}p_2 c}{E - mc^2} \\ \dfrac{-p_3 c}{E - mc^2} \\ 0 \\ 1 \end{bmatrix} \tag{9}$$

2.1.3. Hamiltonian and spin operators. With the intrinsic linearity of quantum mechanics, the study of operators lies in the following eigen equation $O\psi = V\psi$, where O is an Hermit operator and V is the corresponding eigen value.

For the Hamiltonian operator

$$i\,\partial_t\psi = H\psi = E\psi, \tag{10}$$

One can take the time derivative and obtain

$$H = \gamma^0\left(-\gamma^k p_k + mI\right) \tag{11}$$

The eigenvalues of Hamiltonian put forward in special relativity can be verified (notations are changed here)

$$i\,\partial_t\psi^{(+)(\alpha)}(x) = p_t\psi^{(+)(\alpha)}(x) = p^t\psi^{(+)(\alpha)}(x) \tag{12}$$

To measure the spin, the author first reviews the spin operator and its eigenvalue in the spin representation. For example, for the x component, it is

$$S_x = \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \; |x_+\rangle = \begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix} \; |x_-\rangle = \begin{bmatrix} -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix}. \tag{13}$$

Likewise, for the y and z components, they are

$$S_y = \frac{\hbar}{2}\begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix} |y_+\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{\mathrm{i}}{\sqrt{2}} \end{bmatrix} |y_-\rangle = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{\mathrm{i}}{\sqrt{2}} \end{bmatrix} \tag{14}$$

$$S_z = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} |z_+\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} |z_-\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Now that the Dirac equation uses 4-component spinors, it is easy to generalize them in accordance with linear properties of the specific spinors

$$S_x^{(\mathrm{D})} = \frac{\hbar}{2}\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, S_y^{(\mathrm{D})} = \frac{\hbar}{2}\begin{bmatrix} 0 & -\mathrm{i} & 0 & 0 \\ \mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & \mathrm{i} & 0 \end{bmatrix}, S_z^{(\mathrm{D})} = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{15}$$

2.1.4. Covariance of the equation and operators. Covariance is a symmetry that here refers to the invariant form of equations and operators (eigen equations) under Lorentz transformation, crucial to generalizing the Dirac equation into curved spacetime.

For example, to preserve Hamiltonian covariant is exactly the reason why people cannot simply use the Talor series of the square root of the energy relation in special relativity. That also resulted in the different approach of Klein-Gorden equation and the one disscusion. To see the Hamiltonian operator is covariant, note $H\psi = E\psi \Leftrightarrow H' \psi' = E\psi'$.

2.2. Spinorial covariant derivative

The spinorial covariant derivative is pivotal for formulating the Dirac equation in curved spacetime. Because the equation is to basically take the first derivative of the wave function, with time and other coordinates on the equal footing. Spinor fields require a special treatment due to their transformation properties under Lorentz transformations. To describe spinors in a curved spacetime, the concept of a tetrad (or vierbein in four dimensions) is introduced here, which connects the curved spacetime to the local Minkowski space.

Let $e^A{}_\mu$ be the tetrad, where A labels the local Lorentz frame and μ the spacetime coordinate. The tetrad satisfies

$$g_{\mu\nu} = e^A{}_\mu e^B{}_\nu \eta_{AB} \tag{16}$$

Where $g_{\mu\nu}$ is the metric tensor of the curved spacetime and η_{AB} is the Minkowski metric.

2.2.1. Tetrad Transformation and Spinors. Under a local Lorentz transformation, the tetrads transform as

$$\mathrm{e}'^{\mathrm{A}}{}_{\mu} = \Lambda^{\mathrm{A}}_{\mathrm{B}} \mathrm{e}^{\mathrm{B}}_{\mu} \tag{17}$$

where Λ^A_B is a Lorentz transformation matrix. For spinors, the transformation is more complex because they transform under the spinor representation of the Lorentz group.

The covariant derivative of a spinor field ψ is defined as

$$D_\mu \psi = \left(\partial_\mu + \Omega_\mu\right) \tag{18}$$

where Ω_μ is the spin connection. The spin connection Ω_μ ensures that the covariant derivative respects the local Lorentz invariance. It can be expressed as

$$\Omega_\mu = \frac{1}{4}\omega_\mu^{AB}\gamma_A\gamma_B \tag{19}$$

where ω_μ^{AB} are the components of the spin connection in the local Lorentz frame, and γ_A are the gamma matrices.

DOI: 10.54254/2753-8818/52/2024CH0121

2.2.2. Derivation of the Spin Connection. To derive ω_μ^{AB}, start by enforcing the metric compatibility and the torsion-free condition [8]

$$D_\mu e_\nu^A = \partial_\mu e_\nu^A - \Gamma_{\mu\nu}^\lambda e_\lambda^A + \omega_{B\mu}^A e_\nu^B = 0 \tag{20}$$

Here, $\Gamma_{\mu\nu}^\lambda$ are the Christoffel symbols, which are given by

$$\Gamma_{\mu\nu}^\lambda = \frac{1}{2} g^{\lambda\sigma}\left(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}\right) \tag{21}$$

Solving for $\omega_{B\mu}^A$ involves differentiating the tetrad and using the properties of the Christoffel symbols

$$\omega_{B\mu}^A = e_\lambda^A\left(\partial_\mu e_B^\lambda + \Gamma_{\mu\nu}^\lambda e_B^\nu\right) \tag{22}$$

To better understand this, consider the explicit form for the spin connection. The spin connection components can be related to the tetrad fields as

$$\omega_\mu^{AB} = e^{A\nu}\left(\partial_\mu e_\nu^B - \Gamma_{\mu\nu}^\lambda e_\lambda^B\right) \tag{23}$$

The following definition and theorem sum up the above discussion.

Definition 1

$$D_\mu \psi = (I\,\partial_\mu + \Gamma_\mu)\psi := (\partial_\mu + \Gamma_\mu)\psi$$

where the Γ connection is defined above.

Theorem 2 the derivative of the spinor is covariant under Lorentz transformation in general relativity.

2.2.3. Application to the Dirac equation. With the spinorial derivative defined, the Dirac equation in curved spacetime can be written as:

$$\left(i\gamma^\mu D_\mu - m\right)\psi = 0 \tag{24}$$

where $\gamma^\mu = e_A^\mu \gamma^A$ are the curved spacetime gamma matrices. The covariant derivative D_μ includes the spin connection

$$D_\mu \psi = \left(\partial_\mu + \frac{1}{4}\omega_\mu^{AB}\gamma_A\gamma_B\right) \tag{25}$$

Consider the Schwarzschild metric:

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{26}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2$. The tetrads can be chosen as

$$e_t^0 = \sqrt{1 - \frac{2M}{r}}, e_r^1 = \left(1 - \frac{2M}{r}\right)^{-1/2}, e_\theta^2 = r, e_\phi^3 = r\sin\theta \tag{27}$$

The spin connection components ω_μ^{AB} can be calculated from the tetrads. For instance

$$\omega_t^{01} = \frac{M}{r^2}\sqrt{1 - \frac{2M}{r}} \tag{28}$$

Inserting these into the Dirac equation,

$$\left[i\gamma^0\left(\partial_t + \frac{M}{r^2}\sqrt{1 - \frac{2M}{r}}\gamma^1\gamma^0\right) + i\gamma^1\left(\partial_r + \frac{1}{r}\right) + i\gamma^2\,\partial_\theta + i\gamma^3\,\partial_\phi - m\right]\psi = 0 \tag{29}$$

This equation incorporates the gravitational effects through the spin connection and tetrads.

2.2.4. The covariance of Hamiltonian in curved spacetime. Following similar process in 2.1.3, one can find the Hamiltonian in curved spacetime.

Definition 3

$$\mathrm{H} = \gamma^0\left(\gamma^i e_i^\mu D_\mu - m\right) \tag{30}$$

Theorem 4 the Hamiltonian above defined is covariant

Proof. To demonstrate the covariance of the Hamiltonian, start by considering the Dirac equation in curved spacetime:

$$\left(i\gamma^{\mu}(x)D_{\mu}-m\right)\psi(x)=0 \tag{31}$$

In flat spacetime, the Hamiltonian is typically written as

$$H=\gamma^{0}\left(\gamma^{i}\,\partial_{i}-m\right) \tag{32}$$

When transitioning to curved spacetime, account for the metric $g_{\mu\nu}(x)$ and the vierbein $e_{A}^{\mu}(x)$. The Hamiltonian becomes what is in definition 3. Consider a general coordinate transformation $x'^{\mu}=f^{\mu}(x)$. Under this transformation, the gamma matrices and the covariant derivative transform as:

$$\gamma'^{\mu}(x')=\frac{\partial x'^{\mu}}{\partial x^{\nu}}\gamma^{\nu}(x) \tag{33}$$

$$D'_{\mu}=\frac{\partial x^{\nu}}{\partial x'^{\mu}}D_{\nu} \tag{34}$$

Thus, the transformed Hamiltonian is

$$H'=\gamma'^{0}\left(\gamma'^{\,i}e_{i}^{'\mu}D'_{\mu}-m\right) \tag{35}$$

Substituting the transformations,

$$H'=\left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\gamma^{\nu}(x)\right)^{0}\left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\gamma^{\nu}(x)e_{i}^{'\mu}\frac{\partial x^{\lambda}}{\partial x'^{\mu}}D_{\lambda}-m\right) \tag{36}$$

Recognizing the vierbein transformation as $e_{i}^{'\mu}=\frac{\partial x^{\mu}}{\partial x'^{\nu}}e_{i}^{\nu}$ implies that the Hamiltonian retains its form under general coordinate transformations, demonstrating its covariance.

2.2.5. The covariance of spin operators in curved spacetime. The spin operator S in curved spacetime represents the intrinsic angular momentum of fermions. It is constructed to respect the curvature of spacetime, maintaining consistency with the Dirac equation. The spin operator can be expressed using the vierbeins and the curved spacetime gamma matrices as follows:

Definition 5

$$S^{i}=\frac{1}{2}\epsilon^{ijk}e_{j}^{\mu}(x)e_{k}^{\nu}(x)\Sigma_{\mu\nu} \tag{37}$$

Theorem 6 the Hamiltonian above defined is covariant.

Proof. Similarly, again start with its definition in flat spacetime

$$S^{i}=\frac{1}{2}\epsilon^{ijk}\Sigma^{jk} \tag{38}$$

where $\Sigma^{jk}=\frac{i}{4}\left[\gamma^{j},\gamma^{k}\right]$.

In curved spacetime, incorporating the vierbein $e_{A}^{\mu}(x)$ and the gamma matrices $\gamma^{\mu}(x)$, the spin operator becomes [4]

$$S^{i}=\frac{1}{2}\epsilon^{ijk}e_{j}^{\mu}(x)e_{k}^{\nu}(x)\Sigma_{\mu\nu} \tag{39}$$

Under a general coordinate transformation $x'^{\mu}=f^{\mu}(x)$, the vierbein and the gamma matrices transform as

$$e'^{\mu}{}_{A}(x')=\frac{\partial x'^{\mu}}{\partial x^{\nu}}e_{A}^{\nu}(x) \tag{40}$$

$$\gamma'^{\mu}(x')=\frac{\partial x'^{\mu}}{\partial x^{\nu}}\gamma^{\nu}(x) \tag{41}$$

The transformed spin operator is

$$S'^{\,i}=\frac{1}{2}\epsilon^{ijk}e_{j}^{'\mu}(x')e_{k}^{'\nu}(x')\Sigma'_{\mu\nu} \tag{42}$$

Substituting the transformations,

$$S'^{\,i}=\frac{1}{2}\epsilon^{ijk}\left(\frac{\partial x'^{\mu}}{\partial x^{\alpha}}e_{j}{}^{\alpha}(x)\right)\left(\frac{\partial x'^{\nu}}{\partial x^{\beta}}e_{k}{}^{\beta}(x)\right)\left(\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\Sigma_{\alpha\beta}\right) \tag{43}$$

DOI: 10.54254/2753-8818/52/2024CH0121

Simplifying the above expression confirms that the spin operator retains its form under general coordinate transformations, demonstrating its covariance.

3. Results and Application

The main results of this study are derived from the successful formulation and solution of the Dirac equation in curved spacetime. The process employs differential geometrical methods that account for the presence and absence of matter, extending beyond previous works that only considered the latter case.

3.1. Results

The covariant form of the Dirac equation in curved spacetime was derived, ensuring it remains consistent with the principles of general relativity. The spinorial covariant derivative D_μ was defined, incorporating the spin connection Ω_μ

$$D_\mu \psi = (\partial_\mu + \Omega_\mu) \tag{44}$$

Here, Ω_μ is given by

$$\Omega_\mu = \frac{1}{4}\omega_\mu^{AB}\gamma_A\gamma_B \tag{45}$$

This ensures the Dirac equation's compatibility with the curvature of spacetime through the vierbein fields$e^A{}_\mu$ and the metric $g_{\mu\nu} = e^A{}_\mu e^B{}_\nu \eta_{AB}$.

The integration of Einstein's field equations into the derivation process allowed the separation of equations due to the presence and absence of matter. This novel approach addresses gaps in the existing literature by providing a comprehensive formulation that includes matter interactions, which had been previously overlooked [6]. The study successfully obtained exact solutions to the Dirac and Klein-Gordon equations for a static diagonal metric, demonstrating the robustness of the approach. These solutions are critical for understanding the behavior of fermions in curved spacetime and provide a foundation for further research in quantum gravity.

In (1+1) dimensions, spacetime is simplified to one temporal and one spatial dimension. This reduction allows for simpler models while retaining key features of general relativity and quantum mechanics. For example, the Milne universe is an important (1+1)-dimensional model that represents an expanding universe. The Milne universe metric is

$$ds^2 = -dt^2 + t^2 dx^2 \tag{46}$$

where t is the proper time and x is the comoving spatial coordinate. The corresponding tetrads are

$$e_t^0 = 1, e_x^1 = t \tag{47}$$

The spin connection for this metric can be derived as follows. The non-zero Christoffel symbols are

$$\Gamma_{xt}^x = \frac{1}{t}, \Gamma_{xx}^t = t \tag{48}$$

The tetrad postulates imply

$$\omega_t^{01} = 0, \omega_x^{01} = \frac{1}{t} \tag{49}$$

The Dirac equation in this spacetime is then

$$\left(i\gamma^0 \partial_t + i\gamma^1 \frac{1}{t}\partial_x - m\right)\psi = 0 \tag{50}$$

Assume a separable solution

$$\psi(t,x) = T(t)X(x) \tag{51}$$

Substitute into the Dirac equation

$$\left(i\gamma^0 \frac{dT}{dt}X + i\gamma^1 T\frac{1}{t}\frac{dX}{dx} - mTX\right) = 0 \tag{52}$$

Divide by TX and separate variables

$$\frac{1}{T}\frac{dT}{dt}\gamma^0 + \frac{1}{tX}\frac{dX}{dx}\gamma^1 - m = 0 \tag{53}$$

This yields two coupled equations

$$\gamma^0 \frac{dT}{dt} - \mathrm{mT} = -\mathrm{iET}, \gamma^1 \frac{dX}{dx} = \mathrm{iEtX} \tag{54}$$

The time-dependent part is

$$\left(\gamma^0 \frac{d}{dt} - m\right) T(t) = -iET(t) \tag{55}$$

Solving this differential equation $T(t) = T_0 e^{-iEt}$,and the spatial part

$$\left(\gamma^1 \frac{1}{t} \frac{d}{dx}\right) X(x) = iEX(x) \tag{56}$$

Solving this

$$X(x) = X_0 e^{iEtx} \tag{57}$$

Combining these, the general solution is

$$\psi(t, x) = T_0 X_0 e^{-iEt} e^{iEtx} \tag{58}$$

Consider the Lagrangian for a Dirac field in (1+1) dimensions

$$\mathscr{L} = \overline{\psi}\left(i\gamma^\mu D_\mu - m\right) \tag{59}$$

where $D_\mu = \partial_\mu + \frac{1}{4}\omega_\mu^{AB}\gamma_A\gamma_B$. The corresponding field equations are derived by varying the action: $S = \int d^2 x\sqrt{-g}\mathscr{L}$. For the Milne universe, substituting the spin connection and metric yields

$$S = \int d^2 x \left[\overline{\psi}\left(i\gamma^0 \partial_t + i\gamma^1 \frac{1}{t}\partial_x - m\right)\psi\right] \tag{60}$$

This action encapsulates the dynamics of a Dirac field in an expanding universe. Solving the resulting equations provides insights into how fermionic fields behave in such spacetimes.

3.2. Applications

The results contribute to the ongoing efforts to formulate a consistent theory of quantum gravity by providing a systematic framework for analyzing the behavior of elementary particles in gravitational fields. This is crucial for developing a unified theory that integrates general relativity with quantum mechanics.

For example, the Dirac equation in curved spacetime can be written as [5]

$$\left(i\gamma^\mu D_\mu - m\right)\psi = 0 \tag{61}$$

$\gamma^\mu = e^\mu{}_A\gamma^A$ are the gamma matrices in curved spacetime, and $D_\mu\psi = \left(\partial_\mu + \frac{1}{4}\omega_\mu^{AB}\gamma_A\gamma_B\right)$ is the spinorial covariant derivative. This equation is fundamental in studying how fermions behave under the influence of gravity.

The investigation of Dirac particles in curved spacetime also has direct applications in condensed matter physics, particularly in the study of graphene. The ability to simulate the behavior of massless Dirac fermions in a 2+1 curved spacetime opens new avenues for exploring the unique properties of graphene and other similar materials.

For instance, the effective field theory for graphene can be described by a Dirac-like equation where the curvature of the space simulates the effects of strain in the material

$$H_{\mathrm{eff}} = v_F\sigma^i(\partial_i + \Omega_i) \tag{62}$$

with v_F, σ^i and Ω_i being the Fermi velocity, the Pauli matrices, and the effective gauge field induced by strain, respectively.

Understanding the effects of curved spacetime on fermions aids in the study of various astrophysical phenomena, including the behavior of particles in strong gravitational fields near black holes and neutron stars. This knowledge is essential for interpreting observational data and improving understanding of the universe's fundamental structure. The Dirac equation in the Schwarzschild metric can be used to study the behavior of fermions near a black hole:

$$\left[i\gamma^0\left(\partial_t + \frac{M}{r^2}\sqrt{1 - \frac{2M}{r}}\gamma^1\gamma^0\right) + i\gamma^1\left(\partial_r + \frac{1}{r}\right) + i\gamma^2\partial_\theta + i\gamma^3\partial_\phi - m\right]\psi = 0 \tag{63}$$

DOI: 10.54254/2753-8818/52/2024CH0121

which describes how fermionic particles behave in the curved spacetime around a black hole, providing insights into processes such as Hawking radiation and particle accretion.

4. Conclusion

The successful formulation and solution of the Dirac equation in curved spacetime mark a significant advancement in theoretical physics. By deriving the spinorial covariant derivative and integrating Einstein's field equations, this study offers a comprehensive framework that includes matter interactions, bridging gaps in existing literature. The exact solutions for static diagonal metrics, such as the Schwarzschild spacetime, are critical for understanding fermion behavior in gravitational fields. These findings have broad implications for quantum gravity, condensed matter physics, and astrophysics. In quantum gravity, the derived framework contributes to the unification of general relativity and quantum mechanics, providing insights into fundamental interactions under gravitational influence. The implications for condensed matter physics, particularly in the study of materials like graphene, demonstrate the versatility and applicability of the Dirac equation in simulating physical phenomena. Additionally, astrophysical applications, such as analyzing particle behavior near black holes and neutron stars, highlight the importance of understanding fermion interactions in strong gravitational fields. The detailed mathematical framework and exact solutions provided herein lay a solid foundation for future research and practical applications. By enhancing the understanding of the fundamental interactions between quantum mechanics and gravity, this study opens new avenues for exploration and contributes to the ongoing efforts to develop a unified theory of physics.

References

[1] Alhaidari, A., & Jellal, A. (n.d.). Dirac and Klein-Gordon Equations in Curved Space.
[2] Collas, P., & Klein, D. (2018). The Dirac equation in general relativity A guide for calculations.
[3] Finster, F., Kamran, N., Smoller, J., & Yau, S. T. (1998). The Dirac equation in the Kerr-Newman metric. *Communications in Mathematical Physics, 203*(4), 681-707.
[4] Cordova, J., Gamba, A., & Passos, E. (2023). Dirac equation in curved spacetime: the role of local Fermi velocity. *The European Physical Journal Plus, 138*(2).
[5] Faba, J., & Sabín, C. (2019). Dirac equation in exotic spacetimes. *Physical Review D*, 99(2), 025008.
[6] Antoine, X., Fillion-Gourdeau, F., Lorin, E., & MacLean, S. (2020). Pseudospectral computational methods for the time-dependent Dirac equation in static curved spaces. Journal of Computational Physics, 411, 109412.
[7] Horwitz, L. P. (2020). Fourier transform, quantum mechanics and quantum field theory on the manifold of general relativity. The European Physical Journal Plus, 135(6),479.
[8] Nyambuya, G. G. (2008). New Curved Spacetime Dirac Equations. Foundations of Physics, 38(7), 665–677.
[9] Pollock, M. D. (2010). On the Dirac equation in curved spacetime. *V.A. Steklov Mathematical Institute, Russian Academy of Sciences*.
[10] Saaty, J. (2001). Differential Geometrical Methods for Deriving Dirac's Equation in Curved Spacetime to Account for the Presence of Matter, arXiv:math-ph/0111034.
[11] Sabín, C. (2017). Mapping curved spacetimes into Dirac spinors. Scientific Reports, 7(1), 40346.

Probing the Strong-field QED Based on Ultra Intense Laser

Zikai Zhou

School of Physics, Jilin University, Changchun, China

zhouzk1120@mails.jlu.edu.cn

Abstract. Contemporarily, with the rapid development of laser techniques, the laser plasma interaction is entering the quantum electrodynamics regime. This study provides a comprehensive analysis of recent advances in the study of strong-field quantum electrodynamics (SF-QED) effects in recent years, especially since 2020. The research details the remarkable developments in laser technology, especially the chirped pulse amplification (CPA) technique, WHICH emphasizes the study of ultrahigh-intensity lasers interacting with electron beams, reaching laser intensities in excess of $10^{23} W/cm^2$. The paper also explores the theoretical framework of QED, focusing on particle generation, spin-related events and their importance in applications such as vacuum birefringence and photon angular momentum effects, and observes key QED effects like nonlinear Compton scattering or multiphoton B-W pair generation. In addition, important QED applications including vacuum birefringence and angular momentum photon effects are discussed. Finally, it is summarized that although the theoretical part of QED is well established and has strong theoretical support, experimental studies are still being followed up. The current strong-field QED experiments have limitations, and looking ahead, with the progress of laser science, future breakthroughs in laser intensity and experimental methods will further promote the development of QED research.

Keywords: Strong filed, ultra intense laser, QED, spin.

1. Introduction

Starting from the first pulsed laser in 1960, laser technology has had a rapid development in the last decades. CPA technology (Chirped pulse amplification) is the key technology driving the development of laser physics. Starting with the inventors Mourou et al. [1], lasers have gradually broken through the initial kilowatt power limitations into today's megawatt, megawatt, and even exawatt power ranges through CPA. Jin is a leader in researching ultra-high lasers. In 2019, the research by applying wavefront correction to a multi-beat-watt (PW) titanium gemstone laser and using an off-axis parabolic mirror for tight focusing, the research team achieved $5.5 \times 10^{22} W/cm^2$ of laser intensity [2]. In 2021, his team used a CoReLS beat-watt (PW) laser with a two-stage adaptive optics system with f/1.1 off-axis parabolic mirrors for wavefront correction and tight focusing to achieve $1.1 \pm 0.2 \times 10^{23} W/cm^2$ peak intensity of the laser by measuring 80 consecutive laser pulses [3].

Ultra-intense laser fields are extremely important in the QED effect. The QED effect may occur in the interaction of an intense laser pulse with a retrograde electron beam [4]. The effect of a normal, low-power laser field on electrons is not strong. In contrast, if the power of the laser field reaches $5 \times 10^{22} W/cm^2$, laser-electron interactions will take place and result in the creation of multiple electron

DOI: 10.54254/2753-8818/52/2024CH0142

pairs from a single beam of electrons. In QED, such a field is considered strong if the electric field E exceeds the Schwinger limit $E \geq E_S = e\lambda_C m_e c^2$ [5]. Such a field has the potential to separate virtual electron-positron pairs and provide energy in excess of the rest mass energy of the electron $m_e c^2$. An electron's scattering process occurs in a strong laser field absorbs multiple photons and emits a single high-energy photon, which is known as nonlinear Compton scattering. An electron moving at relativistic speeds will radiate a high-energy photon, which is the radiative inversion of the electron, and it will greatly affect the electron energy and trajectory.

Typical QED strong-field effects include the creation of pairs of electrons and positrons from the emission of high-energy photons from the electrons or positrons, high-energy photons, i.e., Compton scattering, and the development of a cascade caused by these two processes. In QED, electron-positron pairs are called Polarized Particles, and they exhibit polarization effects in strong fields, i.e., the electrons and positrons are separated by the field force, resulting in a change in the distribution of charge and magnetic moment in the vacuum. This change affects the nature of light propagation in a vacuum. In QED theory when a strong electromagnetic field (e.g., a strong laser field or a strong magnetic field) is applied to the vacuum, virtual particle pairs (electron-positron pairs) are created, and these virtual particle pairs lead to the polarization of the vacuum, which causes the vacuum to exhibit a birefringence effect similar to that of an anisotropic medium. Using flying focused pump lasers, it is expected that this phenomenon will be directly observed in future experiments [6].

This paper is devoted to summarize the recent advancements in the study of the strong-field QED effect in recent years, especially since 2020, to discuss the basic theoretical framework of QED, and then to focus on the particle generation part, spin, with important applications, to study the applications and results of strong-field QED. Attention is also given to future approaches that may have a significant impact on the study of strong-field QED.

2. Description of electrons in a QED-intense laser field

When the electrical field surpasses the threshold known as the Schwinger limit: $E \geq E_s = m_e c^2/e\lambda_c$, it will be considered a strong field. In the focal region of the laser, a QED strong field can be generated. For very strong laser intensities, e.g. $J \approx 10^{22} W/cm^2$ or even higher, a retrograde particle will experience a QED field in a one-dimensional wave. Apart from this, the electric field in the laboratory frame of reference is not considered a strong field in the context of QED [4]. This is because a particle $a(\eta), \eta = \omega t - (k \cdot x)$ traveling retrograde in a one-dimensional wave rises in frequency in the electron's frame of reference, such that the electron's unitless energy E and momentum are correlated with $m_e c^2$ and $m_e c$. k is the projection in the direction of the wave propagation. QED strong field: $E_0 = |dA/d\eta|\omega(E - p_{//})/c$ If the $\chi = E_0/E_s \geq \mathbf{1}$, upon Lorentz transformation, the field surpasses the Schwinger limit. After conversion, the value of χ can be expressed as:

$$\chi = \frac{1.5\lambda_c}{\lambda\left(E - p_{//}\right)\left|\frac{\mathrm{dA}}{\mathrm{d\eta}}\right|} \approx \frac{\left(E - p_{//}\right)}{1.4 \times 10^3}\sqrt{\frac{\mathrm{J}}{10^{23}\mathrm{W/cm^2}}} \tag{1}$$

Here, J is the local instantaneous intensity of the laser. When increasing the laser field intensity to $5 \times 10^{22} W/cm^2 (|a| \approx 110)$, $\chi \approx 90$. When increasing the laser field intensity to $1 \times 10^{23} W/cm^2$, χ will be extremely large and the QED effect is extremely significant.

Nonlinear Compton scattering (NLC) and multiphoton Breit-Wheeler (BW) pair generation are two common phenomena in Quantum Electrodynamics (QED). Among them, the multiphoton Breit-Wheeler (BW) process is an important strong-field quantum electrodynamics (QED) phenomenon describing the generation of electron-positron pairs by the annihilation of high-energy photons in a strong electromagnetic field. The classical Brett-Wheeler process entails the interaction of two high-energy photons, resulting in the production of an electron-positron pair. Conversely, in the multiphoton BW process, a single high-energy photon engages with multiple low-energy laser photons to produce the same result. This nonlinear effect requires extremely high laser field intensity and high-energy photons. Experiment E-144, discussed at the SLAC meeting, an important conference for the study of QED, measured these two strong-field QED processes for the first time. the NLC scattering, i.e., photon emission due to electron-laser interactions, takes the process form $e + n_L\gamma_L \to e' + \gamma$. the multiphoton BW process takes the form $\gamma + n_L\gamma_L \to e^+e^-$ generating electron-positron pairs. In order to exceed the threshold, the participation of $n_L = 5$ laser photons were required in the E-144 experiment, which was confirmed experimentally [7]. The sketch of the processes is shown in Fig. 1.

Figure 1. Nonlinear Compton scattering (emission of photons, left panel) and nonlinear Breit–Wheeler pair generation (right panel) [7].

3. Particle generation processes

In the QED effect, an important route for particle generation is the aforementioned nonlinear Compton scattering (NLC) and multiphoton Brett-Wheeler (BW) pair generation. Theoretically, this two-step process can be achieved with just one step via virtual photon exchange. This process is known as the three-step process for Volkov electrons. Fig. 2 illustrates the Feynman diagram of this trident process. Using the saddle-point method, an analytic result for its generation probability can be obtained:

$$f(a_0) = 4a_0\left\{\left(2 + a_0^2\right)\sinh^{-1}\left(\frac{1}{a_0}\right) - \sqrt{1 + a_0^2}\right\} \tag{2}$$

In experiment E-144, the probability of the experiment is consistent with the theory.

Using the Weizsäcker-Williams approximation, it is possible to estimate the scattering cross section generated by the direct trident pair. Using an extremely simplified form, the nonlinear inverse Compton scattering differential cross section formula can be written as [8]:

$$\frac{d\sigma}{d\Omega} \sim \sum_{n} |Mn|^2\, \delta\left(E_f - E_i + n\hbar\omega\right) \tag{3}$$

Results show that the cross-sectional area generated by direct trident pairs is diminished by a factor of a thousand relative to that of the two-step process (nonlinear Compton scattering followed by BreitWheeler pair generation) [7]. The generation rate R is denoted as:

$$R \sim \mathbf{exp}\left(\frac{f(a_0, b_0)}{\chi}\right) \tag{4}$$

where the function $f(a_0)$ is of the form:

$$f(a0) = -\frac{8}{3}\left(1 - \frac{1}{15a_0^2} + O\left(a_0^{-4}\right)\right) \tag{5}$$

The results of the E-144 experiment show that the observed energy spectrum of the scattered photons agrees with the theoretical predictions, thus validating the mechanism of scattering of energetic electrons under intense laser conditions. In recent years, a concept for probing SFQED by utilizing a highly charged, ultra-relativistic electron beam interacting with a solid conductive target has been proposed, which can attain interaction scenarios similar to those anticipated in beam-beam interactions, utilizing only a single-beam apparatus, in order to achieve a method of investigating SFQED under low-laser-intensity or no-laser-field conditions [9]. Tim Adamo et al. studied the strong-field QED processes of electrons and photons in the context of an ultrahigh-energy accelerated particle beam. The authors consider the electromagnetic field of an ultrahigh-energy particle beam as a superposition of random plane waves, and obtain a wave-function representation of fermions by solving the Drake equation in the background in order to compute the amplitudes of scattering in elastic collisions, nonlinear Compton scattering, and nonlinear BW processes. This is done by expressing the amplitudes of these processes as amplitude averaged over a plane wave background [10]. Seen from left panel of Fig. 3, as the overall charge of the beam increases, it is well represented by the shockwave. Regarding to the right panel of Fig. 3, at constant field strengths, the cross sections of the beam and shockwave vary at high momentum transfer q, examining the beam's microstructure [10].

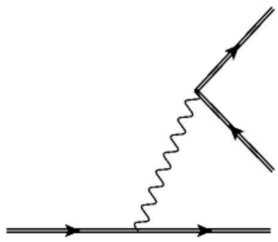


Figure 2. Provides a depiction of the trident mechanism involving Volkov electrons. The phenomena illustrated in Figure 1 can be produced by severing the internal photon line [7].

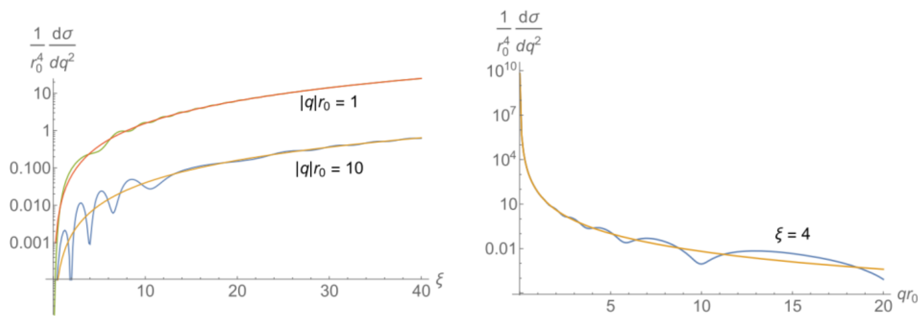


Figure 3. A contrast of the differential cross section (integrated along the z-axis) in the contexts of beam and shockwave [10].

4. Spin-relevant evens

In quantum electrodynamics, spin is of great physical importance. During nonlinear Compton scattering, where an electron absorbs multiple laser photons and then emits a single high-energy photon, the spin state of the electron has a significant effect on the scattering cross section. Previous indicates that in strong-field QED, the radiation probability of an electron in an ultra-intense laser field significantly depends on its spin state [11]. Exact solution in the Volkov electron plane wave background field. In LCFA, it is assumed that the strong field can be considered as a local constant field within the formation length, thus simplifying the calculations. The spin-dependent radiation probability can be expressed as:

$$\frac{dP}{d\tau} = -\alpha b \int dt \left[Ai_1(z) + \frac{gAi'(z)}{z} + \Xi\zeta \frac{sgn(\dot{h})Ai(z)}{\sqrt{z}} \right] \tag{6}$$

where $Ai(z)$ is the Airy function, Ξ is the component of the initial electron polarization vector, ζ is the electron spin direction. In an ultra-intense laser field, the spin state of the electron flips asymmetrically, similar to the Sokolov-Ternov effect in a storage ring. Radiation probabilities for spin-flip and non-flip are given in the literature [11]:

$$\frac{dP_{\mathrm{nonflip}}}{d\tau} = \frac{-\alpha}{2b} \int dt \left\lfloor \begin{array}{c} 2Ai_1(z) + \frac{(g+1)Ai'(z)}{z} + \frac{\zeta\Xi(2t - t^2)}{1-t}\frac{Ai(z)}{\sqrt{z}} \\ + \frac{\Xi^2\kappa t^2}{(1-t)}\left(Ai_1(z) + \frac{Ai'(z)}{z} \right) \end{array} \right\rfloor \tag{7}$$

$$\frac{\mathrm{dP_{flip}}}{\mathrm{d\tau}} = -\alpha/2b \int dt \left[\frac{\mathrm{t}^2}{1-t} \frac{Ai'(z)}{z} + \zeta \Xi \frac{Ai(z)}{\sqrt{z}} - \frac{\Xi^2 \kappa \mathrm{t}^2}{(1-t)} \left(Ai_1(z) + \frac{Ai'(z)}{z} \right) \right] \quad (8)$$

These equations elucidate the alteration in an electron's spin configuration after photon emission. The parameters Ξ and ζ specifically and capture the influence of the initial polarization and spin orientation of the electron. Nonlinear Compton scattering can quickly induce the spin polarizability of the electron. In addition, the polarization generation of gamma photons is closely related to the polarization of laser photons and the spin state of electrons. Regarding Strongly NLC, when an ultra-intense laser pulse collides with an ultra-relativistic electron beam, the polarization of gamma photons is primarily governed by the electron spin configuration. Given that the trajectory deflection angle of the electron in the laser field is significantly greater than the angle of the radiation cone, the formation length is minimal with respect to the inhomogeneous length of the field. In this case, the spin state of the electron has a significant effect on the polarization of the gamma photon. Based on the derivation, the polarizability of the generated gamma photons can be obtained:

$$\boldsymbol{P_\gamma} = \sum_{sf} \frac{(|\boldsymbol{M_{si\to sf}}|^2 - |\boldsymbol{M_{si\to -sf}}|^2)}{(|\boldsymbol{M_{si\to sf}}|^2 + |\boldsymbol{M_{si\to -sf}}|^2)} \quad (9)$$

where $\boldsymbol{P_\gamma}$ is the polarizability of the gamma photon and $\boldsymbol{M_{si\to sf}}$ denotes the matrix element from the initial state spin $\boldsymbol{si}$ to the final state spin $\boldsymbol{sf}$. A comparison is shown in Fig. 4. Probabilistic formulae for the emission of fully polarization-resolved high-energy photons and for the generation of electron-positron pairs are derived under the quasi-classical approximation and the local constant field approximation (LCFA) employing the quantum operator method of Baier and Katkov [13]. Polarization generation of gamma photons is also investigated as affected by the spin state of the electron. The formulas for the polarizability of the generated gamma photons are the same as in Ref. [11] and the results can be verified to be correct.

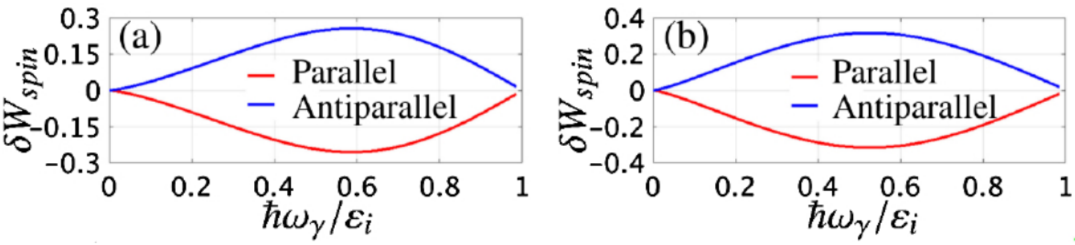


Figure 4. The comparative magnitude of the spin-related component in the probability of radiation for χe = 1 (left) and 0.1 (right). The red/blue curves indicate Si aligned and anti-aligned with SQA, specifically [12].

5. Other proposal applications

5.1. Vacuum Birefringence in QEDs

The vacuum is not empty in the framework of quantum electrodynamics, but is filled with the rise and fall of virtual particles. According to quantum field theory, Under the influence of an applied electromagnetic field, virtual e^- e^+ pairs can cause polarization of the vacuum. leading to a change regarding the transmission characteristics of EM waves in vacuum: the propagation of a photon in the vacuum appears similar to the propagation of light in a birefringent medium, called vacuum birefringence. The vacuum birefringence effect was first proposed by Heisenberg and Euler in 1936, who used a one-loop effective Lagrangian quantity to describe the electrodynamic corrections under strong fields. This correction considers the polarization effect of virtual e^-e^+ pairs within the vacuum, which results in vacuum exhibiting nonlinear optical properties [14]. The vacuum birefringence effect in QED has not yet been verified directly by experiments, but numerous experiments such as the PVLAS experiment [15] and others are now available, which provide valuable experience for the development of direct observations and theoretical studies of vacuum birefringence.

5.2. Angular momentum photon effect in QEDs

The angular momentum photon effect is mainly concerned with the fact that by interacting with a plasma using a circularly polarized LG laser, the spin angular momentum (SAM) and orbital angular momentum (OAM) driving the laser can be transferred to electrons, which in turn are transferred to high-energy gamma photons by quantum radiation. This interaction can be illustrated using a 3D quantum electrodynamic particle-in-cell (PIC) simulation that depicts the topological charge, chirality, and CEP of the photon. Moreover, the orbital angular momentum (OAM) of the photons is transferable to the electrons via the interaction, thus affecting the trajectory and radiation properties of the electrons. The forces on the electrons in this laser field have a helical structure, causing the electrons to move along the helical path. During the process, the radiated photons also inherit part of the angular momentum, leading to the spiral characteristics of the radiated photons. This radiation pattern is different from the conventional linear radiation pattern. In a high-intensity laser field, the electrons are able to obtain higher energies, resulting in the production of higher-energy gamma photons. The helical structure of the OAM photons enhances the generation efficiency of such high-energy photons [16].

6. Limitations and prospects

Currently, the research on strong-field QED is developing rapidly, and the theoretical part of QED has been developed in a very powerful way. The description of QED is very precise and is supported by a strong theory, which allows simulations to be performed to get the expected correct results. A variety of effects have been predicted, and current experiments are gradually confirming them. However, the experimental part of the study is not perfect, and with respect to the laser part, current lasers have recently reached the order of $10^{23} W/cm^2$, but to detect the more pronounced and relevant cascade processes, and other effects derived from QED, lasers of the order of $10^{24} W/cm^2$ and even higher may be needed. of order of magnitude. In addition, there are many QED effects that have not been directly observed experimentally, such as vacuum birefringence with more intense cascade processes. However, the problem is now being gradually solved: laser technology is progressing and will reach the order of $10^{24} W/cm^2$ soon, and, low laser fields versus no laser field and only one ultra-relativistic electron beam are now being gradually explored to study SFQEDs. It has been mentioned in Sec, 3. This research will lead to a breakthrough in the future development of QEDs.

7. Conclusion

To sum up, a more comprehensive discussion of strong-field QED based on ultra-intense lasers is presented in this study. With a laser of sufficient power, the field exceeds the Schwinger limit and the particle will exhibit QED effects. The laser interacts with matter, generating positive and negative electron pairs and gamma photons with a series of cascade effects, using the QED parameter χ to measure the strong-field QED effect and a discussion of its basic definition. The role of spin in QED is investigated, with implications for scattering cross sections and models, and applications such as vacuum birefringence and the angular momentum photon effect, which are important in QED, are discussed. the limitations of current strong-field QED are summarized, and more methods for implementing QED are discussed. this paper summarizes QED research since the 20th century, and highlights new developments since 2020, in anticipation of a more robust future for QED. The paper summarizes the QED research since the 20th century and focuses on the new developments since 2020 in the hope that more excellent methods will emerge in the future.

References

[1] Strickland D and Mourou G 1985 Compression of amplified chirped optical pulses Optics Communications vol 56(3) pp 219-221

[2] Jin W Y, Cheonha J J, Shin S, et al. 2019 Achieving the laser intensity of 5.5×1022 W/cm2 with a wavefront-corrected multi-PW laser. Opt Express vol 27 pp 20412-20420.

[3] Jin W Y, Yeong G K, Li W C, et al. 2021 Realization of laser intensity over 1023 W/cm2 Optica vol 8 pp 630-635.

[4] Sokolov I V, Naumova N M, Nees J A and Mourou G A 2010 Pair Creation in QED-Strong Pulsed Laser Fields Interacting with Electron Beams Physical Review Letters vol 105(19) p 195005.

[5] Schwinger J 1951 On gauge invariance and vacuum polarization Physical Review vol 82(5) p 664.

[6] Jin B and Shen B 2023 Enhancement of vacuum birefringence with pump laser of flying focus Physical Review A vol 107(6) p 062213

[7] Altarelli M, Assmann R, Burkart F, Heinemann B, Heinzl T, Koffas T and Wing M 2019 Summary of strong-field QED Workshop arxiv preprint arxiv:190500059.

[8] Ehlotzky F, Krajewska K and Kamiński J Z 2009 Fundamental processes of quantum electrodynamics in laser fields of relativistic power Reports on Progress in Physics vol 72(4) p 046401

[9] Matheron A, San Miguel Claveria P, Ariniello R, Ekerfelt H, Fiuza F, Gessner S and Corde S 2023 Probing strong-field QED in beam-plasma collisions Communications Physics vol 6(1) p 141.

[10] Adamo T, Ilderton A and MacLeod A J 2021 Particle-beam scattering from strong-field QED Physical Review D vol 104(11) p 116013

[11] Sun T, Zhao Q, Xue K, Lu Z W, Ji L L, Wan F and Li J X 2022 Production of polarized particle beams via ultraintense laser pulses Reviews of Modern Plasma Physics vol 6(1) p 38.

[12] Li Y F, Shaisultanov R, Hatsagortsyan K Z, Wan F, Keitel C H and Li J X 2019 Ultrarelativistic electron-beam polarization in single-shot interaction with an ultraintense laser pulse Phys Rev Lett vol 122(15) p 154801.

[13] Chen Y Y, Hatsagortsyan K Z, Keitel C H and Shaisultanov R 2022 Electron spin-and photon polarization-resolved probabilities of strong-field QED processes Physical Review D vol 105(11) p 116013.

[14] Dinu V, Heinzl T, Ilderton A, Marklund M and Torgrimsson G 2014 Vacuum refractive indices and helicity flip in strong-field QED Physical Review D vol 89(12) p 125003.

[15] Bragin S, Meuren S, Keitel C H and Di Piazza A 2017 High-energy vacuum birefringence and dichroism in an ultrastrong laser field Physical Review Letters vol 119(25) p 250403.

[16] Liu C, Shen B, Zhang X, Shi Y, Ji L, Wang W and Xu Z 2016 Generation of gamma-ray beam with orbital angular momentum in the QED regime Physics of Plasmas vol 23(9).

Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations
DOI: 10.54254/2753-8818/52/2024CH0133

Research on Art Classification Based on the Improved Resnet50 Network

Yingkai Liang[1], Hangxi Ren[2,5], Ziwei Wan[3], Rui Zhou[4]

[1]Helen Way Klingler College of Arts and Sciences, Marquette University, Milwaukee,USA
[2]Software College, Xinjiang University, Urumqi, China
[3]College of Information Engineering, Sichuan Agricultural University, Sichuan, China
[4]College of Computer Science, Nanjing University of Post and Telecommunication, Nanjing, China

[5]20222501439@stu.xju.edu.cn

Abstract. In the field of computer vision, deep learning technology has made breakthroughs in image classification. ResNet50, an efficient deep residual network architecture, has demonstrated excellent performance in many image classification tasks. However, In the task of classifying artworks, ResNet50, still shows some limitations. It has difficulty identifying different artistic styles and is lacking in capturing the delicate details of the artwork. This article proposes an improved resnet50 model. By introducing SE blocks and extended convolution parameters, this article have made a deep revamp to ResNet50 to enhance its performance in art classification tasks. SE blocks increase the sensitivity of the network to differences in art styles by dynamically adjusting the dependencies between channels. Concurrently, the enhanced convolutional operation expands the model's sensory scope, so that the network can capture more delicate artistic details. The experimental results show that the improved model has achieved significant performance improvement on wikiart artwork classification datasets. In upcoming research, the paper plan to enhance art classification by broadening the dataset to encompass diverse art movements and forms, and by optimizing convolutional neural networks for large-scale data. This aims to improve the model's generalization and applicability across various visual arts, ensuring practical effectiveness and reliability.

Keywords: Computer vision, Image classification, Resnet 50, Residual network.

1. Introduction

Artwork classification is an important application field in computer vision [1], and automatic classification of artworks helps with their identification, management, and protection [2]. Traditional image classification methods often perform poorly due to the complexity and diversity of images when processing art images [3].

In traditional image classification methods, researchers usually need to manually extract image features and formulate classification rules. This method often appears to be unsatisfactory when processing large -scale and diverse image data, not only low efficiency, but also difficult to guarantee accuracy [4]. However, with the rapid development of deep learning technology, Particularly the

DOI: 10.54254/2753-8818/52/2024CH0133

efficient use of convolutional neural networks (CNNs) in the realm of image classification, these problems have been effectively resolved [5]. CNN has greatly improved the accuracy and efficiency of classification through the hierarchical structure characteristics of automatic learning images [6].

However, art images have unique characteristics, such as complex texture, diverse styles and delicate details. These characteristics make art classification tasks more challenging. Traditional ResNet50 may not be able to fully capture abstract information such as details and styles in art images when processing these features. Therefore, the goal of this experiment is to use the deep learning framework of PyTorch in the field of computer vision to complete the image classification tasks in specific scenarios by deep learning neural network ResNet, focusing on improving and optimizing some modules of the Resnet network [7]. To adapt to specific application scenarios to achieve higher classification accuracy and better generalization capabilities [8].

This experiment was developed based on the PyTorch library of Python. Pytorch is a widely used opensource machine learning library. Especially in the field of deep learning, it is favored by researchers and developers for its dynamic calculation diagram and ease of use. Pytorch provides rich APIs and tools, making the construction and training deep learning models intuitive and efficient.

During this experiment, the structure of the resnet network was first analyzed in depth to identify key modules that may affect classification performance. Then, try different improvement strategies, such as adjusting the depth and width of the network, introducing attention mechanisms, optimizing the initialization of the network, etc., in order to improve the expression ability and generalization of the model. In addition, the number of overpairies in data pre -processing, enhancement strategies, and training in the training process has been carefully adjusted and optimized.

2. Related works

Image classification is an important research direction in the field of computer vision, which aims to classify images into predefined categories. With the development of deep learning technology, the image classification has made significant progress [9].

Before the popularization of deep learning, image classification mainly depends on handmade characteristics extraction and traditional machine learning algorithms. Popular feature detection approaches encompass the Scale-Invariant Feature Transform(SIFT) method, Histogram of Oriented Gradients, Local Binary Patterns(LBP) and so on. These methods are represented by the local characteristics or global characteristics of the image, and then the image is represented as a feature vector, following that, algorithms such as Vector Support Machines (SVM), Nearest Neighbor (NN), and Stochastic Forests are applied for classification.

However, most of these traditional methods have these limitations: handmade feature extraction of the complexity and diversity of images cannot be completely captured; these methods have lower efficiency when dealing with large -scale data sets; these methods have noise and noise and noise in the image in the image. Deformation is more sensitive and has poor robustness [10].

In recent years, the emergence of deep learning methods, especially the convolutional neural network (CNNS), has greatly promoted the progress of image classification [11]. Deep learning methods can automatically learn the characteristics from the data, avoiding the limitations of handmade feature extraction. Typical deep learning models include Lenet, Alexnet, VGG, Googlenet and Resnet [12].

The RESNET is proposed by the Microsoft Research Institute [13] that the residual connection technique is employed to address issues of vanishing and exploding gradients within deep neural network architectures during the training process, so that the network can be deeper and easy to train. Resnet has achieved leading performance in multiple image classification benchmark tests.

In the field of art classification, researchers have proposed a variety of methods and technologies. Early studies primarily employed a method that relies on manual feature extraction and traditional machine learning approaches, including Support Vector Machines (SVM) and K-Nearest Neighbors (KNN), along with others. Although these methods have achieved certain effects in certain tasks, their classification performance is often not ideal due to the limitations of manual characteristics.

DOI: 10.54254/2753-8818/52/2024CH0133

With the development of deep learning technology, models based on convolutional neural networks have gradually become the mainstream method of art classification. Classic network structures such as VGG, Inception, and ResNet are widely used in various image classification tasks. Especially for resnet, by introducing residual connections, the disappearance of gradient disappearance in deep networks enables the network to extract image features deeper [14]. In the art classification task, many researchers also adopted Resnet as the basic model and improved on this basis.

However, directly applying standard ResNet50 models still have some problems in art classification tasks. For example, the model is easy to ignore the details of the details when processing artistic images with complex texture and diverse style. In addition, the ResNet50 framework underutilized the hierarchical feature nuances present in the image. Therefore, how to improve on the basin of ResNet50 to better adapt to artistic classification tasks and become a problem worth studying.

3. Research methods

Dilated Convolution, this study broadens the perspective by incorporating the parameter count or loss resolution within the convolutional core, thereby augmenting the model's feature extraction capabilities. The realization of expansion convolution is as shown in formulas (1):

$$Oi\text{dilated conv} = siOi\text{in} + 2pi - di \times (ki - 1) + 1 \tag{1}$$

Oidilated conv:Output size of the dilated convolution. si: Size of the input feature map. Oi: Size of the output feature map. pi: Padding size of the input feature map. di: Dilation rate. Ki: Dimension of the convolutional filter.

The convolution workflow and related parameters are shown in Figure 1.

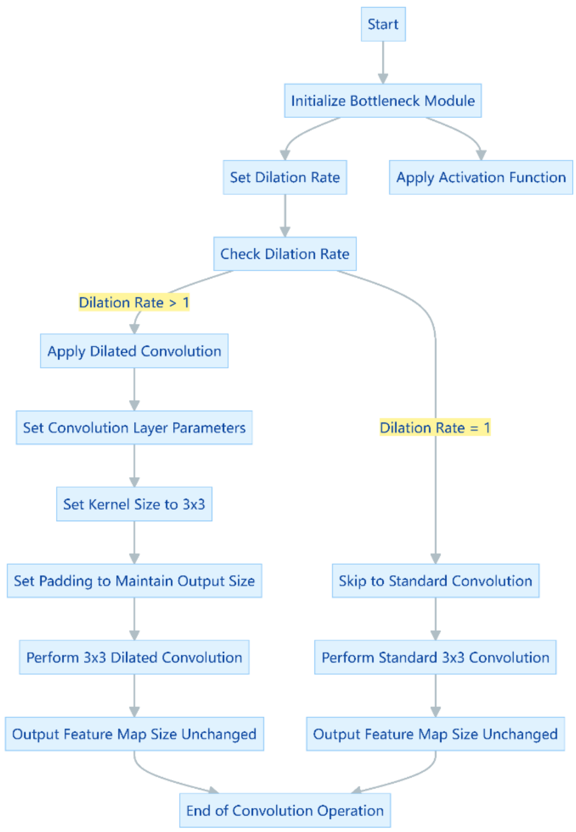


Figure 1. Convolution workflow and related parameters

Add Squeeze-And-EXCity Block. SE blocks adjust the weight of the characteristic channel through adaptive land, to ensure our models prioritize relevant features while discarding irrelevant ones, thereby enhancing the model's discriminative power exponentionally, while maintaining lower calculation costs and good versatility.

$$uc = \mathrm{Fscale}(\mathrm{Fexcite}(\mathrm{Freduce}(X)), X) \quad (2)$$

X**:** Enter the feature diagram to contain multiple channels.

Freduce: Squeeze operation is usually achieved by global average pooling, which is used to compress the activation value of each channel into a single value to capture the activation distribution of the channel.

Fexcite: Excite operation is usually achieved through a full connection layer (or "squeeze" layer) to learn the dependence between channels and generate a weight for each channel.

Fscale: SCALE operation will multiply the weight of the operation generated by the input characteristics, so as to obtain the output feature diagram.

The SE block introduces an attention mechanism to the network in this way, so that the network can pay more attention to the more important feature channel for the current task, thereby improving the performance of the network.

Se block is shown in Figure 2.

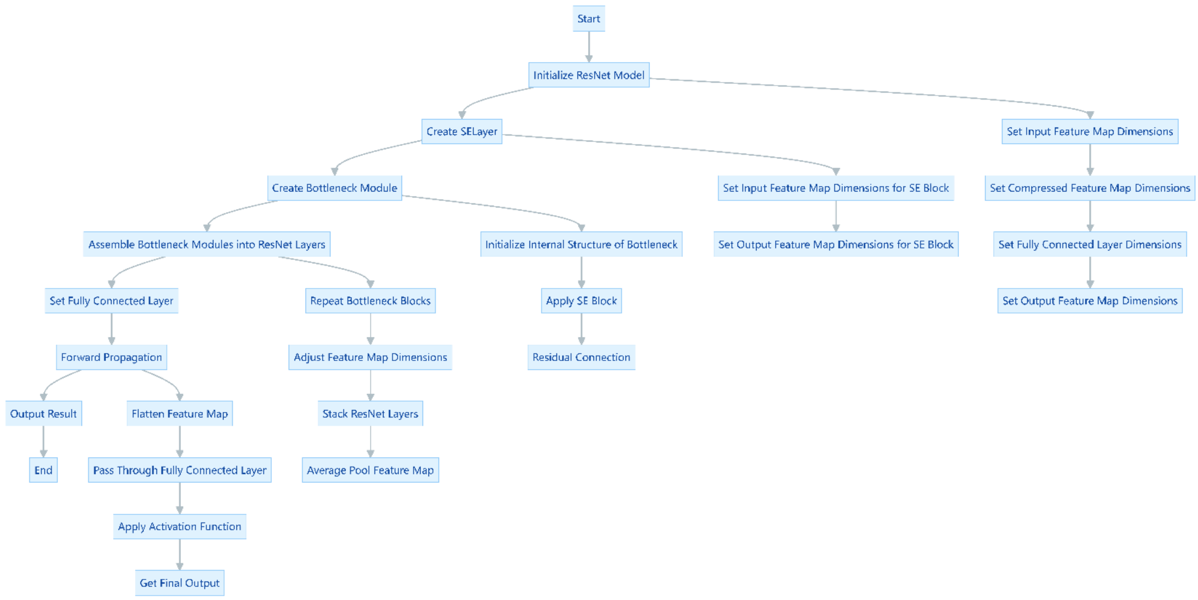


Figure 2. SE block

Data enhancement technology: The pre-training weight of the official website was used. The pre-training weight of the official website can not only significantly improve the ability to extract feature extraction, but also accelerate the training process, improve the accuracy and robustness of the model, and reduce overfitting risks and calculation costs.

4. Experimental results

4.1. data set

In order to verify the effectiveness of the model, and to improve the data quality, improve the processing efficiency and simplify the model training, the paper preprocessed the wikiart dataset and discarded some files that could not be decoded, so as to achieve the above goal.

Wikiart dataset shows the rich diversity and historical depth of art with its large scale of more than 81000 art images. Spanning a wide range of periods from the Renaissance to modern art, it records the evolution of artistic styles, covering 27 different artistic styles, such as abstract expressionism, Baroque, cubism and Impressionism. Each style is a unique chapter in the history of art. These works are from

DOI: 10.54254/2753-8818/52/2024CH0133

the hands of hundreds of artists, who belong to different cultural and historical backgrounds, and show a variety of artistic perspectives and innovative techniques.

The data set of more than 60 art genres has further enriched its content, including traditional portraits, landscape paintings and modern street art, which not only reflects the diversity of artistic expression, but also provides an opportunity to study the interaction and integration between different genres. The universality and application potential of Wikiart data set lies in its extensive art coverage. It is not only limited to the study of art history, but also provides training data for tasks such as image recognition, style classification, and art creation in the field of machine learning, becoming a bridge connecting art and technology. This interdisciplinary resource is a valuable asset for educators, students, artists, researchers and technology developers.

4.2. Evaluation index

4.2.1. Precision. Precision is the most intuitive evaluation index, which measures the proportion of the number of samples correctly predicted by the model to the total number of samples. The accuracy rate is very effective for the situation of uniform data distribution, but when the number of samples in different categories varies greatly, the accuracy rate may be misleading, because it will not distinguish whether the model performs well in a few categories.

4.2.2. Accuracy. The accuracy metric represents the ratio of the true positive instances among all instances predicted as positive. Precision is particularly crucial in scenarios where minimizing false positives (erroneously predicting the impending negative class as positive) is of paramount importance.

4.2.3. Recall rate. The recall metric quantifies the ratio of true positive instances that are accurately identified as such by the model. The recall is especially critical in scenarios where the minimization of false negatives (that is, the incorrect classification of positive instances as negative) is a primary concern.

4.2.4. F1-score. F1-score is the harmonic average of accuracy rate and recall rate. It tries to find a balance between accuracy rate and recall rate. When the accuracy rate and recall rate are both high, F1 score will also be high, which means that the model performs well in both aspects.

4.3. Experimental results and analysis

The experimental results show that the accuracy and other indicators of the improved resnet50 model are improved compared with other models. After a hundred rounds of training, the test indicators obtained are as follows:

Table 1. performance of the improved algorithm, the original algorithm and other algorithms on different datasets

Model name used	Training Accuracy	Validation Accuracy	Precision	Recall	F1 Score
Improved resnet-50	0.9998	0.6455	0.6303	0.6226	0.6206
Original resnet-50	0.9774	0.6071	0.5959	0.5828	0.5840
vgg-16	0.774	0.5365	0.5297	0.5163	0.5368
ConvNext	0.954	0.6032	0.5986	0.5878	0.5923

Table 1 shows the test performance of different neural networks after training. It can be seen that compared with the original resnet-50, vgg-16 and CNN networks, the performance of the improved resnet-50 has been improved to a certain extent, especially compared with the original resnet-50, the verification accuracy of the improved network has been improved by about 4%. In addition, compared with other neural networks, the improved resnet-50 has a training accuracy of 0.9998, which means that the model can be nearly completely correct on the training data set. The very high accuracy shows that

DOI: 10.54254/2753-8818/52/2024CH0133

the model performs well on the training set, which further shows that the model shows strong pertinence in predicting the data of known categories.

5. Conclusion

This paper proposes an improved resnet50 model to improve its performance in art classification tasks. Through the introduction of multi-level feature fusion mechanism, attention mechanism and a variety of data enhancement technologies, the improved model has shown significant performance improvement on multiple public art classification datasets. Specifically, the extended convolution mechanism enhances the feature capture ability of the model; Adding Se blocks enhances the attention of the model to important features and further improves the accuracy of classification; Data enhancement technology increases the diversity of training samples and improves the generalization ability and robustness of the model.

The experimental results show that the improved ResNet50 model outperforms the original ResNet50 model and other existing models in terms of classification accuracy, precision, recall, and F1 score on the WikiArt, Rijksmuseum, and Art500K datasets, validating the effectiveness of the improved method.

In addition, this study also has certain practical application value. The improved model can be applied to fields such as museums, art galleries, and art trading platforms to assist in automatic classification and management of artworks, improving work efficiency. At the same time, this model can also provide technical support for art appraisal and protection, promoting the digital protection and inheritance of cultural heritage.

The experimental results show that the ResNet model built using the PyTorch framework has achieved satisfactory training results in both local environments and cloud platforms. The model has demonstrated high accuracy in image classification tasks in specific scenarios, verifying the effectiveness of the improved strategy. Meanwhile, the model also demonstrated good generalization ability when faced with new data, providing a solid foundation for practical applications.

In summary, this experiment successfully improved the performance of image classification tasks through the application of deep learning and PyTorch framework, providing new ideas and methods for research and application in the field of computer vision. In the future, it may be possible to continue exploring deeper network structures and more advanced optimization strategies in order to achieve greater breakthroughs in the field of image classification.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Surapaneni, S., Syed, S., & Lee, L. Y. (2020). Exploring themes and bias in art using machine learning image analysis. In *2020 Systems and Information Engineering Design Symposium (SIEDS)* 1-6.

[2] Tan, W. R., Chan, C. S., Aguirre, H., & Tanaka, K. (2019). Improved ArtGAN for Conditional Synthesis of Natural Image and Artwork. IEEE Transactions on Image Processing, 28(1), 394-409.

[3] Qian, X. C., & Mi, S. (2019). Color style feature extraction of thangka based on deep neural network. Computer Science and Application, 9, 2129.

[4] Cetinic, E., Lipic, T., & Grgic, S. (2018). Fine-tuning convolutional neural networks for fine art classification. *Expert Systems with Applications*, *114*, 107-118.

[5] Koonce, B., & Koonce, B. (2021). ResNet 50. Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization, 63-72.

[6] Zhao, Z. C., Luo, Z., Wang, P. Y., & Li, J. (2019). A review of image classification algorithms based on deep residual networks. *Computer System Applications*, *29*(1), 14-21.

[7] Sandoval, C., Pirogova, E., & Lech, M. (2019). Two-stage deep learning approach to the classification of fine-art paintings. *IEEE Access*, *7*, 41770-41781.

DOI: 10.54254/2753-8818/52/2024CH0133

[8] Li, Q., Ji, Y. B., Guo, C. W., & Liu, K. X. (2022). Pattern recognition and classification of women's shirts based on deep learning. *Advanced Textile Technology*, *30*(4).

[9] Zhu, J. (2022). SECure: SE block-based classification of buildings post hurricane. In International Conference on Cloud Computing, Performance Computing, and Deep Learning (CCPCDL 2022) 12287, 284-288.

[10] Guo, Y. X., Yang, W., Liu, Q., & Wang, Y. (2020). A Review of Residual Network Research. *Application Research of Computers/Jisuanji Yingyong Yanjiu*, *37*(5).

[11] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) 770-778.

[12] Yang, H. J., Wang, E., Sui, Y. X., Yan, F., & Zhou, Y. (2022). Simplified residual structure and fast deep residual network. *Journal of Jilin University (Engineering Edition)*, *52*(6), 1413-1421.

[13] Cheng, D. Q., Wang, Y. C., Kou, B. Q., Fu, X. Z., Chen, L. L., & Zhao, K. (2021). Mine image classification based on improved deep residual network. *Application Research of Computers/ Jisuanji Yingyong Yanjiu*, *38*(5).

[14] Li, X., Su, J. N., Zhang, Z. P., & Zhou B. Y. (2024). Fine recognition of product style based on deep learning. *Computer integrated manufacturing system*, *30*(3), 1011.

Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations
DOI: 10.54254/2753-8818/52/2024CH0125

Factors influencing housing prices: A case study of Shanghai

Xinning Fang

Liaoning Experimental Middle School, Shenyang, Liaoning, 110841, China

XinningFang@outlook.com

Abstract. Housing is a major consumer good for residents, and housing prices affect people's quality of life and happiness. This paper analyzes the factors influencing housing prices based on the housing price dataset for Shanghai in 2023-2024. Comparison is made via multiple linear regression model to discover the difference between with and without considering interaction terms, assessing the importance and accuracy of the two models' results, and optimizing the model. According to the analysis, living area, decoration type, and building usage are the factors most positively correlated with housing prices. The number of bathrooms, living rooms, and bedrooms also shows a significant positive correlation with housing prices. Additionally, building type and the construction age of the house are negatively correlated with housing prices, and this correlation is very significant. Finally, the housing area weakens the positive impact of the number of bathrooms and bedrooms on housing prices.

Keywords: Housing prices, multiple linear regression, influence factors, Shanghai.

1. Introduction

Housing prices are crucial to people's life and have always been an important indicator and public concern. Since 2004, housing prices nationwide have been continuously rising: housing price in China experienced unprecedented growth from 2002 to 2010, by nearly 1.5 times and seems to be stabilizing recently [1, 2]. The daily consumption of residents in both China's urban and rural regions were affected by the high private housing prices in the real estate sector. However, the public is not particularly aware of the influencing factors [3]. This paper analyzes the various potential factors that may affect the housing prices in Shanghai, in attempt to help residents with tools to assess and predict the trend of housing prices in Shanghai, and possibly other cities in China.

The real estate sector is a sophisticated regime: a myriad of factors may contribute to the mobile housing prices. Housing price assessment and prediction is also a subject of research academia. Lü claims that area of the houses is significant for analyzing the development of the Chinese real estate market [4]. Yan et al. also used housing price dataset in Beijing, combining the hedonic models, the lasso regression, and the random forests in their analysis, and optimizing their results [5]. They studied such variables as the number of living rooms, bedrooms, and bathrooms on housing prices. The combination of multiple tools increased the accuracy and reliability of their analysis. As for second-hand houses, Zhen et al. suggested that the availability of elevators is an important factor for prices, as shown in their study of Chengdu [6]. Multiple linear regression, decision trees, and extreme gradient boosting models are used to fit the price prediction curve of influencing factors. The XGBoost algorithm stands out, when compared with other models, as the most accurate, widely applicable, and reliable in data prediction, while reducing overfitting. However, it might slightly lack precision when analyzing results. Domestic scholars believed that the structure and the age of the buildings are among the influencing factors for housing prices [6, 7]. It is claimed by Zheng that the renovation of the houses is a crucial factor leading to differences of their prices [8]. Multiple linear regression and semi-log models are employed in the study, and ordinary least squares (OLS) regression is used to estimate parameters and verify the model's accuracy [9]. However, the paper used a small data set of 219 samples. While studying housing mortgage prices, Fan et al. discovered that property rights significantly affect housing prices [10]. Wang et al. Focuses their study on external factors, using a global regression model to analyze the significant spatial distribution pattern of housing prices, taking their adjacency to subway stations into consideration [11].

In this research, nine variables are assessed for their impact on housing prices and model to study the correlation between these factors and housing prices, including housing area, building type, decoration type, building orientation, housing purpose, building type, number of bathrooms, number of living rooms, and number of bedrooms, using multiple linear regression model, to discover the influencing factor for housing prices in Shanghai.

2. Methodology

2.1. Data source

The dataset used in this study is collected from various websites showing Shanghai housing prices data and spans from 2023 to 2024. The dataset contains 395,001 entries, and 48,234 of these were selected as the samples for this study.

2.2. Variable selection

The original dataset is huge and contains many variables with missing values, such as construction time, building type, and building structure. These variables were removed in this study. Ultimately, random sampling was used to obtain the data. The data includes nine variables (housing area, building type, decoration type, building orientation, housing purpose, number of bathrooms, number of living rooms, and number of bedrooms) and one dependent variable (the housing price). The dataset used is described in Table 1:

Table 1. Variables description.

Variables	Symbols	Meaning
price	Y	House price
sqft	x_1	House area
Building type	x_2	Assigned value of 1 for panel buildings, 0 otherwise
Decoration type	x_3	Assigned value of 1 for luxury decoration, 0 otherwise
Building orientation	x_4	Assigned value of 1 for south-facing, 0 otherwise
Building usage	x_5	Assigned value of 1 for residential use, 0 otherwise
Housing life span	x_6	Housing life span
Restrooms	x_7	Number of restrooms
Living rooms	x_8	Number of living rooms
bedrooms	x_9	Number of bedrooms

2.3. Method introduction

The multiple linear regression model is employed to compare results with and without considering interaction terms, aiming to compare the importance and accuracy of the two models' results. Ultimately, the optimization of the models is conducted.

The multiple linear regression model falls into the category of linear regression models. The multiple explanatory variables used in the analysis enables it to illustrate the linear relationship between the

DOI: 10.54254/2753-8818/52/2024CH0125

explanatory variables and the dependent variable [12]. The model estimates multiple parameters through Ordinary Least Squares (OLS), thereby minimizing the sum of squared residuals between the dependent and independent variables.

3. Results and discussion

3.1. Correlation analysis

The analysis in this paper shows that there are many factors influencing housing prices. As shown in the Table 2:

Table 2. Correlation Analysis between Dependent and Independent Variables

	lnsqrt	buildtype	decoration	buildingor	buildingus	housinglife
lnprice	1					
lnsqrt	0.102***	1				
buildingtype	0.023***	-0.008*	1			
decoration	0.179***	0.173***	0.00400	1		
buildingor	-0.010**	0.092***	0.281***	-0.010**	1	
buildingus	0.283***	0.078***	0.229***	0.032***	0.200***	1
housinglife	0.178***	-0.535***	0.046***	-0.158***	-0.051***	0.136***
restrooms	0.202***	0.648***	0.027***	0.143***	0.018***	-0.020***
livingrooms	0.114***	0.720***	0.046***	0.198***	0.090***	0.103***
bedrooms	0.070***	0.810***	0.048***	0.130***	0.076***	0.049***

The data reveals that living area, decoration type, and building usage are the factors most positively correlated with housing prices. Therefore, it seems that people are particularly interested in the living area, decoration type, and building usage of a house. The number of bathrooms, living rooms, and bedrooms also shows a significant positive correlation with housing prices. Additionally, building type and the age of the house are negatively correlated with housing prices, and this correlation is very significant. In summary, the factors influencing housing prices are comprehensive. Nowadays, people desire a perfect house from various perspectives. After analyzing the Pearson correlation matrix of the factors, a multiple regression analysis is conducted.

3.2. Linear model results

The results of the linear regression are shown in Table 3. The regression coefficients of the multiple linear regression equation model indicate that the area of the house negatively affects its price, while the building type positively affects the price. Houses that are in slab-type apartment buildings are more expensive than those that are not. The decoration type also has a positive impact on housing prices, with luxury-decorated houses being more expensive than non-luxury-decorated ones. The orientation of the house has a significant positive impact on the price as well, with south-facing houses being more expensive than others. The regression coefficient for housing usage is the largest, indicating that residential-use houses are more expensive than non-residential-use houses. The construction age of the house negatively affects the price, but the impact is relatively small. The number of bathrooms, living rooms, and bedrooms positively influences housing prices.

DOI: 10.54254/2753-8818/52/2024CH0125

Table 3. Regression Coefficients Table

	FE	RE
	lnprice	lnprice
lnsqrt	-0.060***	-0.029***
	(-9.58)	(-4.72)
buildingtype	0.063***	0.046***
	(12.93)	(9.63)
decorationtype	0.037***	0.042***
	(23.31)	(25.96)
buildingorientation	0.026***	0.018***
	(7.11)	(4.92)
buildingusage	1.128	1.030
	(120.23)	(115.42)
housinglifespan	-0.001***	0.002***
	(-2.63)	(8.22)
restrooms	0.042***	0.057***
	(15.31)	(21.01)
livingrooms	0.032***	0.034***
	(15.44)	(16.15)
bedrooms	0.015***	0.004*
	(6.75)	(1.81)
cons	9.761	9.732
	(398.99)	(393.99)

3.3. Linear regression with interaction terms

Interactions between some independent variables might also affect housing prices; these interaction terms are known as interaction effects. In fact, the number of bedrooms and bathrooms may be somewhat related to the size of the house, meaning the effect of the number of bedrooms and bathrooms on the house price depends on the housing area. Interaction terms between the number of bedrooms and housing area and between the number of bathrooms and housing area are added to the regression equation. The regression results are shown in Tables 4 and 5.

The interaction term between the number of bedrooms and the housing area (x1x5) is significantly negative, while the coefficient for the number of bedrooms is significantly positive. This indicates that the housing area significantly weakens the positive effect of the number of bedrooms on housing prices. When the housing area is larger, the impact of the number of bedrooms on the price may become secondary, as shown in Table 4.

The interaction term between the number of living rooms and the housing area (x1x8) is significantly positive, while the coefficient for the number of living rooms is significantly positive. This suggests that the housing area significantly weakens the positive effect of the number of living rooms on housing prices. When the housing area is larger, the impact of the number of living rooms on the price diminishes. These results are shown in Table 5.

Table 4. Regression Results for Interaction Term Between Number of Bedrooms and Housing Area

	FE	RE
	lnprice	lnprice
lnsqrt	-0.009	0.010
	(-1.42)	(1.47)

DOI: 10.54254/2753-8818/52/2024CH0125

Table 4. (continued).

x1x9	-0.001***	-0.000***
	(-21.23)	(-15.99)
buildingtype	0.067***	0.049***
	(13.67)	(10.27)
decorationtype	0.037***	0.042***
	(23.19)	(25.86)
buildingorientation	0.024***	0.017***
	(6.53)	(4.51)
buildingusage	1.087	0.999
	(114.10)	(109.60)
housinglifespan	-0.001**	0.002***
	(-2.14)	(8.65)
restrooms	0.069***	0.078***
	(23.13)	(25.96)
livingrooms	0.077***	0.067***
	(26.04)	(22.76)
bedrooms	0.017***	0.005**
	(7.92)	(2.49)
cons	9.555	9.574
	(364.86)	(360.83)

Table 5. Regression Results for Interaction Term Between Number of Living Rooms and Housing

	FE	RE
	lnprice	lnprice
lnsqrt	-0.020***	0.003
	(-3.19)	(0.47)
x1x8	-0.000***	-0.000***
	(-23.06)	(-18.05)
buildingtype	0.065***	0.048***
	(13.43)	(10.12)
decorationtype	0.037***	0.042***
	(23.16)	(25.84)
buildingorientation	0.025***	0.018***
	(6.96)	(4.82)
buildingusage	1.092	1.000
	(115.53)	(110.55)
housinglifespan	-0.001**	0.002***
	(-2.54)	(8.27)
restrooms	0.075***	0.084***
	(24.58)	(27.16)
livingrooms	0.031***	0.033***
	(15.06)	(15.76)
bedrooms	0.046***	0.027***
	(17.97)	(10.80)

DOI: 10.54254/2753-8818/52/2024CH0125

Table 5. (continued).

cons	9.588	9.594
	(377.06)	(372.33)

4. Conclusion

Ownership of an apartment or house in a major city is a lifelong pursuit for many Chinese urban residents. The soaring housing prices in megacities like Shanghai and Beijing make it difficult for many to afford. This paper examines over 48,000 data sets of Shanghai's real estate transactions from 2023-2024 to understand the factors driving these rising house prices.

This paper uses a multiple linear regression model to compare results with and without considering interaction terms. 9 variables were selected for the analysis. Results reveals shows that Living area, decoration type, and building usage are the factors most positively correlated with housing prices. Therefore, it seems that people are particularly interested in the living area, decoration type, and building usage of a house. The number of bathrooms, living rooms, and bedrooms also shows a significant positive correlation with housing prices. Additionally, building type and the construction age of the house are negatively correlated with housing prices, and this correlation is very significant. Finally, the housing area weakens the positive impact of the number of bathrooms and bedrooms on housing prices.

References

[1] Wu Z K, et al. 2007 Analyzing the impact of housing price factors on homebuyer orientation using the priority factors method. Journal of Tianjin University of Commerce, 27(3).
[2] Hu Q 2017 Analysis of housing price factors based on SVAR model. Time Finance.
[3] Yang D F and Zhang Z M 2013 Empirical research on incorporating housing price factors into China's CPI. Statistics and Information Forum, 28(3).
[4] Lü C Y, et al. 2022 Analysis and prediction of housing price influencing factors based on machine learning. Proceedings of the Third International Symposium on Information Science and Engineering Technology, 117-121.
[5] Yan Z Y and Zong L 2020 Spatial prediction of housing prices in Beijing using machine learning algorithms. In the Proceedings of the 2020 Fourth Conference on High-Performance Computing and Cluster Technology and the Third International Conference on Big Data and Artificial Intelligence (HPCCT & BDAI '20). ACM, New York, NY, USA, 64-71.
[6] Peng Z, Huang Q and Han Y C 2019 Research on the price prediction model of second-hand housing in Chengdu based on XGBoost algorithm: Proceedings of the 11th International Conference on Advanced Information Technology (ICAIT). IEEE.
[7] Pan J, et al. 2023 Analysis and prediction of second-hand housing prices in Qingdao based on ensemble algorithms. Advances in Applied Mathematics, 12(4), 1671-1682.
[8] Wang X J 2019 Study on the impact of second-hand housing prices in Chongqing. Journal of Normal University (Natural Science Edition), 19(3).
[9] Zheng Y F 2007 Study on spatial differences of housing prices in different districts of Hangzhou. Economic Forum, 20, 32-34.
[10] Fan G Z, et al. 2022 Housing property rights, collateral, and entrepreneurship: Evidence from China. Journal of Banking and Finance.
[11] Wang N, et al. 2018 Heterogeneity of the impact of major transportation facilities crossing rivers on housing prices: A case study of Nanchang Riverside New Town. Urban Research Press, 10, 123-130.
[12] Yang C G and Li H B 2019 Population migration, changes in housing supply and demand, and regional economic development: An economic analysis of the current "battles for talents" in domestic cities. Theoretical Research, 3, 93-98.

DOI: 10.54254/2753-8818/52/2024CH0128

Research Progress on the Development of Fatigue Driving Detection Based on Deep Learning

Zhibin Guan

School of Software Engineer, Sichuan University, Sichuan, China

zhibin032@gmail.com

Abstract. Inattentiveness, weariness, and drowsiness are prevalent causes of car accidents; however, timely warnings can prevent these dangerous situations. Advances in artificial intelligence (AI) and computer vision have made it feasible to continuously monitor a driver's condition and alert them when their attention starts to wane. AI technology can assess a driver's level of fatigue by analyzing various facial expressions such as yawning, the closing of eyes, and head movements. Additionally, it can collect data on vehicle behavior and the driver's biological signals. The primary metric for evaluating the effectiveness of these systems is detection accuracy, which indicates how reliably the system can identify signs of driver fatigue. This study provides a comprehensive review of recent techniques for detecting fatigue driving, with a strong focus on deep learning methodologies. It highlights the application and reliability of these advanced technologies in real-world scenarios. Furthermore, the study delves into the latest issues and challenges faced in this field, including limitations in current technology and potential obstacles to widespread adoption. Finally, it proposes suggestions and outlines future research directions to enhance the effectiveness and reliability of fatigue detection systems. This expanded analysis aims to contribute to the development of safer driving environments by leveraging cutting-edge AI and computer vision technologies.

Keywords: Fatigue driving, deep learning, vehicle accidents, neural networks, image detection.

1. Introduction

An important field in image processing and human activity recognition is drowsiness detection. A person's reaction time is greatly slowed down during the period between wakefulness and sleep, which is commonly referred to as drowsiness. The main causes of fatal traffic accidents include driver negligence, fatigue, and sleepiness. These elements have a big effect on road safety because they cause drivers to pay far less attention to the road and have less control over their vehicles. Driving while weary or drowsy is the practice of operating a vehicle while experiencing exhaustion or sleepiness.

Fortunately, it is now possible to detect signs of driver fatigue and drowsiness at an early stage and alert them, which can significantly reduce the occurrence of accidents. This is achieved by identifying various symptoms of drowsiness, such as frequent eye closures and prolonged yawning. Depending on the area of detection, these signs of fatigue can be categorized into four main types. These signs can be detected through images or video sequences captured by cameras monitoring the driver's facial expressions.

DOI: 10.54254/2753-8818/52/2024CH0128

Furthermore, biological signals can be recorded by sensors attached to the driver, and the vehicle's movement and behavior can be monitored. For the same objective, hybrid approaches are also employed. Based on how they are used, Ramzan et al. categorized the state-of-the-art driver drowsiness detection (DDD) methods in a thorough analysis [1]. There exist three primary categories of these techniques: physiological signal-based, vehicle behavior-based, and behavioral parameter-based. The authors examined the most effective supervised methods for identifying sleepiness and talked about the benefits and drawbacks of the three methods they offered. A review research on the identification of driver weariness and drowsiness was carried out by Sikander et al. [2]. There are five types of DDD techniques: hybrid, vehicle-based, physical, biological, and subjective reporting. A study by Otmani et al. used deep learning methods to measure drowsiness and weariness, including recurrent neural networks (RNN) and convolutional neural networks (CNN) [3].

This paper makes significant contributions to the driver fatigue detection systems introduced in recent years, particularly those modern systems, as detailed below:

- Introduces modern deep learning techniques used to detect driver drowsiness for ensuring safe driving.
- Classifies driver fatigue detection systems into four groups based on the tools used to assess drowsiness: eye and facial detection, mouth detection, head detection, and hybrid detection.
- Focuses on the datasets primarily used for detecting drowsiness and introduces their characteristics.
- Compares the applicability and dependability of the four categories for detecting driver weariness.
- explains the drawbacks of drowsiness detection methods, their implications, and the challenges facing the field of detecting driver weariness.

2. Driver Drowsiness Detection (DDD) Systems

When a person is mentally fatigued, they may fall asleep and their attention to tasks decreases. It is crucial to understand that fatigue is not a disease but a state that can be restored through rest and sleep. However, when driving vehicles that require continuous attention, such as cars, buses, or trains, this recovery can become very difficult, potentially leading to serious accidents. Currently, numerous studies are exploring the physiological mechanisms of drowsiness and assessing the degree of fatigue. The literature confirms that there are numerous elements that influence weariness and drowsiness/lethargy in different ways, making it difficult to determine the exact degree of these states. As shown in Figure 1, in order to effectively and reliably depict the driver's condition, fatigue monitoring requires a system that can gather data from various sources and integrate it. Driver fatigue detection systems include several stages:

a. Data Collection: Data is acquired through convenient sensing devices, known as input modules.
b. Preprocessing Module: This module prepares the data by applying landmarks, dimensionality reduction, face detection, feature extraction, feature transformation, and feature selection.
c. Detection or Classification: Relies on machine learning and deep learning techniques for detection or classification.
d. Output Module: This module determines the driver's condition and indicates whether or not they are tired.
e. Alarm Activation Module: Activates an alarm when drowsiness is detected. As shown in Figure 1, this series of processes works in coordination to ensure the driver's safety.

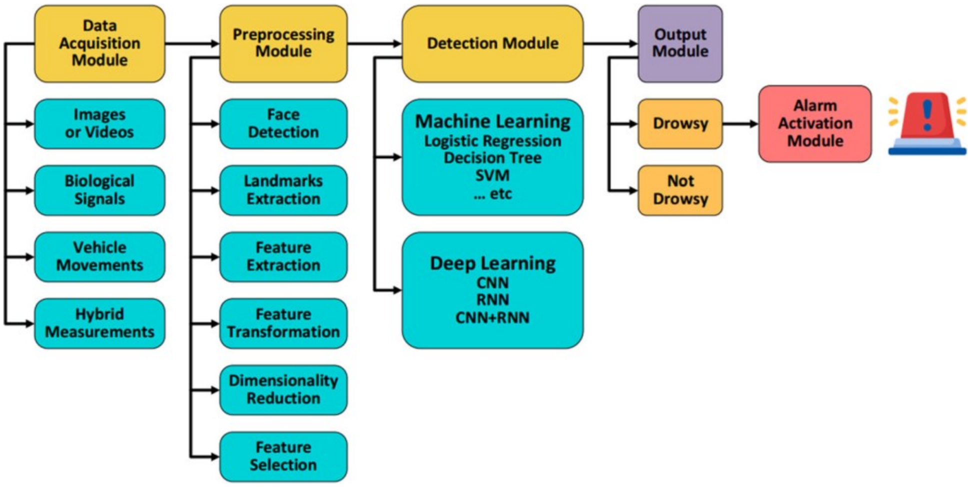


Figure 1. Data flow and general block diagram in DDD systems [4]

2.1. Eye and Facial Detection

For many computer vision applications, including human-computer interface (HCI), facial emotion recognition (FER), and fatigue and sleepiness monitoring, automatic classification of eye states is essential. These applications are mentioned in Figure 2 and Table 2.

Saurav S. et al. concentrated on creating and implementing a vision-based real-time eye state identification system for embedded platforms with limited resources in order to overcome these problems [5]. To extract meaningful data from the ocular regions, the system makes use of two lightweight CNNs. Fine-tuned transfer learning (TL) was used to train the CNNs using a small sample eye state database without overfitting.

60-second picture sequences of the face were employed by Magán E. et al. They created two distinct techniques to determine whether the driver is sleepy in order to reduce false positives [6]. While the second method uses DL techniques to extract numerical features that are then fed into a fuzzy logic-based system, the first strategy relies on RNN and CNN.

Dua M. et al. introduced a driver fatigue detection system that utilizes a structure comprising four deep learning models: VGG-FaceNet, AlexNet, FlowImageNet, and ResNet. To identify drowsiness and weariness, they used RGB videos of the driver as input [7]. They took into account four main types of features while applying these techniques: hand and head movements, behavioral traits, and facial expressions.

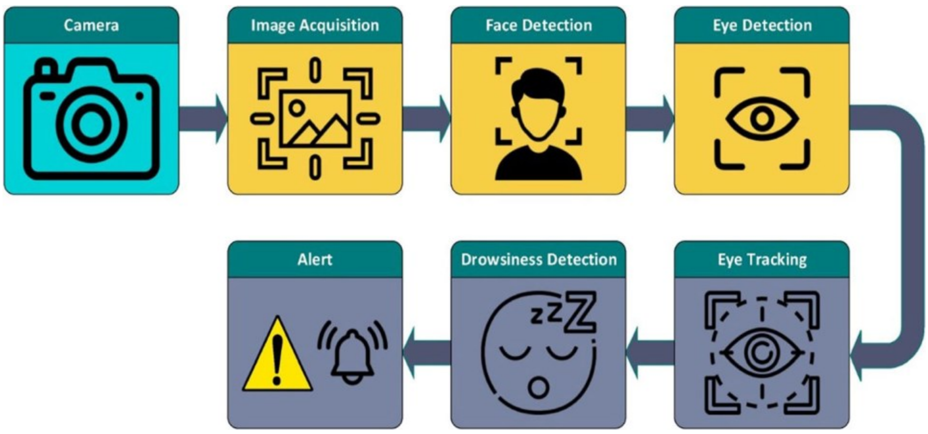


Figure 2. Procedure for observing blinking of the eyes [4]

2.2. Mouth (Yawning) Detection

The several methods for determining tiredness based on mouth and yawning analysis are shown in Table 2 and Figure 3, along with the benefits and drawbacks of each method.

The goal of the work by Savaş BK et al. was to develop a yawning-based driver detection system that uses a CNN model for classification to detect fatigue [8]. The datasets YawDD, Nthu-DDD, and KouBM-DFD were used to train and evaluate the suggested model.

An enhanced algorithm that can recognize the mouth and face was used in the work by Omidyeganeh M. et al. To determine yawning, the back-projection technique was used to quantify the rate of mouth alterations [9]. Congivue's intelligent embedded Apex camera was employed in this project, fully using the high computational power and low memory of the embedded platform. It could only identify specific alertness levels, though.

In the study by Knapik M. et al., a yawning detection fatigue system based on thermal imaging was described. This device can function without disturbing the driver during the day or night because it makes use of thermal imaging [10]. Facial alignment is first accomplished by identifying the corners of the eyes, and yawning is then detected using the suggested thermal model.

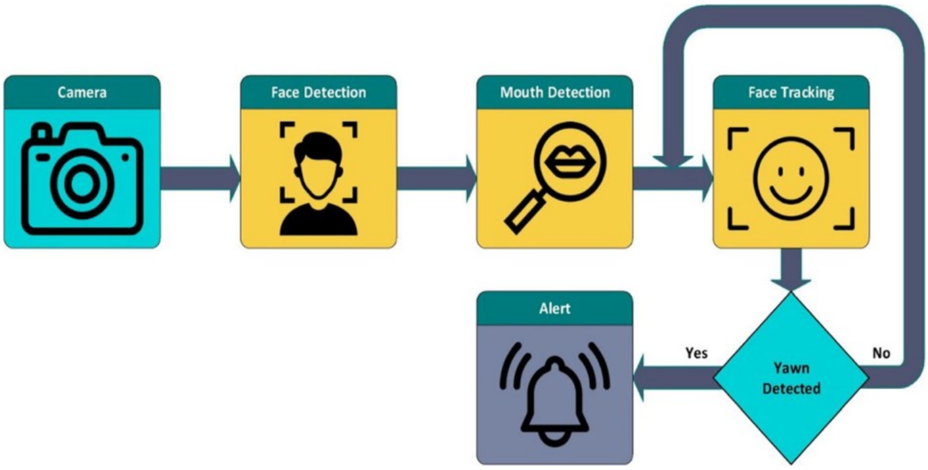


Figure 3. Procedure for observing mouth and yawn [4]

2.3. Head Detection

Numerous techniques have been researched in recent years to detect driver drowsiness. The primary method used by these tactics to identify fatigue is driving behavior observation. Disturbances in a driver's driving style are frequently used to detect driver weariness. Numerous authors cited in Table 2 use gaze and head position variables to detect driver weariness, as illustrated in Figure 4.

Even when the driver wears sunglasses that cover their eyes, the study by Wijnands JS et al. provides an example explaining how to balance real-time inference requirements and good prediction accuracy using deeply separated 3D convolutions and early spatial and temporal feature fusion [11].

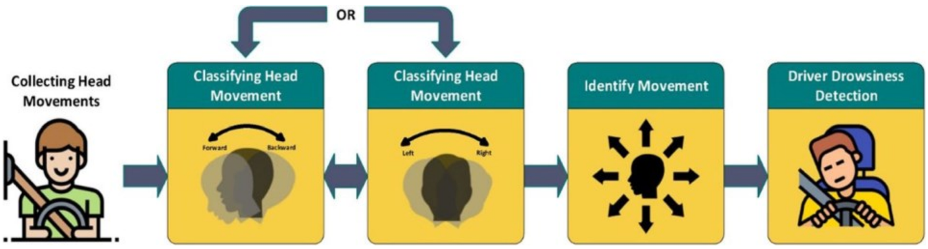


Figure 4. Head pose and gaze observation procedure [4]

DOI: 10.54254/2753-8818/52/2024CH0128

2.4. Hybrid Eye, Mouth, and Head Detection

Table 2 presents driver fatigue detection systems that rely on and incorporate hybrid architectures with parameters from eye, mouth, and head movements.

De Lima Medeiros PA et al. used volunteers' blinking as a human-computer interaction interface and developed an advanced computer vision detector capable of processing information acquired by a standard camera in real-time [12]. Blink detection was carried out using the following processes as an improved technique for eye state recognition: face recognition, modeling, region of interest (ROI) extraction, and eye state categorization. Moving average filter, ROI evaluator, and rotation compensator were other components of this method. Moving average filter, ROI evaluator, and rotation compensator were other components of this method. Two other datasets were also generated: the Autonomous Blinking Dataset (ABD) and the YouTube Eye State Classification (YEC) dataset, which was produced by taking facial photos out of the AVSpeech collection. CNN and SVM models were trained on the YEC dataset, and experiments were carried out on multiple datasets, such as CEW, ZJU, Talking Face, Eyeblink, and ABD, to assess the algorithms' performance.

Moujahid A. et al. introduced a facial monitoring method that relies on compact facial texture descriptors, which can capture the most apparent drowsiness features [13]. This compactness was further improved by applying a feature selection procedure to the initial recovery features and using a multi-scale pyramid facial representation with fundamental sections comprising both local and global information. The model was separated into four stages: (i) facial description utilizing a multi-level multi-scale hierarchy for feature extraction; (ii) pyramid multi-level (PML) facial modeling; (iv) feature subset selection and classification. In the study by Ed-Doughmi Y. et al., RNNs were used to evaluate a sequence of driver facial images and predict the driver's fatigue status. The model was designed and evaluated using a dataset to detect driver drowsiness. A multi-layer model based on a recurrent neural network architecture that relies on 3D convolutional networks was constructed [14].

3. Results Analysis

3.1. Analysis of Commonly Used Datasets

Table 1 lists the most commonly used datasets in driver fatigue detection systems, providing descriptions for each. Various datasets have been utilized in multiple driver fatigue detection systems.

DOI: 10.54254/2753-8818/52/2024CH0128

Table 1. Description of datasets

Ref/Year	Dataset Name	Description
[15], 2016	NTHUDDD public dataset	This dataset includes 36 people from different ethnic backgrounds who participated in a range of driving simulation scenarios wearing and not wearing glasses or sunglasses. These situations include everyday driving, yawning, blinking slowly, falling asleep, laughing fits, and more, and they are performed in both day and nighttime scenarios. • There are five distinct scenarios and eighteen participants in the training dataset. Over the course of about a minute, the slow blink rate, nodding signals, and yawning of each participant are recorded. • Each of the two most important scenarios is demonstrated by a 1.5-minute sequence: one shows a mix of non-drowsiness-related behaviors (such as talking, laughing, and looking in both directions), and the other a combination of drowsiness-related indicators (like nodding, yawning, and slow blinking).
[16], 2020	ZJU gallery	NeuralBody introduced a multi-view dataset known as LightStage. A multi-camera system with more than 20 synchronized cameras was used to record a variety of dynamic human movies, which is how this dataset was created. Driving data was evaluated in both daytime and nighttime scenarios.
[17], 2020	YawDD	An in-car camera was used to record a number of movies featuring actual drivers with a variety of facial features, including men, women, people of different ethnic backgrounds, and even robots. Talking, singing, yawning, and other actions were among the activities. The videos were split up into the following two sets: • The camera was installed beneath the front mirror of the vehicle in the first batch. Three different scenarios are covered in this collection of 322 videos: normal driving (no talking), talking or singing, and yawning while driving. There are three or four videos in each scenario.
[18], 2015	CelebA dataset	There are 200,000 celebrity photos in the CelebA collection, and each one has 40 attributes added to it. With 202,599 face photos, 40 binary feature labels per image, 5 landmark locations, and 10,177 identities, this extensive collection of facial attributes—also referred to as CelebFaces—offers a broad range of attribute annotations.

3.2. Evaluation Metrics

Various metrics have been adopted to evaluate the system's ability to recognize drowsy and fatigued participants, including precision, accuracy, F1 score, and sensitivity. Equation (1) provides the accuracy evaluation metric. Accuracy is the most widely used metric for evaluating driver fatigue detection systems. It provides a good indicator of the system's ability to distinguish between true positives (TP) and true negatives (TN). True positives, true negatives, false positives, and false negatives are denoted by the letters TP, TN, FP, and FN, in that order.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{1}$$

3.3. Experimental Results Analysis

The work on driver drowsiness detection (DDD) that combines deep learning (DL) and machine learning (ML) is shown in Table 2. covering most algorithms based on eye, facial, mouth, and head measurements.

DOI: 10.54254/2753-8818/52/2024CH0128

The table clearly indicates the source, year, measurement area, algorithm classification, implementation environment, accuracy, and dataset for each algorithm.

Following data analysis with the NTHUDDD public dataset shown in Table 2, the following conclusions can be drawn:

The accuracy of deep learning is superior to that of traditional machine learning. This is because deep learning has more parameters, allowing it to learn from more data and features, thus enhancing its learning performance significantly compared to machine learning.

Measuring mouth features plays a crucial role in improving the model's accuracy. Yawning is a prominent and critical feature of driver fatigue, serving as an important criterion for determining whether a driver is fatigued. Therefore, mouth measurement holds significant weight in DDD systems.

Table 2. Image and video-based DDD systems

Ref, Year	Image and video parameters	Classification methods	Accuracy	Datasets
[5], 2022	Eye and Face	Dual CNN Ensemble (DCNNE)	CEW:97.56% ZJU:97.99% MRL:98.98%	CEW, ZJU, MRL
[7], 2021	Eye and Face	Deep-CNN-based ensemble	85%	NTHU-DDD video dataset
[4], 2019	Mouth and Eye	Multiple CNN-kernelized correlation filters method	92%	CelebA dataset, YawDD dataset
[6], 2022	Face	RNN and CNN	60%	UTA-RLDD dataset
[8], 2021	Mouth	ConNN model	99.35%	YawDD, Nthu-DDD and KouBM-DFD datasets
[9], 2016	Mouth	SVM Viola-Jones	75%	Self-prepared dataset
[10], 2019	Mouth	Cold and hot voxels	Cold voxels: 71% Hot voxels: 87%	Self-prepared dataset
[19], 2019	Eye and mouth	Hybrid structure of CNN and LSTM	84.45%	NTHUDDD public dataset
[11], 2020	Facial features, and head movements	3D CNN	73.9%	NTHUDDD public dataset
[12], 2022	Eye, head, and mouth	CNN and SVM	97.44%	new datasets were created: YEC and ABD
[13], 2021	Eye, head, and mouth	SVM	79.84%	NTHUDDD public dataset
[14], 2020	Eye, head, and mouth	Multi-layer model based on 3D convolutional networks	97.3%	NTHUDDD public dataset

4. Discussion

In-depth literature research has revealed various strategies for detecting driver fatigue and mitigating potential issues. With advancements in technology and ongoing research in the field of artificial intelligence, the performance of these systems has significantly improved, successfully addressing many challenges. Considering factors such as invasiveness, interference, cost, ease of use, and accuracy, the practicality of these systems highlights their capability to accurately identify actual levels of drowsiness.

4.1. Image-Based Measurement Methods

Image-based measurement methods are considered practical because they are non-invasive, non-interfering, cost-effective, and do not require sensor setup each time the system is used. Common tools include webcams, smartphone cameras, and thermal cameras. Cameras that gather data are positioned at a specific distance from the driver in order to prevent obscuring their view.

4.2. Relevant Parameters and Accuracy

These systems are dependent on factors associated with weariness, like eye closures, head motions, and yawning. Their accuracy typically ranges from 80% to 99%. However, as mentioned earlier, these systems are influenced by various parameters. Typically, controlled conditions are used for their construction and evaluation, or video records of drivers detecting weariness are used. Therefore, the response to fatigue and its symptoms is mainly simulated, with drivers being prompted to mimic certain symptoms during the data collection phase. In simulated test environments or when using actual databases, the accuracy of image-based systems can exceed 90%. When tiredness manifests visually, it is easy to identify.

4.3. Advantages of Hybrid Systems

Overall, using an appropriate fatigue detection system is the most important step in minimizing car accidents caused by exhaustion. Furthermore, hybrid systems are seen to be the best option for detecting tiredness because of their outstanding track record of dependability.

5. Conclusion

This work offers a thorough analysis and current assessment of systems for detecting driver fatigue, with an emphasis on research findings from 2012 to 2024. Given that this field directly impacts human safety, extensive research is currently being conducted. Given the direct impact on human safety, extensive research is currently being conducted in this field. Driver carelessness or inattention is the primary cause of most traffic accidents. In this review, a comprehensive classification of driver fatigue detection systems is provided, discussing four fundamental methods based on different features for identifying fatigue and drowsiness. These four methods are: eye or facial detection, mouth detection, head detection, and hybrid detection systems. This study presents a comprehensive review of all the systems that are involved, along with the datasets that are utilized, the deep learning models that are implemented, the parameters that are used, and the accuracy of each system's detection. Additionally, typical datasets used for detecting driver fatigue are introduced. Finally, the paper proposes new research directions and potential solutions to address the current issues. Through continuous innovation and improvement, driver fatigue detection systems will be able to more effectively prevent fatigue-related road accidents, ensuring the safety of human lives.

References

[1] Ramzan, M., Khan, H. U., Awan, S. M., Ismail, A., Ilyas, M., & Mahmood, A. (2019). A survey on state-of-the-art drowsiness detection techniques. IEEE Access, 7, 61904-61919. https: //doi.org/10.1109/ACCESS. 2019.2917046

[2] Sikander, G., & Anwar, S. (2018). Driver fatigue detection systems: a review. IEEE Transactions on Intelligent Transportation Systems, 20(6), 2339-2352. https://doi.org/10.1109/ TITS.2018.2878884

[3] Otmani, S., Pebayle, T., Roge, J., & Muzet, A. (2005). Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. Physiology & Behavior, 84(5), 715-724. https://doi.org/10.1016/j.physbeh.2005.02.021

[4] Pandey, N. N., & Muppalaneni, N. B. (2022). A survey on visual and non-visual features in Driver's drowsiness detection. Multimedia Tools and Applications, 1-41. https://doi.org/10.1007/s11042-021-11606-1

[5] Saurav, S., Gidde, P., Saini, R., & Singh, S. (2022). Real-time eye state recognition using dual convolutional neural network ensemble. Journal of Real-Time Image Processing, 1-16. https://doi.org/10.1007/s11554-021-01120-4

[6] Magán, E., Sesmero, M. P., Alonso-Weber, J. M., & Sanchis, A. (2022). Driver drowsiness detection by applying deep learning techniques to sequences of images. Applied Sciences, 12(3), 1145. https://doi. org/10.3390/app12031145

[7] Dua, M., Singla, R., Raj, S., & Jangra, A. (2021). Deep CNN models-based ensemble approach to driver drowsiness detection. Neural Computing and Applications, 33(8), 3155-3168. https://doi.org/10.1007/s00521-020-05448-8

[8] Savaş, B. K., & Becerikli, Y. (2021). A deep learning approach to driver fatigue detection via mouth state analyses and yawning detection. https://doi.org/10.1109/ACCESS.2021.3095074

[9] Omidyeganeh, M., Shirmohammadi, S., Abtahi, S., Khurshid, A., Farhan, M., Scharcanski, J., et al. (2016). Yawning detection using embedded smart cameras. IEEE Transactions on Instrumentation and Measurement, 65(3), 570-582. https://doi.org/10.1109/TIM. 2015.2504018

[10] Knapik, M., & Cyganek, B. (2019). Driver's fatigue recognition based on yawn detection in thermal images. Neurocomputing, 338, 274-292. https://doi.org/10.1016/j.neucom. 2019.02.048

[11] Wijnands, J. S., Thompson, J., Nice, K. A., Aschwanden, G. D., & Stevenson, M. (2020). Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks. Neural Computing and Applications, 32(13), 9731-9743. https://doi.org/10.1007/s00521-019-04194-3

[12] de Lima Medeiros, P. A., da Silva, G. V. S., dos Santos Fernandes, F. R., Sánchez-Gendriz, I., Lins, H. W. C., da Silva Barros, D. M., et al. (2022). Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam. Expert Systems with Applications, 188, 116073. https://doi.org/10.1016/j.eswa.2021.116073

[13] Moujahid, A., Dornaika, F., Arganda-Carreras, I., & Reta, J. (2021). Efficient and compact face descriptor for driver drowsiness detection. Expert Systems with Applications, 168, 114334. https://doi.org/ 10.1016/j.eswa.2020.114334

[14] Ed-Doughmi, Y., Idrissi, N., & Hbali, Y. (2020). Real-time system for driver fatigue detection based on a recurrent neuronal network. Journal of Imaging, 6(3), 8. https://doi.org/10.3390/jimaging6030008

[15] Weng, C. H., Lai, Y. H., & Lai, S. H. (2016). Driver drowsiness detection via a hierarchical temporal deep belief network. In Asian Conference on Computer Vision (pp. 117-133). Springer, Cham. https://doi.org/ 10.1007/978-3-319-54187-7_8

[16] Zhou, K., Wang, K., & Yang, K. (2020). A robust monocular depth estimation framework based on light-weight erf-pspnet for day-night driving scenes. Journal of Physics: Conference Series, 1518(1), 012051. https://doi.org/10.1088/1742-6596/1518/1/012051

[17] Abtahi, S., Omidyeganeh, M., Shirmohammadi, S., & Hariri, B. (2014). YawDD: a yawning detection dataset. In Proceedings of the 5th ACM Multimedia Systems Conference (pp. 24-28). https://doi.org/10. 1145/2557642.2563671

[18] Liu, Z., Luo, P., Wang, X., & Tang, X. (2015). Deep learning face attributes in the wild. In Proceedings of the IEEE International Conference on Computer Vision (pp. 3730-3738). https://doi.org/10.1109/ICCV. 2015.425

[19] Guo, J. M., & Markoni, H. (2019). Driver drowsiness detection using hybrid convolutional neural network and long short-term memory. Multimedia Tools and Applications, 78(20), 29059-29087. https://doi.org/10.1 007/s11042-018-7016-1

Advances in Holographic Theory for Explaining Quantum Entanglement

Jingjing Wang

School of Applied Physics, University of Science and Technology Beijing, Beijing, China

u202141452@xs.ustb.edu.cn

Abstract. The non-locality in quantum entanglement, where particles exhibit instantaneous correlations over large distances, has long been a significant challenge in quantum mechanics. This paper explores the contributions of holographic theory in resolving the non-locality of quantum entanglement and the black hole information paradox. Through the ER=EPR conjecture, holographic theory proposes that quantum entanglement can be explained by tiny wormholes, offering a new geometric perspective to understand non-local effects. Moreover, holographic theory has made substantial progress in addressing the black hole information paradox, particularly through new concepts such as quantum extremal surfaces and replica wormholes. These theoretical approaches indicate that the entropy of black hole radiation follows the Page curve, supporting the idea of information preservation and resolving the paradox without violating the fundamental principles of quantum mechanics. Overall, holographic theory provides crucial theoretical tools and explanatory frameworks for quantum entanglement and black hole physics, advancing the understanding of the unification between quantum gravity and general relativity.

Keywords: Quantum entanglement, holographic theory, AdS/CFT correspondence, holographic wormhole, black hole entropy.

1. Introduction

One of the most astonishing properties of quantum entanglement is its non-locality. It can be said that entanglement is the "breeding ground" for action at a distance [1, 2]. Essentially, remote action is instantaneous and discontinuous, making it impossible to describe through any spatial propagation process. Therefore, action at a distance cannot be approached by any continuous process with finite speed, lacking any characteristics of continuous motion. Moreover, entanglement does not involve the concept of speed; the interaction between entangled particles occurs simultaneously. This phenomenon was referred to by Einstein as "spooky action," with the lower limit of this action's speed being four orders of magnitude greater than the speed of light [3]. However, within the framework of relativity, any superluminal motion and action at a distance are deemed impossible, meaning energy and information cannot travel faster than the speed of light. To explain the simultaneity of wave function collapse, Bell's inequality and Bell's theorem have become crucial for understanding the existence of action at a distance or non-locality. So far, experimental results of Bell's theorem indicate that quantum non-locality or action at a distance truly exists.

Unlike the orthodox Copenhagen interpretation of quantum mechanics, holographic theory provides a new explanation for the principles of action at a distance. Within the framework of holographic theory, it can be assumed that the entangled quantum states are connected in a higher-dimensional space, which can be considered as AdS space. Holographic theory explains the EPR paradox through the AdS/CFT correspondence, where the entangled states in quantum field theory correspond to geometric structures in the higher-dimensional gravitational theory [4]. This means that the information of the entangled states is not local but realized through the geometric structure of the higher-dimensional space, thus explaining the nature of non-locality. Moreover, the holographic principle posits that all the information in a higher-dimensional space can be fully encoded on its lower-dimensional boundary [5, 6]. In this framework, quantum entanglement can be understood as the relationship between the entanglement entropy in the boundary CFT and the geometric background of the higher-dimensional gravity. This non-local transfer of information does not require superluminal propagation but is achieved through changes in the geometric structure.

Contemporarily, holographic theory has achieved several significant advancements in explaining quantum entanglement. Scientists have simulated holographic wormhole states on quantum computers to understand how quantum entanglement manifests in spacetime structures, providing new experimental validation for quantum gravity theories [7]. Additionally, researchers have demonstrated two new types of holographic entropy inequalities related to the topological properties of entanglement wedge nesting, further deepening the relationship between quantum entanglement and the holographic principle [8].

2. Methodology

2.1. Definition of Quantum Entanglement

Quantum entanglement is a quantum state where two or more particles are interconnected in such a manner that their properties cannot be described independently of each other. A unique aspect of entangled states is action at a distance: even if the particles are far apart, measuring one particle will still affect the state of the other particle. Quantum states have the property of superposition. According to the superposition principle, if a quantum state is in state $|\psi_1>$ and state $|\psi_2>$, it can also be in any linear combination of these states:

$$|\psi>=c_1|\psi_1>+c_2|\psi_2> \tag{1}$$

Here, c_1 and c_2 are complex numbers, known as superposition coefficients. Bell states are a set of special states where two qubits (quantum bits) are in a maximally entangled state, serving as an important tool for describing quantum entanglement phenomena. There are four Bell states, which are:

$$|\phi^+>=\frac{(|00>+|11>)}{\sqrt{2}};|\phi^->=\frac{(|00>-|11>)}{\sqrt{2}};|\psi^+>=\frac{(|01>+|10>)}{\sqrt{2}};|\psi^+>=\frac{(|01>-|10>)}{\sqrt{2}} \tag{2}$$

Bell states involve two qubits in a completely entangled state, regardless of the distance between them. These states exemplify the typicality and complexity of quantum entanglement [9].

The strength of entanglement is typically quantified by entanglement entropy. Von Neumann entropy is a measure used to describe the information of mixed states, and the binary von Neumann entanglement entropy is given by the von Neumann entropy of one of the reduced states [10]. For pure states:

$$\mathrm{rh}_{\mathrm{OAB}}=|\mathrm{P_{si}}><\mathrm{P_{si}}|_{\mathrm{AB}},\mathrm{S}(\rho_\mathrm{A})=-\mathrm{Tr}\big[\rho_\mathrm{A}\log\rho_\mathrm{A}\big]=-\mathrm{Tr}\big[\rho_\mathrm{B}\log\rho_\mathrm{B}\big]=\mathrm{S}(\rho_\mathrm{B}) \tag{3}$$

Here, $\rho_A=Tr_B(\rho_{AB})$ and $\rho_B=Tr_A(\rho_{AB})$ describe the reduced density matrices for each partition. In holographic theory, entanglement entropy can be calculated through the geometric structure on the holographic surface.

2.2. EPR paradox and Bell's inequality

The concept of quantum entanglement was first proposed by Einstein, Podolsky, and Rosen in 1935, known as the EPR paradox, to question the completeness of quantum mechanics [3]. Specifically, examine a system composed of two spin-1/2 particles, A and B. After a certain time, A and B become completely separated and no longer interact with each other. When an observer measures a particular spin component of particle A, the spin value of particle B in the corresponding direction can be determined with certainty based on the conservation of angular momentum. Quantum entanglement demonstrates the non-locality in quantum mechanics, and Bell's inequality is an experimental tool to test this non-locality. Under the dual assumptions of locality and realism, this theorem establishes a strict limit on the possible correlation of results when two separated particles are simultaneously measured. Bell's inequality can be written as:

$$\left|P_{xz} - P_{zy}\right| \leq 1{+}P_{xy} \tag{4}$$

If Ax is positive: The spin state of quantum particle A observed on the x-axis is positive. P_{xy} corresponds to the correlation between Ax being positive and Bx being positive.

Additionally, de Broglie believed that the apparent randomness on the surface of quantum effects is entirely due to some unknown variables. If those extra variables are taken into account, the entire system is deterministic and predictable, adhering to strict causality. Such a theory is called the "hidden variables theory [11]. However, experiments have shown that Bell's inequality excludes local hidden variables as a credible explanation for quantum mechanics. As Bell himself stated, "No local hidden variable theory can reproduce all the statistical predictions of quantum mechanics" [2].

As shown in the Fig. 1, for the quantum correlation of spin (assuming 100% detection efficiency), the prediction of local hidden variable theory is represented by the solid line, while the prediction of quantum mechanics is shown by the dashed line. Assuming the angle θ between the two axes lies within the basic case angles mentioned above, the validity of the local hidden variable theory implies that quantum correlations should vary linearly. However, according to quantum mechanics, quantum correlations should vary as the cosine of the angle, cos θ.

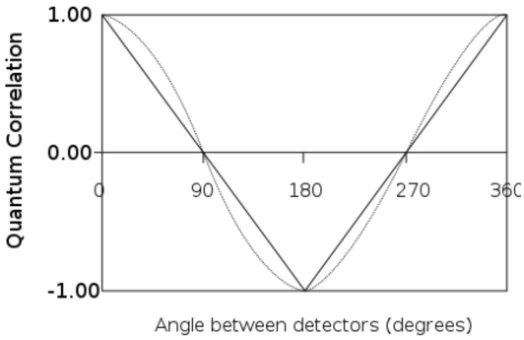


Figure 1. Quantum correlation as a function of the Angle between detectors (degrees) (Photo/Picture credit: Original).

2.3. Holographic theory

Holographic theory is an important theoretical framework in physics used to describe how a physical system in higher-dimensional space can encode its information on a lower-dimensional boundary [5, 6]. The basic concept of holographic theory is that all physical information of a physical system in n-dimensional space can be completely encoded in the (n-1)-dimensional boundary entropy. The core idea of holography is that the universe is an indivisible whole with tightly interconnected parts, where each part contains information about the whole. The implicit order must be extended to a higher-dimensional reality, which in principle is an indivisible whole that contains the entire universe with all its 'fields' and 'particles'. The process of encoding information in higher-dimensional space onto a lower-dimensional boundary is mainly realized through the AdS/CFT correspondence [4]. This means that under certain conditions, physical phenomena in higher-dimensional space can be described by physical theories on the lower-dimensional boundary, thereby greatly simplifying the analysis of complex systems. The metric of the d-dimensional AdS space is:

$$\mathrm{ds}^2 = \frac{\mathrm{L}^2}{\mathrm{z}^2}\left(\mathrm{dz}^2 + \eta_{\mu\upsilon}\mathrm{dx}^{\mu}\mathrm{dx}^{\upsilon}\right) \tag{5}$$

Here, L is the radius of space AdS, z is the radial coordinate, $\eta_{\mu\upsilon}$ is the Minkowski metric. On the boundary CFT:

DOI: 10.54254/2753-8818/52/2024CH0132

$$S_{CFT} = \int d^{d-1}x\, L_{CFT} \tag{6}$$

For a quantum state ρ_A, the entanglement entropy S_A is characterized as:

$$S_A = -\mathrm{Tr}(\rho_A \log \rho_A) \tag{7}$$

To calculate the entanglement entropy of a booundary, one can use the Ryu-Takayanagi Eq. (8):

$$S_A = \frac{\mathrm{Area}(\gamma_A)}{4G_N} \tag{8}$$

Here, γ_A is the minimal surface in AdS space for subregion A and G_N is the Newton constant. This method provides a geometric approach to understanding quantum entanglement phenomena.

Through holographic correspondence, the CFT states on the boundary of AdS space correspond one-to-one with the gravitational states inside the AdS volume. Each quantum state on the boundary maps to a state inside the AdS space [4]. This correspondence can be expressed as the formula:

$$Z_{\mathrm{gravity}}[\phi_0] = \langle e^{\int d^d x \phi_0(x) O(x)} \rangle_{\mathrm{CFT}} \tag{9}$$

where the left side of the equation represents the partition function of the gravitational theory in AdS space, with ϕ_0 as the boundary condition. The right side represents the expectation value in the boundary CFT, with O(x) as the operator in the CFT. This formula expresses how the gravitational theory in AdS maps to the boundary, offering a powerful tool for studying quantum gravity and quantum field theory.

3. Applications

3.1. Holographic Wormhole

The theory of wormholes can, to some extent, explain the non-local effects of quantum entanglement through the ER=EPR conjecture. This conjecture was initially proposed by physicists Juan Maldacena and Leonard Susskind, aiming to connect the spacetime structures in general relativity with the phenomenon of entanglement in quantum mechanics [12]. The core idea of the ER=EPR conjecture is that there exists a tiny, invisible wormhole between each pair of entangled particles. Although these wormholes cannot be traversed in the classical sense, they provide a natural explanatory mechanism at the quantum level, offering a theoretical foundation for the non-local effects observed in entangled particles. The concept of wormholes is based on the spacetime structures permitted by general relativity, which can connect two different points in spacetime. The initial theoretical framework for wormhole formation is the Einstein-Rosen bridge, which proposed a possible bridge between black holes and white holes [13]. The Einstein-Rosen bridge was initially introduced as a mathematical construct to explain the bipolarity of charges (positive and negative charges). It demonstrated a potential spacetime topology allowing two different points to be connected through a "shortcut." The ends of this bridge are located inside the event horizons of two black holes. This model is based on the Schwarzschild solution, a classical solution in general relativity used to describe non-rotating, uncharged black holes.

The Schwarzschild metric represents a static, spherically symmetric gravitational field, with its metric given by:

$$ds^2 = -(1 - \frac{2GM}{r})c^2dt^2 + (1 - \frac{2GM}{r})^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{10}$$

Here, G is the gravitational constant, M is the mass of the black hole, c is the speed of light while (t, r, θ, ϕ) are spherical coordinates. Einstein and Rosen extended the Schwarzschild metric to two symmetrical regions of the universe, thereby constructing a bridge. The mathematical formulation of this bridge involves introducing a new coordinate u such that:

$$u^2 = r - 2GM \tag{11}$$

In this new coordinate system, the metric becomes:

$$ds^2 = -(\frac{u^2}{u^2+2GM})c^2dt^2 + 4(u^2 + 2GM)^2du^2 + (u^2 + 2GM)(d\theta^2 + \sin^2\theta d\phi^2) \tag{12}$$

This indicates that at u = 0, i.e., at r = 2GM, there exists a bridge connecting two different regions of the universe. The holographic principle suggests that certain non-gravitational quantum states can be represented alternatively as higher-dimensional gravitational states, a concept referred to as a holographic dual.

DOI: 10.54254/2753-8818/52/2024CH0132

A team of researchers employed Google's Sycamore quantum processor, which consists of nine quantum bits, to simulate a simplified model of a traversable wormhole [14]. Seen from Fig. 2, the substantial quantum entanglement linking the two exterior regions of a black hole prevents travel between them. This limitation is 'holographically dual' to the prohibition against using entanglement to transmit messages faster than the speed of light. The top half of the image shows two particles, each representing a half quantum state, entangled with each other. The entanglement means the states of these two particles are correlated, irrespective of the distance separating them. The quantum entanglement of these particles has a holographic dual, which is depicted as an emergent wormhole in the lower half of the image. The left and right exterior regions are linked to the left and right quantum state respectively. These regions are parts of spacetime that are connected through the wormhole, illustrating the EP=EPR conjecture which points that each pair of entangled particles is connected by a tiny, non-traversable wormhole.

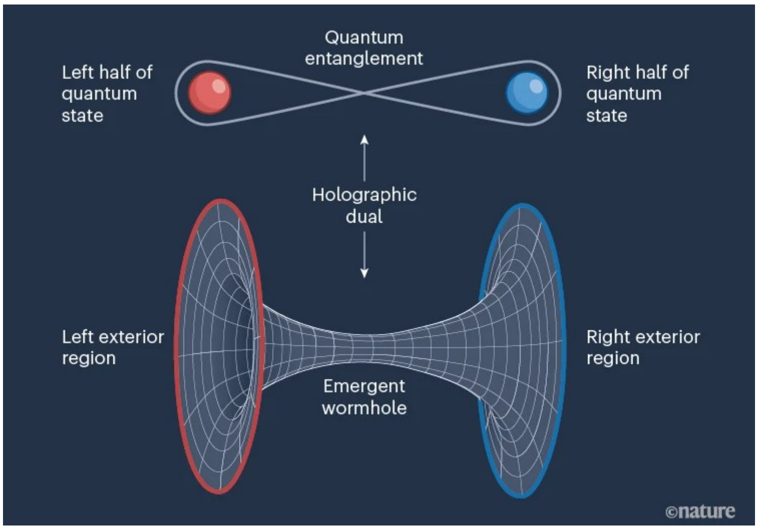


Figure 2. The concept of the EP=EPR conjecture, connecting quantum entanglement and wormholes (Einstein-Rosen bridges) [14].

3.2. Black hole information paradox

The black hole information paradox is one of the major challenges at the intersection of general relativity and quantum mechanics [15]. The paradox arises because black hole radiation (Hawking radiation) leads to the evaporation of black holes, resulting in the information loss, which conflicts with the principle of information conservation. In quantum mechanics, entropy is a measure of the amount of information that is inaccessible or the degree of uncertainty. For a black hole, this refers to the information content of Hawking radiation emitted during its evaporation. Stephan Hawking suggested that black holes emit radiation because of quantum effects occurring near the event horizon. This radiation essentially involves the creation of entangled pairs: one particle escapes the black hole while the other is absorbed into it. This implies that entanglement exists between the black hole and radiation. Scientists have made significant progress in understanding information preservation and the geometric description of quantum entanglement.

Black hole entropy is crucial for comprehending black hole thermodynamics and quantum gravity. According to the Bekenstein-Hawking entropy formula, the entropy of a black hole is proportional to the area A of its event horizon:

$$S_{BH} = \frac{A}{4G_N} \tag{13}$$

This formula indicates that the black hole’s entropy is closely related to the geometric properties of its event horizon. Within the framework of the holographic principle, this relationship is further extended to reveal the deep connections between quantum field theory and gravity. The Ryu-Takayanagi Eq. (8) has a form analogous to the black hole entropy formula. Using the Ryu-Takayanagi formula, black hole entropy can be interpreted as the entanglement entropy in the boundary theory, thereby uncovering the holographic nature of black hole information.

By combining holographic theory and the Ryu-Takayanagi formula (9), one can study the changes in entanglement entropy during the formation and evaporation of black holes. During the evaporation of a black hole, the evolution of the minimal surface can be studied to track the flow of information, providing a geometric description of quantum gravity to address the information paradox. Furthermore, by discussing black hole information within the AdS/CFT framework, information can be considered as being encoded within the entangled structure of the boundary field theory, thus avoiding information loss. The latest research has introduced the island formula, which suggests that during black hole evaporation, certain regions (called islands) can contribute to entanglement entropy, thereby affecting the overall flow of information [16]:

$$\mathrm{S(R)} = \min\left\{\mathrm{ext}\left(\frac{\mathrm{Area}(\partial \mathrm{I})}{4\mathrm{G_N}} + \mathrm{S_{QFT}}(\mathrm{RUI})\right)\right\} \tag{14}$$

Here, R is the radiation region; I is the island region; ∂I is the boundary of the island; $\mathrm{S_{QFT}}(\mathrm{RUI})$ is entropy of the quantum field theory region. The island formula provides a mechanism by which information can still be preserved through quantum entanglement during black hole evaporation, offering a powerful tool for resolving the information paradox. An extension of the Ryu-Takayanagi formula is the quantum extremal surface (QES) method. QES not only considers the area of the extremal surface but also includes the quantum correction terms for the entanglement entropy:

$$\mathrm{S_A} = \frac{\mathrm{Area}(\gamma_A)}{4\mathrm{G_N}} + \mathrm{S_{bulk}}(\gamma_A) \tag{15}$$

Here, $S_{bulk}(\gamma_A)$ is the bulk entropy contribution of the extremal surface γ_A. Through the QES method, it is possible to describe the structure and evolution of entropy inside a black hole more accurately. The Page curve is a theoretical construct that describes the entropy of a black hole's radiation over time, particularly when the black hole evaporates through Hawking radiation [17]. It was introduced by Don Page and has significant implications for understanding the black hole information paradox. Fig. 3 is a simplified depiction of the Page curve.

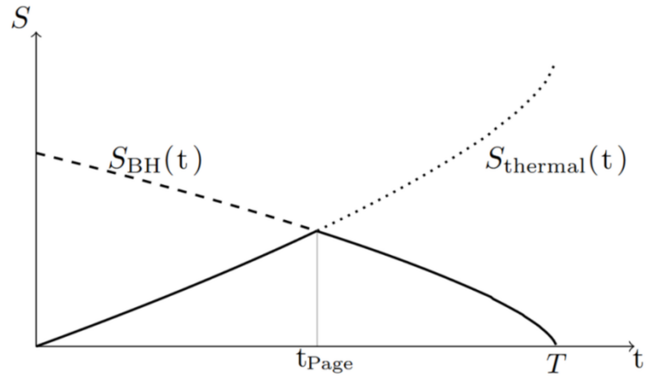


Figure 3. Illustration of the Page curve (Photo/Picture credit: Original).

As the black hole begins to emit Hawking radiation, the entropy of the radiation increases due to the creation of entangled particle pairs. At approximately half the lifetime of the black hole, the entropy reaches its maximum value. After Page time, the entropy decreases, indicating that the remaining black hole has less information content and the information starts being transferred to the radiation until the black hole completely evaporates. The entanglement entropy S of the Hawking radiation before Page time:

$$S_{radiation} \approx \frac{t}{t_{Page}} S_{BH} \tag{16}$$

After Page time:

$$S_{radiation} \approx S_{BH} - \frac{t - t_{Page}}{t_{Page}} S_{BH} \tag{17}$$

Replica wormholes suggest that information can be encoded in the correlation between radiation and the black hole interior. After the Page time, the wormhole configurations dominate, allowing for information to be extracted from the black hole’s interior via spacetime wormholes connecting real and simulated black holes. The replica trick entails computing the von Neumann entropy by analytically continuing the Renyi entropies to n=1. In this approach, gravitational path integrals encompass topologies that link different replicas through spacetime wormholes. This mechanism provides a geometric interpretation of quantum entanglement, where the entanglement entropy of Hawking radiation and the black hole is described by these connected topologies.

4. Limitations and Prospects

Although holographic theory has shown great potential in explaining quantum entanglement, there are still many detailed issues that need to be addressed. The problems that this theory needs to solve in the future can be mainly divided into theoretical, experimental, and application aspects. In the theoretical aspect, the mathematical calculations involved in the holographic principle require more complex and precise mathematical tools, such as CFT and AdS space. CFT is usually defined in high-dimensional spaces, and understanding and handling high-dimensional conformal symmetries and their associated algebraic structures is very complex. Moreover, the algebraic relationships of operators in CFT (such as OPE algebra) are very intricate, and analyzing these algebraic structures requires advanced mathematical tools, such as representation theory and harmonic analysis. The analysis of modal spaces in CFT in high-dimensional cases involves more complex integrals and special functions, and finding appropriate mathematical tools to solve these problems is challenging. On the other hand, AdS space itself is a spacetime with negative curvature, and its geometric properties are quite different from flat space. Analyzing and dealing with the impact of such curvature on physical phenomena requires complex differential geometric tools. In AdS/CFT duality, the boundary conditions of AdS space need to be handled carefully to ensure consistency with CFT. Therefore, boundary theory and geometric methods need to be improved to establish a more accurate holographic theory.

In the problem of the black hole information paradox, how information is transmitted and preserved between AdS and CFT involves many controversies and challenges. According to Don Page's calculations, information is not released uniformly during Hawking radiation but begins to be significantly released after a specific point in time (Page time). The holographic theory has not yet explained this timing and mechanism of information release. Some scholars have proposed the soft hair theory, suggesting that black holes may have "soft hair," i.e., perturbations on the event horizon that can record the information of matter falling into the black hole. This information can be released through Hawking radiation during the black hole's evaporation process. Another hypothesis is the firewall hypothesis, which posits the existence of a "firewall" at the black hole's horizon that destroys any matter approaching the black hole's horizon, ensuring no loss of information. However, this contradicts the smooth horizon hypothesis. Both hypotheses require further theoretical validation.

Experimentally, given the high dimensionality in the calculations of the holographic principle, it is challenging to design suitable experiments to test mathematical predictions. Additionally, holographic theory involves extremely high energy scales and very small spacetime scales, which place extremely high demands on the precision of experimental equipment. Current experimental technology is still difficult to balance with the theoretical level, and only certain aspects of holographic theory can be verified indirectly through experiments. Therefore, designing a direct experiment to verify the core of the holographic theory is a major challenge. The behavior of quantum entanglement in many-body systems is still not fully understood. Although scholars have proposed some possible simulation models, these speculations cannot provide practical significance for the explanation of quantum entanglement. Thus, how the holographic principle can be applied to the quantum entanglement of complex many-body systems is also an unresolved issue. Holographic theory also provides new perspectives for quantum computing, but applying the theory to practical quantum computing devices and algorithms requires further exploration. Of course, the application of holographic theory is not limited to physics;

it can extend to information science, materials science, and other fields. Effectively applying holographic theory in interdisciplinary fields also needs further exploration.

There have been several significant advances in the field of holographic theory and quantum entanglement, with numerous research papers providing insights into current progress and directions. A recent study examines the entanglement entropy of dyonic black holes through the lens of doubly holographic theory. The researchers focused on the entanglement between the eternal black hole and Hawking radiation. They employed doubly holographic theories to analyze the radiation's entanglement entropy, revealing a Page curve that aligns with the principle of unitarity. Additionally, using holographic entanglement density, they demonstrated for the first time that the saturated value of the entanglement entropy is twice the Bekenstein-Hawking entropy when considering a tensionless brane in double holography [18]. Another notable paper discusses the use of quantum gravity concepts to understand teleportation and traversable wormholes. This research emphasizes how holographic theories can simulate quantum teleportation by examining operator growth and its implications for signal transmission across different systems. They explore the potential for simulating holographic quantum gravity with quantum devices and present a specific approach—teleportation by size and the occurrence of size winding [19]. Research on holographic quantum simulations of entanglement renormalization circuits underscores the promise of leveraging holographic techniques to optimize tensor networks on quantum hardware. Utilizing circuits based on the multiscale entanglement renormalization ansatz (MERA), researchers successfully prepare the ground state of an L=32 critical, nonintegrable perturbed Ising model and measure long-range correlations on the ten-qubit Quantum trapped-ion computer. This method enhances the measurement and optimization of quantum systems, advancing the practical application of holographic principles in quantum computing [20].

5. Conclusion

To sum up, holographic theory significantly contributed to addressing quantum entanglement non-locality and the black hole information paradox. By employing the AdS/CFT correspondence, holographic principles reveal profound links between quantum entanglement and geometry. The theory uses the concept of tiny wormholes, or Einstein- Rosen bridges, to explain how entangled particles remain connected across vast distances, thus resolving the non-locality issue. In tackling the black hole information paradox, holographic theory introduces the island formula and replica wormholes. The island formula suggests that the interior of a black hole is part of the entangled system, implying that information is not lost but instead encoded in the radiation emitted. Replica wormholes provide a method to calculate the entanglement entropy of black hole radiation, supporting the idea that information can escape a black hole, thereby preserving quantum coherence. The advancements not only enhance the understanding of fundamental quantum mechanics and gravitational interactions but also offer promising pathways to unify quantum theory with general relativity, potentially leading to a more comprehensive theory of quantum gravity.

References

[1] Horodecki R, Horodecki P, Horodecki M and Horodecki K 2009 Quantum entanglement. Reviews of modern physics vol 81(2) p 865-942
[2] Bell J S 1964 On the einstein podolsky rosen paradox. Physics Physique Fizika vol 1(3) p 195
[3] Einstein A, Podolsky B and Rosen N 1935 Can quantum-mechanical description of physical reality be considered complete? Physical Review vol 47(10) p 777
[4] Maldacena J 1999 The large-N limit of superconformal field theories and supergravity International journal of theoretical physics vol 38(4) pp 1113-1133
[5] Hooft G T 1993 Dimensional reduction in quantum gravity arxiv preprint: gr-qc/9310026
[6] Susskind L 1995 The world as a hologram. Journal of Mathematical Physics vol 36(11) pp 6377-6396
[7] Verlinde H 2021 Wormholes in quantum mechanics arxiv preprint arxiv: 210502129

[8] Susskind L 1995 The world as a hologram Journal of Mathematical Physics vol 36(11) pp 6377-6396
[9] Bennett C H, Brassard G, Crépeau C, Jozsa R, Peres A and Wootters W K 1993 Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. Physical Review Letters vol 70(13) p 1895
[10] Von Neumann J 2013 Mathematische grundlagen der quantenmechanik (Springer-Verlag) vol 38
[11] Bohm D 1952 A suggested interpretation of the quantum theory in terms of" hidden" variables I. Physical Review vol 85(2) p 166
[12] Verlinde H 2020 ER=EPR revisited: on the entropy of an Einstein-Rosen bridge arxiv preprint arxiv:200313117
[13] Einstein A and Rosen N 1935 The particle problem in the general theory of relativity Physical Review vol 48(1) p 73
[14] Brown A R and Susskind L 2022 A holographic wormhole traversed in a quantum computer. Nature vol 621 p 41.
[15] Giddings S B 1995 The black hole information paradox arxiv preprint: hep-th/9508151
[16] Deng F, Chu J and Zhou Y 2021 Defect extremal surface as the holographic counterpart of Island formula Journal of High Energy Physics vol 2021(3) pp 1-26
[17] Almheiri A, Engelhardt N, Marolf D and Maxfield H 2019 The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole. Journal of High Energy Physics vol 2019(12) pp 1-47
[18] Jeong H S, Kim K Y and Sun Y W 2023 Entanglement entropy analysis of dyonic black holes using doubly holographic theory Physical Review D vol 108(12) p 126016.
[19] Nezami S, Lin H W, Brown A R, Gharibyan H, Leichenauer S, Salton G and Walter M 2023 Quantum gravity in the lab II Teleportation by size and traversable wormholes. PRX Quantum vol 4(1) p 010321
[20] Anand S, Hauschild J, Zhang Y, Potter A C and Zaletel M P 2023 Holographic quantum simulation of entanglement renormalization circuits PRX Quantum vol 4(3) p 030334

DOI: 10.54254/2753-8818/52/2024CH0134

Review of GAN-Based Image Super-Resolution Techniques

Huining Feng

Shenzhen MSU-BIT University, Shenzhen, China

1120210272@smbu.edu.cn

Abstract. The objective of Image Super-Resolution (ISR), a significant area of study in computer vision and image processing, is to produce high-resolution images from low-resolution images. The main objective of this paper is to explore the Image Super-Resolution methods based on Generative Adversarial Networks (GANs), especially SRGAN and ESRGAN. In the experimental part, the performance of SRGAN and ESRGAN will be evaluated by using PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index) as evaluation metrics, and the results demonstrate the great potential of Generative Adversarial Networks in Image Super-Resolution tasks, especially in improving the quality of the generated images, and both SRGAN and ESRGAN perform well in recovering image details. Facing the current challenges, researchers can explore new methods and techniques to promote the development and application of image super-resolution technology. The prospect of wide application of GAN models in the field of image processing provides rich opportunities and possibilities for future research and innovation.

Keywords: Image Super-Resolution, Generative Adversarial Networks, Super-Resolution Generative Adversarial Network technique, Enhanced Super-Resolution Generative Adversarial Network technique.

1. Introduction

With the advancement of technology, the quality of images continues to improve, and higher image quality will bring people better visual effects and experience. However, various factors can lead to deterioration of image quality, such as image compression process. Therefore, the deterioration of image quality can be improved by employing super-resolution techniques.

The objective of Image Super-Resolution (ISR), a significant area of study in computer vision and image processing, whose aim is to generate high-resolution images from low-resolution ones. Specifically, it restores and reconstructs as much detail and clarity as possible in an image while enlarging the image size by utilizing computational methods and algorithms. The image super-resolution technique not only visually improves image quality, but also plays an important role in many specialized fields, it can be used in many different fields such as medical imaging, remote sensors [1], surveillance systems and video enhancement.

Although the image super-resolution technique has an important application value, it also faces many challenges. They are mainly reflected in the following aspects:

Estimation of degenerative model: The biggest challenge in image super-resolution restoration is the accurate estimation of the degenerative model. The degenerative model needs to extract motion information, blur information and noise information from low-resolution images. Because of the lack of

DOI: 10.54254/2753-8818/52/2024CH0134

sufficient data, it is very difficult to estimate the degenerative model by relying only on low-resolution images. Most studies focus on designing image priors and ignore the influence of degenerative models on the results. Although many methods perform well under specific conditions, the results on natural images are often unsatisfactory, due to the inconsistency of the degenerative model of the natural image with the hypothetical model. This is one of the reasons why existing algorithms are difficult to be widely used.

Computational complexity and stability: The computational complexity and stability of image super-resolution algorithms are the main factors restricting their wide application. The existing algorithms usually improve the reconstruction quality by increasing the amount of computation, especially when the magnification is large, the amount of computation increases exponentially, resulting in long processing time and limited practical applications. Although some methods can generate high-quality images, the detail reconstruction errors may occur when the input images do not conform to the algorithm's assumptions. In addition, the learning-based methods rely only on the generalization ability of the model to predict high-frequency details when the training dataset is incomplete, making it difficult to avoid the introduction of erroneous details. Therefore, improving the computational efficiency and stability of the algorithm is still an important research direction.

Compression degradation and quantization error: Many images super-resolution algorithms ignore the degradation of quality due to image compression. The transmission of images captured by consumer-grade cameras and Internet images often compress images, which changes the degenerative model of images. The compressed images contain not only additive noise, but also content-dependent multiplicative noise. In addition, most reconstruction methods are based on continuous imaging models and do not consider the digital quantization errors, which can further affect the degenerative model and thus affect the reconstruction results.

Detail reconstruction and visual realism: A key challenge in the process of image upscaling is how to maintain or restore detail without introducing artifacts. The generated high-resolution image needs to be visually natural and realistic, avoiding oversmoothing or unnatural sharpening effects. This requires algorithms to be able to accurately reconstruct details and maintain overall visual consistency while upscaling the image.

Objective evaluation metrics: Currently used evaluation metrics of the super-resolution include Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) [2]. These metrics rely on realistic high-resolution images as a reference to assess the quality of reconstructed images. However, these metrics do not fully reflect the actual results, and there are often cases where the subjective evaluation metrics are high and the objective evaluation metrics are low. In addition, the lack of realistic reference images for natural images limits the application of these metrics. For this reason, it is crucial to develop objective evaluation metrics that do not require reference images and are consistent with subjective evaluation.

The main objective of this paper is to explore image super-resolution methods based on Generative Adversarial Networks (GANs), especially Super-Resolution Generative Adversarial Network(SRGAN) and Enhanced Super-Resolution Generative Adversarial Network(ESRGAN). The results illustrate the great potential of Generative Adversarial Networks in image super-resolution tasks, especially in improving the quality of the generated images, both SRGAN and ESRGAN perform well in restoring image details.

2. Overview of super-resolution techniques

According to the different implementation methods, the super-resolution techniques can be separated into two methods: traditional super-resolution techniques and deep learning-based super-resolution techniques.

2.1. Traditional super-resolution techniques

The traditional image super-resolution techniques include interpolation methods, statistical-based methods and reconstruction-based methods. The interpolation methods [3], such as bilinear interpolation and bicubic interpolation, which are calculated by the values of adjacent pixels, are simple and easy to use, but the generated images are limited in detail and prone to blurring. The statistical-based methods, such as sparse coding and dictionary-based methods, can reconstruct high resolution images by training sparse dictionaries or utilizing pre-trained dictionaries [4], which can generate more details but have the high computational complexity.

The reconstruction-based methods search for the optimal solutions in the high-resolution image space through optimization algorithms and constraints, such as regularization reconstruction and projective reconstruction [5]. This type of method is theoretically robust [6], but the computational complexity is also high. The traditional methods improve the image resolution to a certain extent, but they usually have limitations in detail reconstruction and computational efficiency, and cannot fully meet the needs of practical applications.

Before the emergence of GAN, the image super-resolution methods mainly include bicubic interpolation, sparse coding and dictionary-based methods. The bicubic interpolation [7] is a classical interpolation method that upscale an image by calculating the values of the adjacent 16 pixels through a cubic polynomial, and although the calculation is simple, the generated image usually lacks details and is prone to blurring. The sparse coding methods are based on the theory of sparse representation of images by training sparse dictionaries and using these sparse coefficients, in order to build high-resolution pictures from low-resolution ones, which can reconstruct certain details but have high computational complexity. The dictionary-based methods utilize the pre-trained dictionaries to represent image blocks, and reconstruct the high-resolution images by finding and matching image blocks in the dictionaries, which can effectively reconstruct details but require a large number of computational resources and storage space.

Although these early methods have improved image resolution to a certain extent, they all have their own limitations.

2.2. Deep learning-based super-resolution techniques

Due to the advancements in deep learning, particularly the growth of Convolutional Neural Networks (CNNs), image super-resolution techniques have made breakthroughs. The deep learning-based methods such as CNNs, GANs, Recurrent Neural Networks (RNNs), and Autoencoders have significantly improved the image quality. SRCNN is the first proposed CNN-based super-resolution model, which can learn the mapping from low resolution to high one through multilayer convolutional networks [8]. SRGAN utilizes generative adversarial networks to generate high visual quality images, while ESRGAN takes this a step further to enhance image detail and visual effects. RNNs and LSTMs improve the image reconstruction effect by capturing long-range dependencies, and the autoencoder generates the high-resolution images by learning high-level representations of images.

2.2.1. Overview of GAN. The introduction of GAN, the image super-resolution techniques have made significant progress after CNNs. Deep learning-based methods can generate more realistic and high-quality high-resolution images, greatly improving the restoration of image details and visual effects.

The GAN is a type of deep learning model and is proposed by [9]. Its basic principle is to generate realistic data models by two neural networks competing with each other: Generative Model and Discriminative Model. These two networks compete with each other during the training process, in order to continuously improve the generator's capacity to produce more accurate data, and also to continuously optimize the discriminator's ability to more correctly distinguish between real data and generated data, so they are therefore called "Adversarial" networks. Eventually, the GAN model reaches equilibrium when the data generated by the generators is realistic enough that the discriminators cannot easily distinguish it.

2.2.2. SRGAN Techniques. Despite the increasing depth and speed of CNNs and significant advances in single-image super-resolution as for accuracy and processing speed, how to restore fine texture details at large magnification remains a core challenge. The choice of objective function has a significant impact

DOI: 10.54254/2753-8818/52/2024CH0134

on the performance of traditional optimization-based super-resolution methods. Recent research has focused on minimizing the Mean Squared Error (MSE) of reconstruction. Although this method achieves high PSNR, the generated images usually lack high-frequency details and are not as perceptually effective, as they fail to achieve the desired fidelity and detail richness at high resolution. Therefore, the authors of [10] proposed a GAN for image Super-Resolution (SR), namely SRGAN. Its generator employs a Residual Network (ResNet) structure [11] to mitigate the gradient vanishing problem through skip connections. The generator network inputs a low-resolution image, and then outputs a high-resolution image after multi-layer convolution and activation function processing. Its discriminator uses CNNs to gradually extract the features of the image through multi-layer convolutional layers, and outputs a value between 0 and 1 through fully connected layers and Sigmoid activation function, which indicates the probability of whether the input image is a generated image or a real image.

2.2.3. ESRGAN technique. SRGAN is the first model to apply GAN to image super-resolution, and it achieves significant improvement in visual effect. However, SRGAN still has shortcomings in detail restoration and visual realism. For this reason, the authors of [12] proposed ESRGAN to further improve the quality of super-resolution images. ESRGAN employs a GAN architecture to generate high-resolution images from low-resolution ones, along with several improvements. First, ESRGAN introduces a new perceptual loss function that compares not only the pixel-by-pixel differences, but also the high-level features of the generated images and target images. This method allows the model to generate clearer, which appeals more visually and has a natural feel. Besides, ESRGAN employs a Residual-in-Residual Dense Block (RRDB) architecture in the generator network, which enables the model to capture more complex nonlinear mappings between high-resolution and low-resolution images, thereby significantly improving the image quality. The relative discriminator loss is used to enable the discriminator to learn the relative differences between the generated images and real images, and improve the fidelity of the generated images.

With these improvements, ESRGAN further enhances the details and visual effects of the image while generating high-resolution images, overcomes some limitations of SRGAN and the generated images are visually more natural and realistic.

These methods based on deep learning have significant advantages over traditional methods and are capable of generating more realistic and detail-rich high-resolution images. However, these methods usually require large amounts of training data and computational resources. In practical applications, there is a trade-off between image quality and computational to meet the needs of different application scenarios.

3. Experimental results and analysis

3.1. Basic framework

Figure 1 shows the basic framework of SRGAN and ESRGAN.

The SRGAN framework can be divided into two parts: generator network and discriminator network. The generator network starts from a low-resolution image, passes through multiple Residual Blocks, goes through a convolutional layer, batch normalization and PReLU activation function processing, and finally generates a high-resolution image through an upsampling layer (e.g., Sub-pixel convolution). The generated high-resolution image is fed into a discriminator network that includes multiple convolutional layers, batch normalization and LeakyReLU activation function, and then passes through a fully connected layer and a Sigmoid activation function to output a probability between 0 and 1, which indicates the authenticity of the input image.

The ESRGAN framework has been improved on the basis of SRGAN. The generator network starts with a low-resolution image, passes through multiple RRDB, each of which RRDB contains dense blocks and residual connections, after these processes, generates a high-resolution image through an upsampling layer (e.g. Sub-pixel convolution). The generated high-resolution images are fed into a discriminator network that includes multiple convolutional layers, batch normalization and LeakyReLU

DOI: 10.54254/2753-8818/52/2024CH0134

activation function, and then passes through a fully connected layer, and specifically employs Relativistic Discriminator Loss to output a probability between 0 and 1, which indicates the authenticity of the input image.

During the process of training, the generator and the discriminator are trained alternately. The generator tries to generate more detailed high-resolution images to fool the discriminator, while the discriminator not only distinguishes between real and generated high-resolution images, but also evaluates their relative differences.

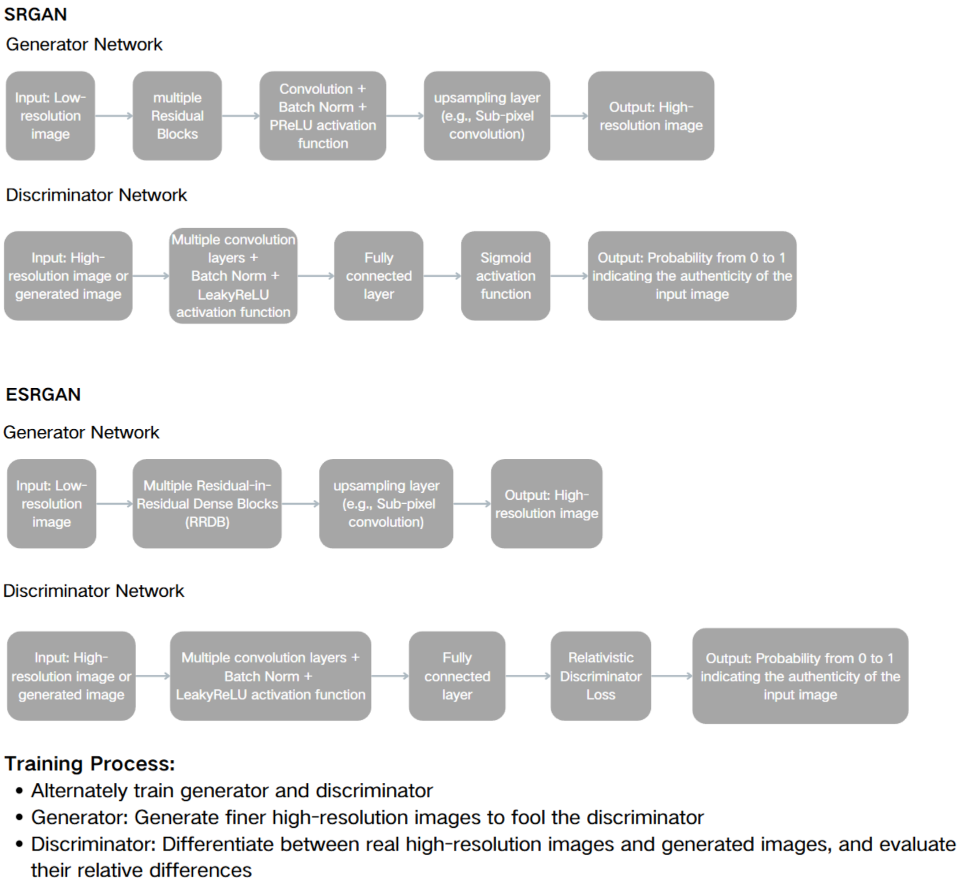


Figure 1. Basic framework of SRGAN and ESRGAN

3.2. Datasets

In order to evaluate the performance of SRGAN and ESRGAN, the datasets used mainly include Berkeley Segmentation Dataset 100(BSD100), Set5 and Set14. These three datasets are the standard datasets commonly used in the field of image processing, especially in image super-resolution tasks.

BSD 100 is a collection of 100 natural images extracted from BSD300, which is rich in image content, including a variety of different scenes and objects, such as natural landscapes, animals, buildings and so on. It provides accurate manual annotation to evaluate the performance of the algorithms in image segmentation and edge detection tasks. Set5 (Set5 Image Dataset) contains five high-resolution images, including close-ups of babies, birds, butterflies, hair, and women's faces. While Set14 (Set14 Image Dataset) contains 14 high-resolution images with a variety of images including animals, buildings, people, etc. Compared with Set5, Set14 provides more image samples, which are suitable for a wider range of evaluations.

In this task, it will be measured by comparing the difference between the original high-resolution image and the model-generated high-resolution image.

3.3. Evaluation Metrics

The paper utilize the PSNR and the SSIM as metrics to evaluate the performance of the two SRGANs and ESRGANs. PSNR is used to measure the similarity between the generated image and the original high-resolution image, and the higher the value of PSNR, the better the quality of the image. Whereas SSIM is employed to evaluate how similar images are in terms of brightness, contrast, and structure, the higher the SSIM value, the better the image quality.

3.4. Analysis of Results

Table 1. Performances of SRGAN and ESRGAN in different datasets

Datasets	Value	SRGAN	ESRGAN
Set5	PSNR	29.4	29.8
	SSIM	0.9019	0.9102
Set14	PSNR	28.49	30.15
	SSIM	0.8184	0.845
BSD100	PSNR	27.68	27.96
	SSIM	0.7620	0.784

As can be seen in Table 1, the PSNR and SSIM values of ESRGAN in all the datasets outperform those of SRGAN, indicating that the images generated by ESRGAN are of higher quality in terms of quantitative metrics. This is mainly due to the improvements of ESRGAN in model architecture and loss function. First, ESRGAN adopts a generative network structure based on residual blocks, which enables the model to better capture the detailed information of the image and reduce image blurring during the training process. Second, ESRGAN introduces that it combines with perceptual loss and adversarial loss, which effectively improves the perceptual quality of the image, so that the generated image not only performs well in objective metrics, but also is more natural and realistic in subjective visual effects. In addition, the ESRGAN generator can better adapt to the characteristics of different datasets during the training process, thereby improving its generalization ability in multiple scenarios.

4. Conclusion

SRGAN and ESRGAN perform well in the image super-resolution task. SRGAN achieves significant enhancement of super-resolution images by introducing generative adversarial networks, while ESRGAN further improves the quality and detail restoration of the generated images by improving the network architecture and the loss function on this basis. The experimental results demonstrate that these models exhibit strong potential in various evaluation metrics.

Future research can continue to go deeper in model optimization and application field expansion. By further optimizing the GAN models and improving the computational efficiency and image quality, they can be applied to more image processing tasks. In the face of the current challenges, researchers can explore new methods and techniques to promote the development and application of image super-resolution technique. The wide application prospect of GAN models in the field of image processing provides rich opportunities and possibilities for future research and innovation.

References

[1] Sumbul, G., Cinbis, R. G. & Aksoy, S. (2019). Multisource region attention network for fine-grained object recognition in remote sensing imagery. IEEE Trans. Geosci. Remote Sens., 57(7), 4929-4937.
[2] Wang, Z., Bovik, A. C., Sheikh H. R. & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process., 13(4), 600-612.
[3] Tsai, R. & Huang, T. (1984). Multiple frame image restoration and registration. Advances in Computer Vision & Image Processing, 317-339.

DOI: 10.54254/2753-8818/52/2024CH0134

[4] Zhang, Y., Wu, W., Dai, Y., Yang, X., Yan B. & Lu, W. (2013). Remote Sensing Images Super-Resolution Based on Sparse Dictionaries and Residual Dictionaries. 2013 IEEE 11th International Conference on Dependable, Autonomic and Secure Computing, 318-323.

[5] Lorent, A. & Cierniak, R. (2014). Regularized Image Reconstruction from Projections Method. 2014 International Conference on Engineering and Telecommunication, 82-86.

[6] Farsiu, S., Robinson, M. D., Elad, M. & et al. (2004). Fast and robust multiframe super resolution. IEEE Transactions on Image Processing, 13(10), 1327-1344.

[7] Rujul, R. & Mehta, D. (2013). Survey on single image Super resolution techniques. IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), 5(5), 23-33.

[8] Dong, C., Loy, C. C., He, K. & Tang, X. (2016). Image super-resolution using deep convolutional networks. IEEE Trans. Pattern Anal. Mach. Intell, 295-307.

[9] Goodfellow, I., Pouget, A. J., Mirza, M. & et al. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 2672-2680.

[10] Ledig, C., Theis, L., Huszr, F., Caballero, J., Cunningham, A., Acosta, A. & et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 105-114.

[11] He, K., Zhang, X. & et al. (2016). Deep residual learning for image recognition. Washington, DC: IEEE Computer Society, 770-778.

[12] Wang, X. T., Yu, K., Wu, S. X., Gu, J. J., Liu, Y. H., Dong C., & et al. (2018). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. arXiv:1809.00219.

DOI: 10.54254/2753-8818/52/2024CH0136

High school student GPA prediction by various linear regression models

Weijia Zhu

Mathematics Science and Statistics Program, University of Toronto, ON, L5L 1C6, Canada

weijia.zhu@mail.utoronto.ca

Abstract. Academic performance (GPA) is a significant index for high school students in North America. The research aims to develop and validate the best predictive regression model to evaluate the GPA of high school students. In addition, the prediction numeric data of GPA can correspond to specific classification data of GPA (Grade Level) to calculate and compare the models' accuracy for predicting Grade Level. The research subjects are high school students from different backgrounds in North America, including their background, study habits, participation in extracurricular activities, etc. The experiment explores the impact of different factors on student GPA and finds that the number of absences from lectures is a key factor in predicting student GPA. Multiple linear regression analysis is used as the main model in the experiment, which may be improved by the stepwise regression methods. The generalization ability of the model is evaluated through cross-validation (CV) methods. Also, the boosting or random forest model is used to be the comparing model for predicting GPA. The experimental result shows that the multiple linear regression model has high accuracy (84%) and reliability (R^2 value is 0.95) in predicting student GPA. The conclusion of the research emphasizes the importance of predicting student GPA in high school education and the potential for guiding educational practice through data analysis. Future work will consider introducing more subjects and variables, such as different subject learning abilities, mental health, and social support, to further improve the predictive accuracy of the model.

Keywords: GPA, linear regression models, stepwise regression, prediction.

1. Introduction

For high school students in North America, there are plenty of memorable experiences and memories in the school. Some students make the most significant decisions in high school, such as becoming an occupational basketball player or singer. However, most high school students may choose to continue studying in the universities or colleges they applied to successfully. In this case, academic GPA is significant for high school students in North America because it is an important index for university admissions decisions [1]. GPA in high school can be calculated by all the course scores which are relative to homework and test scores. In this case, most high school students may hope to have a higher GPA to apply to the university. Also, students who have low GPA hope to change their performance to improve their GPA. However, most students with low GPA do not know how to increase their GPA. In this case, most students hope to understand how to get a higher GPA, and which features of students are most important for improving their GPA.

Sawyer found that GPA in high school is useful for universities to predict high school students' academic success in the future [1]. Kobrin et al. found that the GPA in high school with a correlation of 0.36 is more significant than admission test scores (0.35) in predicting students' first-year GPA in the university [2]. University of Miami researchers found that a point increase in high school students' GPA may affect the increasing 12% - 14% of income in the future [3]. In addition, high school GPA is affected by plenty of factors. Philippe et al. proposed that participation in both civic and non-civic organized activities may positively affect students to increase their GPA in high school [4]. Rahafar et al. found that chronotype, gender, conscientiousness, and test may affect high school student GPA, where gender differences had the most impact on students' GPA by using the structural equation model [5]. Warren showed that students who study in small learning communities could reduce their number of absences and improve their test scores, thereby, increasing their GPA in high school [6]. Therefore, the research plan is to figure out which feature is the most important for affecting high school student's GPA in North America.

In addition, predicting student's GPA by their different features without their course scores and homework performance is the main purpose of the research. In this case, the machine learning methods of the research are significant for predicting student's GPA. Paolo et al. used a logistic multiple regression model to predict the GPA of students and found that cross-validation is important for the generalizability and overall utility of prediction models [7]. Hassan et al. used multiple regression models including linear, ridge, lasso, and LassoCV to predict the GPA of students and found that the LassoCV method is the best model with higher R^2 and lower RMSE in his dataset selected based on the Community of Inquiry framework [8]. Nasiri et al. utilized regression analysis for predicting GPA and the C5.0 algorithm (a type of decision tree) to predict academic dismissal (Classification Data) [9]. Cai et al. found that gradient boosting regression and random forest models improved by the k-fold cross-validation algorithm are better than other models for predicting net ecosystem carbon exchange (NEE) [10].

Based on their ideas of models for predicting data, the experiment decides to use multiple linear regression, random forest, and boosting models improved by the k-fold cross-validation method to predict high school student's GPA, thereby, making the comparison between these model methods to choose the best prediction model.

2. Methodology

2.1. Data source

The dataset collection is from Kharoua in Kaggle about features (Student ID, Age, Gender, Ethnicity, Parental Education, Week Study Time, Absences, Tutoring, Parental Support, Extracurricular, Sports, Music, Volunteering, GPA, and Grade Class) of 2392 high school students in North America.

2.2. Dataset preprocessing

For the details of the dataset, data visualization by histograms is necessary for every feature. In this case, a factor (Student ID), which is the ID for all the subjects, having no distribution in the histogram, being not relevant to the value of the GPA of the students (See Figure 1). Therefore, this factor (Student ID) is dropped before predicting the GPA. The other 12 factors' histograms may show that they may affect the GPA (See Figure 2, Figure 3). In addition, Grade Class (Grade Level) is a classification data of GPA, so Grade Level cannot be used to predict GPA. In this case, Grade Level is dropped in the experiment for predicting GPA. Besides, there are data errors and loss that do not exist since the dataset in Kaggle is reprocessed by Kharoua. The only problem with the dataset is that students' GPA in this dataset are relatively low since Kharoua might collect the data from most students with bad academic performance. The boxplot is necessary for finding outliers for numeric factors but all numeric data in this dataset do

DOI: 10.54254/2753-8818/52/2024CH0136

not have outliers through the boxplot (See Figure 4). Also, the experiment normalizes 12 factors' data (except Student ID, GPA, and Grade Class) in the dataset to eliminate data differences between factors.

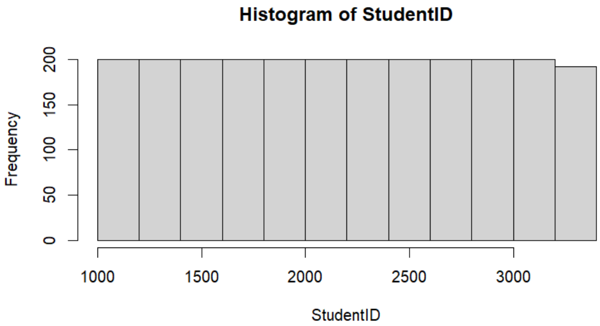


Figure 1. Histogram of Student ID.

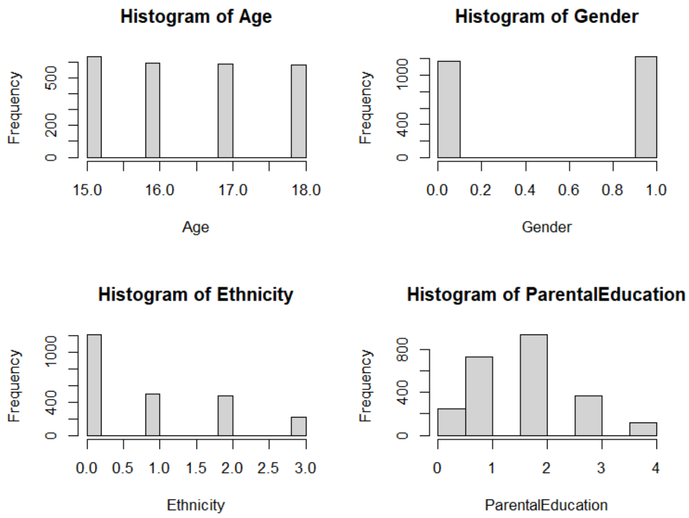


Figure 2. Histogram of age, gender, ethnicity, and parental education.

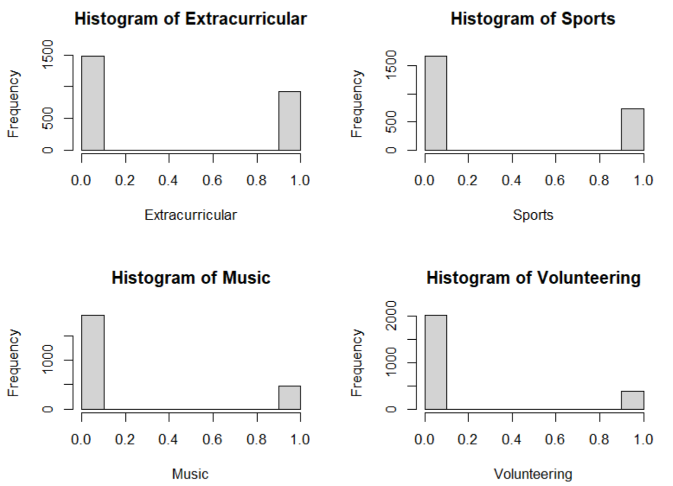


Figure 3. Histogram of extracurricular, sports, music, and volunteering.

DOI: 10.54254/2753-8818/52/2024CH0136

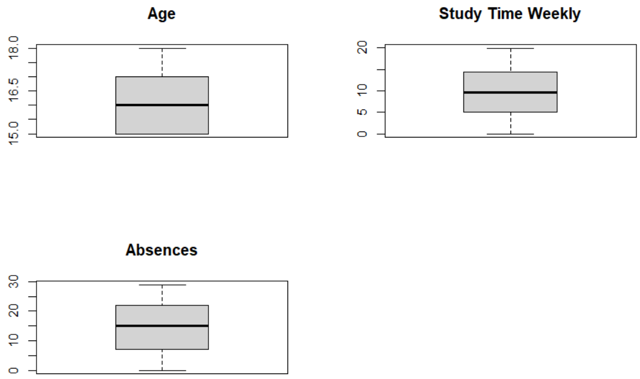


Figure 4. Box plot of age, study time weekly, absences.

2.3. *Variable selection*

The experiment has already dropped 3 factors to have the 12 dataset factors as the variables including Age, Gender, Ethnicity, Parental Education, Week Study Time, Absences, Tutoring, Parental Support, Extracurricular, Sports, Music, and Volunteering. In this case, the experiments make variables have their logograms (See Table 1).

Table 1. Logogram of the 12 factors and target factor.

Factors	Logogram	Meaning
Age	x_1	The age of students (15 – 17 years old)
Gender	x_2	Male (0), Female (1)
Ethnicity	x_3	Caucasian (0), African American (1), Asian (2), and Other (3)
Parental Education	x_4	None (0), high school (1), some college (2), Bachelor's (3), Higher (4).
Week Study Time	x_5	Students' weekly study time in hours (0 to 20 hours)
Absences	x_6	The number of absences during the school year (0 to 30 times)
Tutoring	x_7	No (0), Yes (1)
Parental Support	x_8	None (0), Low (1), Moderate (2), High (3), and Very High (4)
Extracurricular	x_9	Participation: No (0), Yes (1)
Sports	x_{10}	Participation: No (0), Yes (1)
Music	x_{11}	Participation: No (0), Yes (1)
Volunteering	x_{12}	Participation: No (0), Yes (1)
GPA	y	Target factor: the grade value of students (range: 0 – 4)

2.4. *Method introduction*

The experiment utilizes a correlation matrix to determine the feature importance in the dataset. Before creating models to predict GPA, the dataset is divided into two datasets which are the model train dataset (80%) and the test dataset (20%). The experiment trains multiple linear regression models as the base model for predicting the GPA and uses the cross-validation and stepwise regression methods to improve the regression model. For the 12 factors in the model, the multilinearity check is necessary for having a good regression model. In addition, the experiment may use random forest and boosting to create models to compare with the multiple linear regression models to determine the best model for predicting students' GPA. Finally, the experiment calculates the value of mean squared error (MSE) and R^2 from every model for comparison between models. In addition, using the confusion matrix by predicting model data checks every model's accuracy.

3. Results and discussion

3.1. Descriptive analysis

The experiment describes the information of 12 factors and target variable from the original dataset, which includes name, mean, standard deviation (SD), median, trimmed mean (Trimmed), median absolute deviation (Mad), minimum (Min), maximum (Max), range, and standard error (SE) (See Table 2). In addition, the experiment uses the 12 factors' histograms to analyze their distribution.

Table 2. Factor information.

Name	Mean	SD	Median	Trimmed	Mad	Min	Max	SE
Age	16.47	1.12	16	16.46	1.48	15	18	0.02
Gender	0.51	0.5	1	0.51	0	0	1	0.01
Ethnicity	0.88	1.03	0	0.73	0	0	3	0.02
Parental Education	1.75	1	2	1.75	1.48	0	4	0.02
Study Time Weekly	9.77	5.65	9.71	9.73	6.97	0	19.98	0.12
Absences	14.54	8.47	15	14.57	10.38	0	29	0.17
Tutoring	0.3	0.46	0	0.25	0	0	1	0.01
Parental Support	2.12	1.12	2	2.14	1.48	0	4	0.02
Extracurricular	0.38	0.49	0	0.35	0	0	1	0.01
Sports	0.3	0.46	0	0.25	0	0	1	0.01
Music	0.2	0.4	0	0.12	0	0	1	0.01
Volunteering	0.16	0.36	0	0.07	0	0	1	0.02
GPA	1.91	0.92	1.89	1.9	1.07	0	4	0.03

3.2. Correlation analysis

For correlation between the factors and GPA, the correlation matrix is a good method to check which factor is the most important for predicting GPA. In this dataset, the figure shows that absences with a correlation (- 0.92) have the most impact on GPA, which is negatively correlated with students' GPA (See Figure 5).

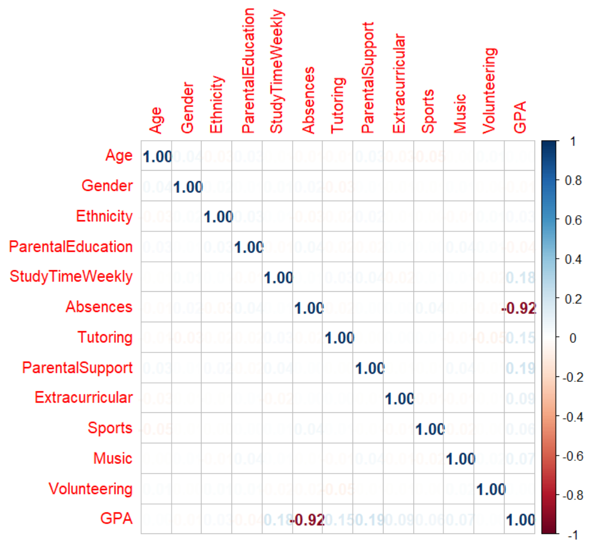


Figure 5. Correlation plot.

DOI: 10.54254/2753-8818/52/2024CH0136

3.3. Model analysis

3.3.1. Multiple linear regression model 1. The experiment uses the training dataset to create a multiple linear regression model (Model 1) between 12 factors and GPA first (See Table 3).

Table 3. Summary of Model 1

Coefficients:	Estimate	Std. Error	t-value	Pr $(>\|t\|)$
(intercept)	2.5214	0.0193	130.303	$< 2 \times 10^{-16}$
Age	-0.0181	0.0120	-1.503	0.133
Gender	0.0089	0.0090	0.992	0.321
Ethnicity	0.0055	0.0131	0.417	0.677
Parental Education	0.0054	0.0181	0.301	0.764
Study Time Weekly	0.5841	0.0156	36.719	$< 2 \times 10^{-16}$
Absences	-2.892	0.0156	-185.981	$< 2 \times 10^{-16}$
Tutoring	0.2468	0.0098	25.073	$< 2 \times 10^{-16}$
Parental Support	0.6002	0.016	37.494	$< 2 \times 10^{-16}$
Extracurricular	0.1957	0.0093	21.034	$< 2 \times 10^{-16}$
Sports	0.1937	0.0098	19.704	$< 2 \times 10^{-16}$
Music	0.1430	0.0112	12.780	$< 2 \times 10^{-16}$
Volunteering	-0.0088	0.0124	-0.706	0.480

3.3.2. Global validation linear model assumption. In this case, the experiment uses the global validation of the linear model assumption function to check whether the experiment needs to improve the linear regression model (See Table 4). In Table 4, the assumption of global stat and skewness is not satisfied, then it shows that some factors do not have a linear relationship with students' GPA.

Table 4. GVLMA of Model 1

	Value	p-value	Decision
Global stat	12.2420	0.0156	Assumptions not satisfied!
Skewness	7.9398	0.0048	Assumptions not satisfied!
Kurtosis	0.6388	0.4241	Assumptions acceptable.
Link function	0.4011	0.5265	Assumptions acceptable.
Heteroscedasticity	3.2623	0.0709	Assumptions acceptable.

3.3.3. Multicollinearity check of Model 1. Checking multicollinearity by using the Durbin-Watson Test (DW test), variance inflation factor (VIF), and kappa coefficient is necessary for finding some factors with autocorrelation. Therefore, there is no multicollinearity in this model because DW test value =1.9664, VIF < 2, and value of kappa < 5 (See Table 5).

Table 5. Value of VIF

Factors	VIF value	Factors	VIF value
Age	1.010	Tutoring	1.007
Gender	1.003	Parental Support	1.006
Ethnicity	1.004	Extracurricular	1.005

DOI: 10.54254/2753-8818/52/2024CH0136

Table 5. (continued).

Parental Education	1.005	Sports	1.008
Study Time Weekly	1.004	Music	1.006
Absences	1.007	Volunteering	1.008

3.3.4. Stepwise regression of model 2. The stepwise regression method is a good idea to reduce several useless factors to create a new linear regression model (Model 2) (See Table 6). In Table 6, the stepwise regression model reduces four factors (Parental Education, Ethnicity, Volunteering, and Gender) to have a new linear regression model.

Table 6. Summary of Model 2

Coefficients:	Estimate	Std. Error	t-value	Pr $(> \lvert t \rvert)$
(intercept)	2.5279	0.0167	151.447	$< 2 \times 10^{-16}$
Age	-0.0180	0.0120	-1.501	0.134
Study Time Weekly	0.5844	0.0159	36.775	$< 2 \times 10^{-16}$
Absences	-2.8912	0.0155	-186.403	$< 2 \times 10^{-16}$
Tutoring	0.2470	0.0098	25.169	$< 2 \times 10^{-16}$
Parental Support	0.6004	0.0160	37.562	$< 2 \times 10^{-16}$
Extracurricular	0.1956	0.0093	21.050	$< 2 \times 10^{-16}$
Sports	0.1938	0.0098	19.718	$< 2 \times 10^{-16}$
Music	0.1429	0.0112	12.799	$< 2 \times 10^{-16}$

3.3.5. Simple linear regression model 3. Figure 2 and Figure 3 show that Absences are the most impact factor for GPA, so it is meaningful for the experiment to create a simple linear regression model (Model 3) between Absences and GPA (See Table 7).

Table 7. Summary of Model 3

Coefficients:	Estimate	Std. Error	t-value	Pr $(> \lvert t \rvert)$
(intercept)	3.3520	0.0164	204.8	$< 2 \times 10^{-16}$
Absences	-2.8801	0.0283	-101.9	$< 2 \times 10^{-16}$

3.3.6. Stepwise linear regression model 4 by 10-fold cross-validation. Table 3 shows that the p-value of five factors (Age, Gender, Ethnicity, Parental Education, and Volunteering) is larger than 0.05, which means that they are not significant for the model. In this case, the experiment utilizes the stepwise regression method by the 10-fold cross-validation to train a new multiple linear regression model (Model 4) (See Table 8).

Table 8. Summary of Model 4

Coefficients:	Estimate	Std. Error	t-value	Pr $(> \lvert t \rvert)$
(intercept)	2.5189	0.0156	161.64	$< 2 \times 10^{-16}$
Study Time Weekly	0.5848	0.0159	36.78	$< 2 \times 10^{-16}$

DOI: 10.54254/2753-8818/52/2024CH0136

Table 8. (continued).

Absences	-2.8911	0.0155	-186.34	$< 2 \times 10^{-16}$
Tutoring	0.2472	0.0098	25.17	$< 2 \times 10^{-16}$
Parental Support	0.5994	0.0160	37.52	$< 2 \times 10^{-16}$
Extracurricular	0.1961	0.0093	21.10	$< 2 \times 10^{-16}$
Sports	0.1946	0.0098	19.83	$< 2 \times 10^{-16}$
Music	0.1429	0.0112	12.80	$< 2 \times 10^{-16}$

3.3.7. Random forest model 5 and 6. Except for the multiple linear regression model, the experiment chooses to use the random forest method to train a new model (Model 5) with all 12 factors for predicting GPA. In addition, Model 4 shows that there are several factors (Age, Gender, Ethnicity, Parental Education, and Volunteering) are dropped, so the experiment trains a relative model (Model 6) with 7 factors left by using the random forest.

3.3.8. Boosting model 7 and 8. Boosting is a good regression method for predicting GPA in this dataset, so a new boosting model (Model 7) is trained with 12 factors. In addition, the experiment trains a relative boosting model (Model 8) with 7 factors left, since the experiment dropped 5 factors in Model 4.

3.4. Comparison analysis

The experiment predicts the value of GPA by different models, thereby, calculating MSE and R^2 by using the prediction data and the test dataset (See Table 9). The experiment uses the prediction of GPA value and test dataset round to the nearest single digit to calculate the accuracy of every model through the confusion matrix (See Table 9).

Table 9. Comparison of Models

Model Methods	Model	MSE	R^2	Accuracy
Multiple Linear Regression Model	Model 1	0.0386	95.6%	83.82%
Stepwise Linear Regression Model	Model 2	0.0388	95.58%	83.82%
Simple Linear Regression Model	Model 3	0.1334	84.79%	69.12%
Stepwise Linear Regression Model	Model 4	0.0388	95.58%	84.03%
Random Forest Model	Model 5	0.0518	94.09%	81.72%
Random Forest Model	Model 6	0.0494	94.36%	81.93%
Boosting Model	Model 7	0.1437	83.62%	67.23%
Boosting Model	Model 8	0.0496	94.35%	82.98%

Compared with other models, Model 4 with only 7 factors has 84.03% accuracy which is the most accurate for predicting Grade Level (See Table 9). Also, Model 1 has the lowest MSE (0.0386) and highest R^2 (95.6%), slightly better than Model 4 in predicting GPA. However, Model 1 is trained by using all 12 factors, which is more complex than Model 4. In this case, the experiment chooses to use Model 4 as the best model in the research (See Function of Model 4 below). Function of Model 4 (See Table 8) (Round to three decimal places):

$$y = 0.585x_5 - 2.891x_6 + 0.247x_7 + 0.599x_8 + 0.196x_9 + 0.195x_{10} + 0.143x_{11} + 2.519 \quad (1)$$

By using the same dataset in Kaggle, Knapp showed that the support vector machine (SVM) is the best model to predict the Grade Level (classification data of GPA), which has 74.5% accuracy.

Compared with his experiment results, Model 4 with 84.03% accuracy in the experiment has increased almost 10% accuracy for predicting Grade Level.

3.5. *Resident test of model 4*

The experiment chooses Model 4 to be the best model in this research, so the experiment needs to use the test plot to check the model assumptions including normal independent identically distribution errors, constant error variance, absence of influential cases, linear relationship between predictors and the outcome variable, and collinearity (See Figure 6). Figure 6 shows that Model 4 successfully passes these five tests. Therefore, Model 4 is a good model that does not need to be optimized deeply.

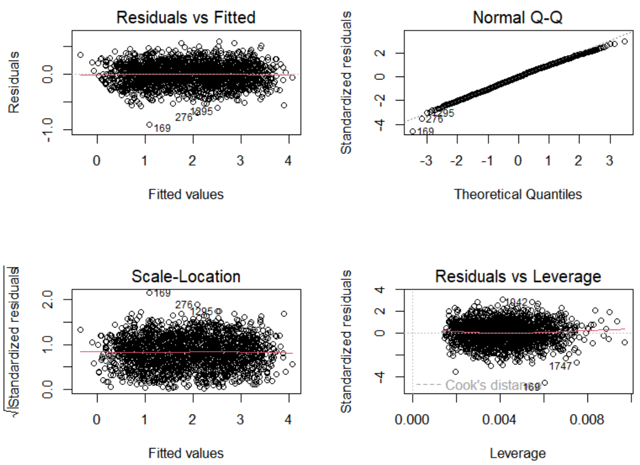


Figure 6. Residual test of Model 4

4. Conclusion

The research works on predicting high school students' GPA by the experiment above. The experiment uses three machine learning methods to train 8 regression models to find the best model for predicting GPA and Grade Level. In this case, Model 4 with 7 factors is the most simple and accurate regression model in the experiment (See the function of Model 4 in the Result). Therefore, Model 4 is useful and important for school administrators to estimate the GPA of students. Also, the experiment can let students know which should be improved to increase their GPA.

For the experiment, several places can be improved in the future. Firstly, subjects in the dataset are not large enough (only 2000+), which may decrease the accuracy of models for predicting data. In addition, plenty of features from students, which is not in this dataset, may significantly affect the GPA in high school, such as subject learning abilities, mental health, and social support. For the machine learning methods, the experiment only uses three different methods, especially for using only one method from the random forest or boosting method. In this case, the research can use more regression methods to train more models to have the best model for predicting the GPA in the future.

References

[1] Sawyer R 2013 Beyond Correlations: Usefulness of High School GPA and Test Scores in Making College Admissions Decisions. Applied Measurement in Education, 26(2), 89-112.
[2] Kobrin J L, et al. 2008 Validity of the SAT® for Predicting First-Year College Grade Point Average. Research Report, College Board.
[3] Miami U O 2014 New Research Shows How Students' High School GPA Could Affect Their Income. Lab Manager.
[4] Philippe F L, et al. 2023 Organized civic and non-civic activities as predictors of academic GPA in high school students. Applied Developmental Science, 27(2), 189-204.

DOI: 10.54254/2753-8818/52/2024CH0136

[5] Rahafar A, Maghsudloo M, Farhangnia S, Vollmer C and Randler C 2015 The role of chronotype, gender, test anxiety, and conscientiousness in academic achievement of high school students. Chronobiology International, 33(1), 1-9.
[6] Warren J E 2016 Small Learning Communities and High School Academic Success. ProQuest Dissertations & Theses.
[7] Paolo A M, Bonaminio G A, Durham D and Stites S W 2004 Comparison and Cross-validation of Simple and Multiple Logistic Regression Models to Predict USMLE Step 1 Performance. Teaching and Learning in Medicine, 16(1), 69-73.
[8] Hassan Y M I, Elkorany A and Wassif K 202. Utilizing Social Clustering-Based Regression Model for Predicting Student's GPA. IEEE Access, 10, 48948-48963.
[9] Nasiri M, Minaei B and Vafaei F 2012 Predicting GPA and academic dismissal in LMS using educational data mining: A case mining. 6th National and 3rd International Conference of E-Learning and E-Teaching, 53-58.
[10] Cai J, Xu K, Zhu Y, Hu F and Li L 2020 Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest. Applied Energy, 262, 114566.

DOI: 10.54254/2753-8818/52/2024CH0127

Smoke Detection and Recognition Based on Flame Feature Extraction

Xinyue Li

School of Automation and Software Engineering, Shanxi University, Taiyuan, China

202102811421@email.sxu.edu.cn

Abstract. The sensor-based fire monitoring scheme requires the installation of temperature and smoke detectors near the monitoring target. Meanwhile, the pictorial fire detection technology is capable of capturing the location and development degree of the fire instantly, enabling rapid detection. Aiming at the issue that the traditional fire monitoring system is prone to failure in large-space indoor and open outdoor environments, a digital image fire monitoring algorithm based on the characteristics of fire flames and smoke is proposed. In the early stage of a fire, smoke is frequently accompanied by its generation and can spread over a large area in the space and is not easily blocked. Therefore, the detection of smoke can realize the early warning of fires and conduct fire rescue work in a targeted manner. In this paper, the background subtraction method is initially employed to extract the moving object, and then the required flame model features are extracted through image preprocessing, and the fire image is identified through multi-feature fusion. The experiment indicates that the flame identification method proposed in this paper can accurately identify the flame under various circumstances and provides a favorable judgment basis for fire alarms.

Keywords: Flame identification, Feature extraction, Object detection.

1. Introduction

Fire is a serious threat to people's production and life safety, at the same time in order to deal with the fire, the input of the fire force is also huge, in recent years, there are many fire rescue operations in the fire sacrifice incident, so improve the existing fire monitoring force, the fire stifle in the cradle, to minimize the fire risk is the current urgent problem.

In order to monitor fire in fire-prone areas, many fire detection techniques have been explored. At present, relatively mature fire detection methods have been found in some places, such as gas-sensitive, temperature-sensitive, smoke-sensitive detectors, in addition to compound detectors, although the physical quantity of detection is increased and the overall performance of the detector is improved, but it cannot effectively eliminate drawbacks. This detection technology has many shortcomings, such as long response time, high equipment cost and single alarm information [1], and there are still technical defects for fire detection in outdoor and large space buildings, and it is difficult to play its due role. For example, these detectors will only issue an alarm when the smoke temperature or gas concentration reaches a certain degree, so the alarm must have a certain delay, which is not conducive to the early rescue of the fire [2].

The image fire detection technology can make up for the shortcomings of these methods, effectively eliminate the influence of interference factors, and is superior to the traditional fire detection technology in the accuracy and reliability of large space fire detection. Image fire detection technology, early mainly focused on the characteristics of flame color identification, such methods cannot distinguish the real flame from similar flame color objects. The fire detection technology based on video image has a unique advantage in quickly and accurately identifying the early flame. Therefore, the fire detection technology of video image has been deeply studied at home and abroad, and more features of flame have been explored and utilized. Li Juhu [3] et al. used the texture features of flame for flame identification, but the texture features of flame are actually unstable, and the extraction and matching of texture features require a lot of calculation work, which is very resource-consuming. Jin [4] et al. used the frame difference method to extract the moving region between two consecutive frames as the suspected flame region. This method can quickly extract some moving objects including flames, but the detected moving objects are not complete. Jin Huabiao [5] determines whether a fire has occurred according to the changes in the area and similarity when the flame spreads. However, to find out the applicable and practical characteristics of the flame image for fire detection, further in-depth research and tests are needed. In addition, for the fire alarm system, various factors should be comprehensively considered, and the speed and accuracy of the fire detector needs to be strengthened. Toreyin [6] et al. used time-domain wavelet change and spatial-domain wavelet change to analyze the flicker of the flame and the color change inside the flame, reducing the probability of false positives. However, since the features of the distant flame may be blocked or weakened, it is impossible to effectively identify the distant flame. This method also easily misidentifies sunlight or other light sources as flames. Mei Jianjun [7] et al. used a more concise visual background extractor algorithm to extract moving objects in images, and segmented moving objects as suspected flame regions, which could realize flame detection in a variety of complex environments. However, direct identification of moving objects as suspected flame regions could not filter out non-flame moving objects, and the recognition accuracy still needed to be improved. These methods can usually detect the flame in most cases, but they are also prone to misjudgment.

The traditional image fire detection method only considers some characteristics of the flame, but does not comprehensively consider these characteristics from the whole point of view, which makes it easy to produce missing or false detection, and the detection speed is relatively slow. In view of this, this paper takes advantage of the changing shape and area of the flame in the early stage of the fire, uses the method of moving target detection to extract the suspected area of the flame, and then extracts the characteristic value of the flame according to the color distribution of the flame, the gradual expansion of the area and the constant change of the edge, and then makes a comprehensive discrimination.

2. Research method

2.1. Extraction of suspected fire areas and contours

The extraction of suspected fire areas and contours is the basis of flame feature extraction and recognition. The quality of flame target extraction plays an important role in improving the recognition accuracy and detection speed of the system. In traditional fire monitoring methods, the target area of the flame is mostly segmented based on the gray value threshold. Although this method can extract the target effectively, it also brings a lot of noise, and increases the calculation workload for the removal of these noises, and also increases the probability of misjudgment. In view of this, this paper uses background subtraction method to initially locate suspected areas in the fire image according to the characteristics of the flame growing from nothing and having movement changes in the initial stage of the fire, and then extracts suspicious areas and contours in the fire image through the area threshold method based on the connected area. This method is fast and can effectively filter out some noise and some static objects with flame color characteristics.

The core of the background subtraction method is to approximate the pixel value of the background image by establishing the parameter model of the background, and then compare the image of the current frame with the background by using the differential comparison method, so as to identify the moving

DOI: 10.54254/2753-8818/52/2024CH0127

region, in which the pixel region with large difference is the moving region, and the pixel region with small difference is the background region [8]. The formula is as follows:

$$diff_{BS} = |f_t - b_t| \tag{1}$$

$$F_{diff_{BS}} = \begin{cases} 1 & if\ diff_{BS} \geq ForeObjThreshold_{BS} \\ 0 & otherwise \end{cases} \tag{2}$$

In the formula:

$diff_{BS}$ is the absolute value of the difference between image and background at time t;

f_t is the pixel value of the image at time t;

b_t is the pixel value of the background at time t;

$F_{diff_{BS}}$ is a binary image with background difference;

$ForeObjThreshold_{BS}$ is the default threshold.

2.2. Feature extraction of flame color

During the combustion process, the color characteristics of the flame are more obvious, and there are more colors in the range of yellow to red. When a fire occurs, there is a distinct visual disparity between the flame area and the background area. The flame luminescence phenomenon is manifested as a red color in the image, and its brightness value is conspicuously higher than that of other surrounding pixels. Therefore, this paper will employ the color characteristics of the flame for the research on fire identification. When extracting flame image features, the collected flame model is stored in the computer by RGB model, which is very beneficial to the extraction of flame features. The 16 images containing flame were analyzed. These images contained about 2.548 million pixels, of which about 824,000 were flame pixels. There are also 8 images that do not contain flames, and these images contain about 1.156 million pixels.

The RGB color model is also referred to as the additive color mixing model. In accordance with the principle of three primary colors, the quantity of light is represented by the primary color light unit. In the RGB color model, any color light F can be obtained through the addition of different components of R, G, and B, and the expression is F = r [R] + g [G] + b [B]. When the three primary color components are 0, F is black light; when all the three primary color components are 1, F is white light. The RGB color model encompasses almost all the colors that human vision can perceive. F is a point in the coordinates of the cube, representing any color, and adjusting the value of any of the coefficients r, g, or b alters the coordinate value of F, thereby modifying its color value. The RGB color space can be depicted by the cube shown in Figure 1.

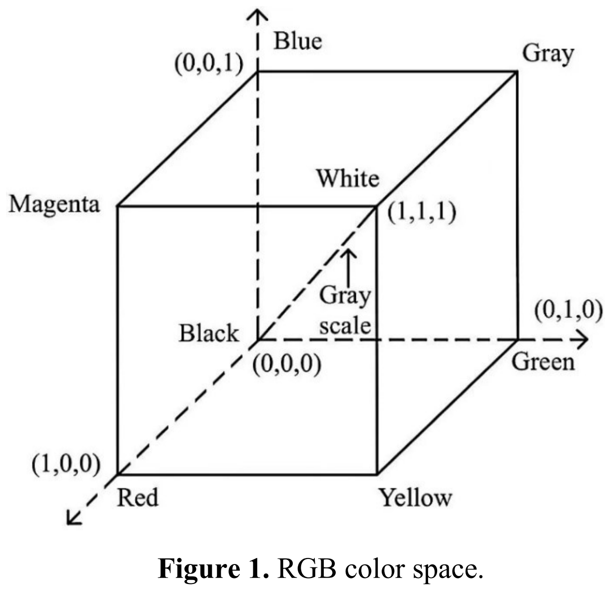


Figure 1. RGB color space.

DOI: 10.54254/2753-8818/52/2024CH0127

2.3. Feature extraction of flame model

The flame model feature extraction algorithm eliminates the influence of strong light or the change of external light intensity on the flame according to the characteristics of the ratio of the three color components of the RGB color model in the flame image. The flow chart of the algorithm is shown in Figure 2:

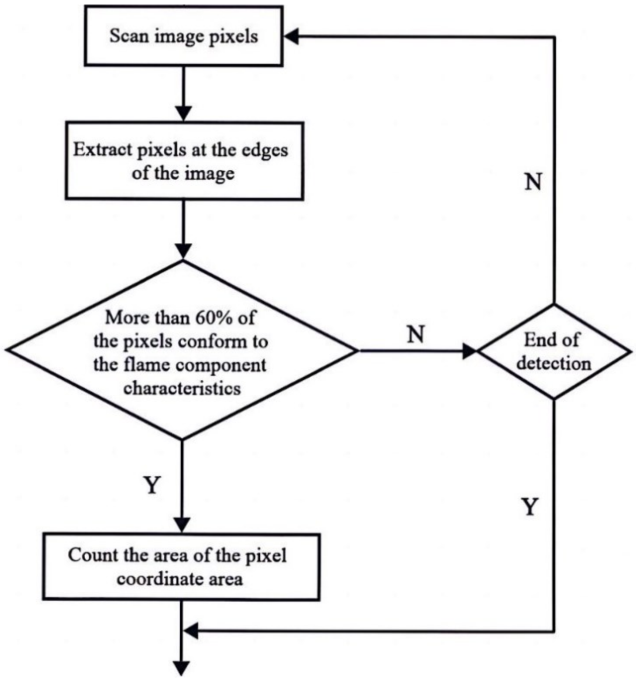


Figure 2. Flow chart of flame model feature extraction algorithm.

As shown in Figure 2, the pixels of the connected area are taken from the top left of the image to the right in turn. If there are no pixels on the right side (that is, the disconnected area on the right side), they are taken down, and all pixels on the image are extracted in this order. If more than 60% of the pixels meet the following characteristics, such a connected region is considered to be a flame region, and the red component value of the peripheral region of the flame will become smaller, that is, if the RGB color component meets the following two conditions, it is considered to be the feature image to be extracted: R (x, y) > G (x, y) > B (x, y) and R (x, y) > 200 G (x, y) < 200, B (x, y) < 100.

2.4. Feature extraction of smoke model

The reference point x is selected by scanning the pixels of the color image and taking the components of the three primary colors of the RGB color model and the H component of the hue of the HSI color model. When the hue of a pixel is in the range of $175° \leqslant H \leqslant 185°$, the parameter α is calculated, and the pixel with the value range of α in $\alpha \in [0,20]$ is selected as the reference point. For pixels whose spatial distance is less than 0.2 (parameter $\alpha \in [0,20]$) and similar to the reference point, the pixels whose spatial distance is >0.2 are reserved as noise cancellation, that is, they are set to white. The algorithm flow chart is shown in Figure 3:

DOI: 10.54254/2753-8818/52/2024CH0127

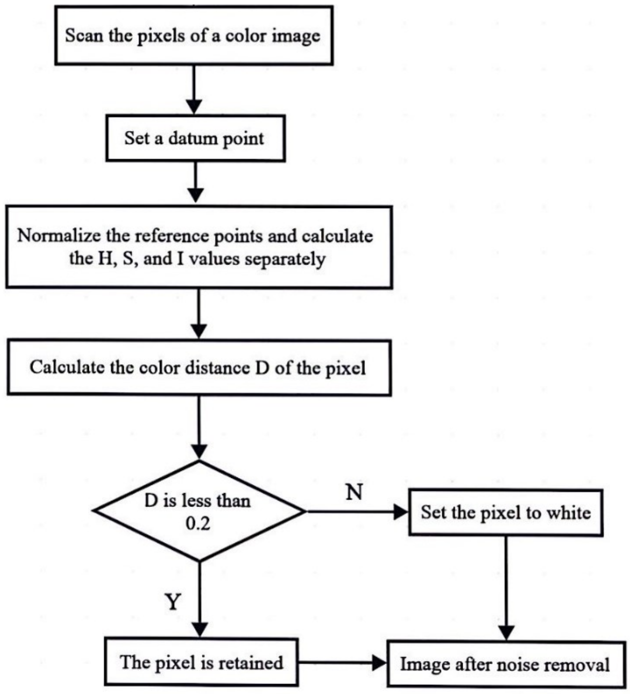


Figure 3. Flow chart of smoke color extraction algorithm.

The spatial distance between the scanned pixels and the reference points is calculated to determine the similarity of HSI colors. Suppose pixel y, its hue, saturation and brightness are h, s, i, x and y respectively, and the space distance between the two points is D, then the space distance D is calculated as follows:

$$D = \sqrt{(V1 + V2 + V3)} \tag{3}$$

In the formula:

$$V1 = (I - i)^2$$

$$V2 = S \times \cos H - s \times cosh \times (S \times cosH - s \times cosh) \tag{4}$$
$$S \times \cos H - s \times sinh \times (S \times sinH - s \times sinh)$$

3. Research result

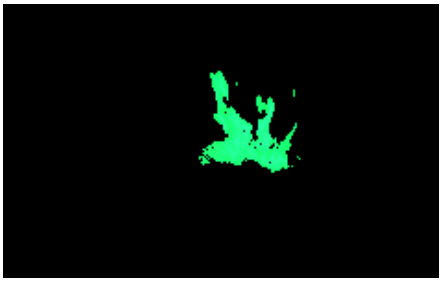
MATLAB is a simulation calculation software, which is widely used in scientific calculation. It has powerful functions such as calculating numerical values and symbols, language programming and simulation drawing. The process of image processing requires a large number of matrix operations, while MATLAB has a strong matrix operation capability, and a large number of functions provided are easier to implement mathematical operations [9]. The proposed algorithm is simulated by MATLAB in the experimental environment of R2016a(9.0.0.341360) and 64-bit(win64). The results are as follows:

3.1. Picture recognition

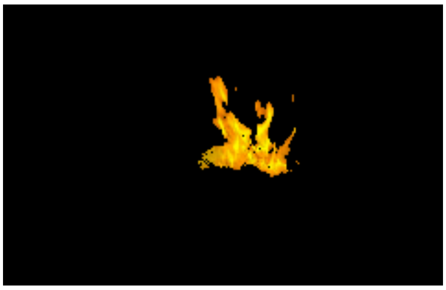
Import a picture and get the number of closed images, whether there is flame, flame roundness and five types of recognition results of this picture, as shown in Figure 4:

DOI: 10.54254/2753-8818/52/2024CH0127

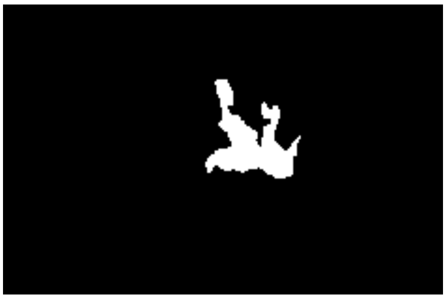
(a) Detection of a picture



(b) HIS conversion diagram



(c) Color segmentation diagram



(d) Binary conversion diagram

(e) Median filter processing diagram

(f) Closed figure diagram

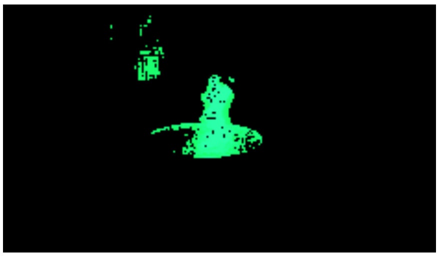
Figure 4. Recognition of an image and five types of graphical results.

3.2. Video recognition

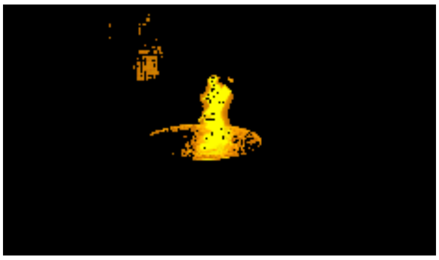
Import a video, select the start frame and end frame to be recognized, and get the number of closed shapes of these selected frames, whether there is a flame, the roundness of the flame and the following five types of recognition results (Figure 5):

DOI: 10.54254/2753-8818/52/2024CH0127

(a) Detection of a frame of a video



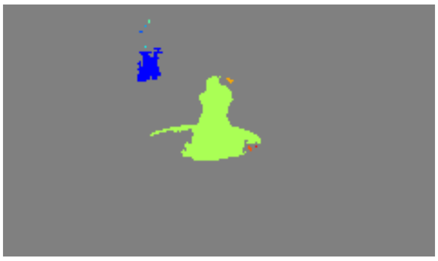
(b) HIS conversion diagram



(c) Color segmentation diagram

(d) Binary conversion diagram

(e) Median filter processing diagram



(f) Closed figure diagram

Figure 5. Recognition of a frame of a video and five types of graphical results.

As depicted in Figures 4 and 5, the background subtraction technique is capable of detecting objects with relatively intact contours and ample fine details. The algorithm boasts low computational complexity and facilitates ease of implementation, thereby satisfying the requirements for real-time video surveillance. In Figure 4, the static background devoid of noise and other interfering factors allows for optimal detection performance. Conversely, Figure 5 presents a challenge due to the figure's coat color closely matching that of the flame, thereby precluding distinct flame identification. Moreover, the recognition outcomes preserve the shadow cast by the moving target, potentially introducing a margin of error in subsequent target tracking processes.

4. Conclusion

The flame color feature is predicated on the triad of primary colorants, i.e., red (R), green (G), and blue (B), within the RGB color model, and the interplay among these components to ascertain the presence of a flame. The smoke model feature is derived from the HIS (Hue, Intensity, Saturation) color space model subsequent to image gray level preprocessing. The Euclidean distance, D, is computed between pixels and reference pixels to diagnose the existence of a fire. The experimental outcomes indicate that the flame and smoke feature extraction algorithm introduced in this paper demonstrates enhanced performance and superior timeliness. It facilitates rapid and efficient monitoring, effectively addressing the challenging task of fire detection within expansive areas. When employing the background subtraction technique for smoke detection, critical considerations include the choice of feature types and sizes. The feature size selections encompass pixel-level, block-level, and cluster-level dimensions, while

DOI: 10.54254/2753-8818/52/2024CH0127

feature types encompass color features, frame features, texture features, and others. The utilization of the standard RGB color feature is susceptible to alterations caused by external environmental factors such as varying light conditions. Consequently, when employing color features to detect moving targets, significant discrepancies may arise in the detection outcomes if there are changes in external lighting. However, the identification robustness of smoke can be significantly improved and the false positive rate diminished by integrating spatio-temporal information, for instance, the inter-frame relationships within a video sequence. Deep learning models, such as the Convolutional Neural Network (CNN) and the Recurrent Neural Network (RNN), are capable of learning intricate feature representations, thereby enhancing the accuracy of smoke recognition. This technology is capable of timely identification of flames and smoke in buildings, forests, and other areas. It monitors the location and scope of fires, offering guidance for firefighting and rescue operations, thereby enhancing the efficiency of fire management. Additionally, it can be employed to detect atmospheric smoke, aiding in the monitoring of air quality and enabling prompt action to mitigate pollution. In the context of surveillance cameras, the technology can identify potential fires or other emergencies, thereby enhancing the efficacy of the security system. In conclusion, this technology, through ongoing research and technological advancement, possesses significant practical utility and broad prospects for application.

References

[1] Mei, J. J. (2020). The early fire detection algorithm based on machine learning research, Tianjin university. DOI:10.27356/d.cnki.gtjdu.2018.001139.
[2] Wang, W. H., Chen, X. B., & Liu, J. L. (2014). Fire detection based on feature fusion of connected region and SVM, Computer simulation, 31(01):383-387.
[3] Li, J. H., Fan, R. X., & Chen, Z. B. (2020). Forest fire image recognition based on color and texture features, Journal of South China University of Technology (Natural Science Edition), 48(01):70-83.
[4] Jin, H., Zhang, R. (2020). A Fire and Flame Detecting Method Based on Video, IEEE, 2347-2352.
[5] Jin, H. B. (2002). Fire detection Technology based on digital image processing, Fire Science and Technology, (03):46-47.
[6] Toereyin, U. B., Dedeoglu, Y., Gueduekbay, U., et al. (2020). Computer vision based method for real-time fire and flame detection, Pattern recognition letters, 27(1):49-58.
[7] Mei, J. J., Zhang, W. (2018). Early fire detection algorithm based on ViBe and machine learning, Acta Optica Sinica, 38(07):60-67.
[8] Liang, X. Y., Liu, K., Wang, X. S., & Liu, H. X. (2023). Research on Video-based forest fire smoke detection algorithm, Forestry Science and Technology, 48(06):45-48.
[9] He, X. H., Zhou, Y. Y., Wang, J. Y., et al. (2006). MATLAB7.x Image Processing, 1st edition., Beijing: Posts and Telecommunications Press, 8-21.

DOI: 10.54254/2753-8818/52/2024CH0140

Research on housing prices prediction-take Boston as an example

Boxuan Liu

Dundee International Institute, Central South University, Changsha, 40086, China

7805220117@csu.edu.cn

Abstract. The problem of predicting Boston housing prices belongs to the field of artificial intelligence, specifically in the domain of regression problems, which is a crucial area of study in machine learning. This article aims to get an accurate prediction model of Boston housing prices by comparing four regression models based on the datasets derived from StabLib library. Multiple Linear Regression model, Random Forest Regression model, Extreme Gradient Boosting Regression model and Support Vector Machine Regression model are taken into consideration. Five evaluation index R-squared, adjusted R-squared, mean absolute error, mean squared error and root mean squared error are compared in terms of generalization ability of model. Eventually, Extreme Gradient Boosting Regression model was found to be the most effective model when predicting housing prices in Boston. The model has certain positive applications in real life, which can help government formulate real estate policies and people make wiser house purchasing strategies.

Keywords: Housing prices, regression model, prediction.

1. Introduction

As an essential aspect of everyday life for the public, housing prices have consistently remained a matter of great importance. From 2004 onward, there has been a sustained nationwide escalation in housing prices in China, making it a central point of interest in the daily lives of individuals [1]. In China, mean house prices almost doubled from 2002 to 2010, indicating substantial fluctuations in house prices over the span of a few years [2]. The fluctuation in housing market prices has affected the daily consumption patterns of both urban and rural dwellers in China [3]. Also, housing price fluctuations significantly inhibit total factor productivity, which hindered economic development [4]. The similar thing happened in America. In the 80s and early 90s of the 20th century, housing prices in the United States maintained a steady upward trend, but Boston's housing prices fluctuated significantly, Boston's housing prices began to rise rapidly in 1983, and Boston's housing prices began to turn around in the third quarter of 1988, and it took four full years to get out of the shadow of the decline, which has had a certain impact on the economic development of Boston area [5]. Hence, housing prices fluctuations not only affect life of residents worldwide, but also hindered economic development in a certain degree. It shows that it is important to get an accurate housing prices prediction model to assist people in assessing the planned purchase of a house and government regulate housing prices for better economic development.

Many house prices prediction works have been carried out. Chen and Qing analyzed the influencing factors of housing prices in different communities in the Boston area in 1980, using quantile regression

DOI: 10.54254/2753-8818/52/2024CH0140

as the basic method, and preliminarily explores the factors that affect the high or low prices of houses in different areas or communities in a region, besides the level of economic development. The use of quantile regression avoids the strict limitations of traditional ordinary least square (OLS) methods on data distribution characteristics and enables targeted research on different data at different quantiles [6]. Wang used the Boston housing price data, this study compares four methods: Levenberg-Marquardt, robust linear regression, Least Mean Squares, and tau. The focus is on analyzing the classic estimation methods and three robust estimation methods, studying the differences and advantages and disadvantages of the four methods [7]. But there exists a shortage of precision. Tian utilized a range of techniques to address the practical issue of predicting house prices in Boston, employing diverse methods for training and testing. It conducted a comprehensive comparison of diverse algorithm models, evaluating their performance from various perspectives and drawing conclusions about their effectiveness. It conducted a horizontal comparison of the advantages and weaknesses of various models and analyzed and summarized the differences in effectiveness [8]. Li and Guo considered the statistical diagnosis and local influence analysis of the single index mode based on data deletion and mode drift models. Simulation studies and results demonstrate the effectiveness and feasibility of the proposed method [9]. This article uses the typical and authoritative Boston housing price data as an example of the application of the semi-parametric varying coefficient quantile regression model and its two-stage estimation, demonstrating the model's good explanatory power in real economic issues [10].

In conclusion, this research compared different algorithmic models from multiple perspectives in order to provide an analysis and summary of their performance in predicting Boston house prices as a regression problem.

2. Methodology

2.1. Data source and description

The data utilized in this paper was sourced from the StatLib library, which is curated at Carnegie Mellon University. It contains 13 continuous attributes, 1 binary-valued attribute and 506 rows.

2.2. Variable selection

The Boston housing price data is relatively clean, so the preprocessing and feature engineering work will be relatively minimal. There are no missing values in the datasets. And the datasets don't contain attributes with low correlation, and there's no need to worry about multicollinearity. Eventually, all features are included. The data that the current paper analyzed contains 13 independent variables and one dependent variable (PRICE). The detailed information of this datasets is displayed in Table 1.

Table 1. Features interpretation.

Variable	Logogram	Meaning
CRIM	x_1	Crime rate per capita in town
ZN	x_2	Percentage of residential land designated for lots exceeding 25000 sq.ft
INDUS	x_3	Ratio of non-retail business acres per municipality
CHAS	x_4	Charles River proximity (1 if tract borders river; 0 otherwise)
NOX	x_5	Concentration of nitric oxides (measured in parts per 10 million)
RM	x_6	Mean quantity of rooms per residence
AGE	x_7	Percentage of residences owned and constructed before 1940
DIS	x_8	Distances weighted by proximity to five employment hubs in Boston
RAD	x_9	Indicator for proximity to radial highways
TAX	x_{10}	Property tax rate per $10,000 of assessed value

DOI: 10.54254/2753-8818/52/2024CH0140

Table 1. (continued).

PTRATIO	x_{11}	Ratio of students to teachers in each town
B	x_{12}	$1000(Bk-0.63)^2$ where Bk represents the percentage of blacks by town
LSTAT	x_{13}	The socioeconomic status of the residents
PRICE	Y	Median value of homes owned by residents in \$1000s

2.3. Method introduction

The paper compares four different methods: Multiple Linear Regression, Random Forest Regression, Extreme Gradient Boosting Regression, Support Vector Machine Regression. After analyzing the advantages and disadvantages of the four methods, the most suitable model was selected to construct a Boston housing prices prediction model based on the datasets.

In machine learning, Multiple Linear Regression is a statistical technique employed to forecast the linear correlation between several independent variables and a dependent variable. Within this method, the association is represented as a linear equation, in which the dependent variable is a weighted total of the independent variables, coupled with an intercept term. Random Forest Regression functions by creating numerous decision trees during the training process and produces the average prediction from each tree for regression assignments. Every tree is constructed using a random subset of the training data and a random subset of the features, aiding in diminishing overfitting and enhancing overall applicability. Extreme Gradient Boosting Regression operates by constructing a sequence of decision trees, with each subsequent tree rectifying the errors of its predecessor. A technique named gradient boosting is used to minimize the loss function by adding new models to the ensemble in a step-by-step fashion. Support Vector Machine regression aims to determine the optimal hyperplane that effectively segregates data points belonging to distinct classes within a high-dimensional space. It is applicable for both classification and regression assignments.

3. Results and discussion

3.1. Multiple linear regression

The mathematical formula for multiple linear regression is:

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{13} x_{13} + e \tag{1}$$

In the above formula: β_0 is a constant term, and e is a residual term.

According to the heatmap, the independent variables do not exhibit highly correlated features among themselves and there are no features with relatively low correlation with the target variable (Figure 1). There is also no need to worry about multicollinearity.

DOI: 10.54254/2753-8818/52/2024CH0140

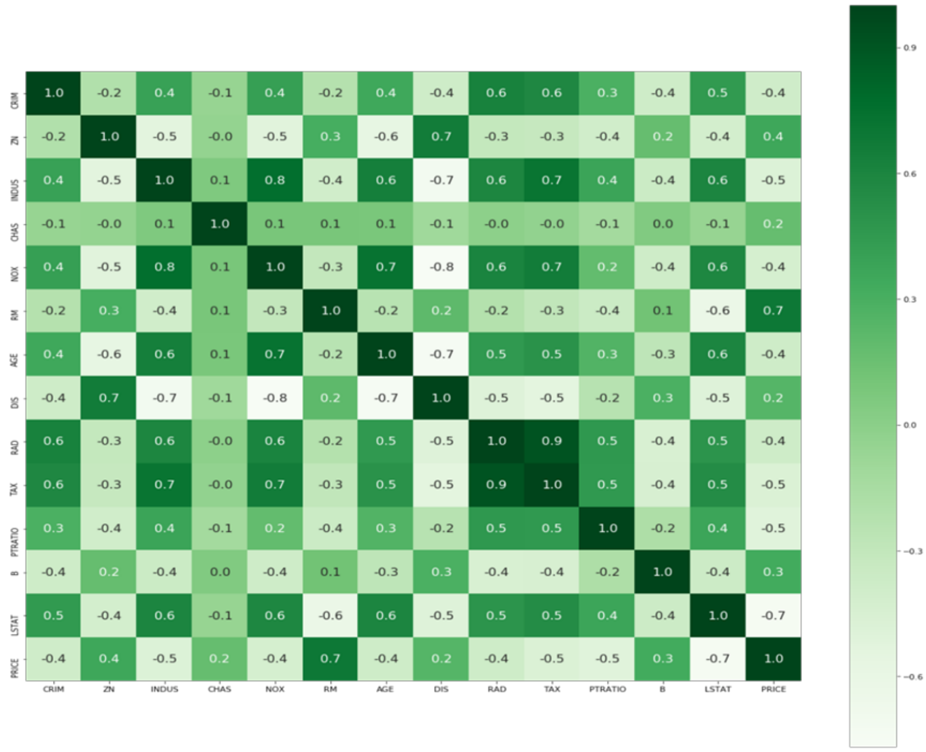


Figure 1. Heatmap of correlation between features.

Table 2. Model results.

	B	S.E.	T	significance
Constant	36.491	5.104	7.149	0.000
X1	-0.107	0.033	-3.276	0.001
X2	0.046	0.014	3.380	0.001
X3	0.021	0.061	0.339	0.735
X4	2.689	0.862	3.120	0.002
X5	-17.796	3.821	-4.658	0.000
X6	3.805	0.418	9.102	0.000
X7	0.001	0.013	0.057	0.955
X8	-1.476	0.199	-7.398	0.000
X9	0.306	0.066	4.608	0.000
X10	-0.012	0.004	-3.278	0.001
X11	-0.954	0.131	-7.287	0.000
X12	0.009	0.003	3.500	0.001
X13	-0.526	0.051	-10.366	0.000

Regression outcomes are shown in the above (Table 2). The significance for the 11 independent variables did not exceed 0.002, which indicates that each of 11 independent variables has a notable influence on the dependent variable Y. To evaluate the model, visualization figure of actual prices compared to predicted prices is shown below (Figure 2).

DOI: 10.54254/2753-8818/52/2024CH0140

Figure 2. Actual prices compared to predicted prices (multiple linear regression).

The visualized graph indicates that this is not a very good model in terms of intuitive effect. The deviation between predicted values and actual values is not slight enough to be ignored, which indicates that the model's fitting effect is not good (Table 3).

Table 3. Model Evaluation (Multiple Linear Regression).

Evaluation	Value
R-squared	0.741
Adjusted R-squared	0.734
MAE	3.273
MSE	21.898
RMSE	4.679

In order to evaluate the model and compare with other three models later more accurately, five data are shown in figure 3. R-squared score is 0.712. Adjusted R-squared score is 0.685. Mean absolute error is 3.867. Mean squared error is 30.068. Root mean squared error is 5.483.

3.2. Random forest regression

Random Forest Regression is renowned for its robustness, flexibility, and capacity to manage high-dimensional extensive datasets, making it a popular choice for predictive modeling in various domains.

Figure 3. Actual prices compared to predicted prices (random forest regression).

DOI: 10.54254/2753-8818/52/2024CH0140

Figure 3 compares actual prices and predicted prices in the training set. If fitting all the points in the figure 3, the line is close to the line y=x. It turns out that the model has good generalization ability by visualization.

Table 4. Model evaluation (random forest regression).

Evaluation	Value
R-squared	0.823
Adjusted R-squared	0.805
MAE	2.469
MSE	18.619
RMSE	4.315

R-squared for Random Forest Regression is 0.822. Adjusted R-squared is 0.805. Mean absolute error is 2.469. Mean squared error is 18.619. Root mean squared error is 4.315 (Table 4).

3.3. *Extreme gradient boosting regression*

Extreme Gradient Boosting Regression is renowned for its high performance, flexibility, and excellent predictive capabilities, making it an essential tool for addressing regression problems.

Figure 4. Actual prices compared to predicted prices (XGBoost regression).

By visualization, it is evident that the Extreme Gradient Boosting model exhibits good fitting and generalization capabilities, notably outperforming the multiple linear regression model (Figure 4).

Table 5. Model Evaluation (Extreme Gradient Boosting Regression).

Evaluation	Value
R-squared	0.849
Adjusted R-squared	0.835
MAE	2.451
MSE	15.716
RMSE	3.964

For Extreme Gradient Boosting Regression model, R-squared is 0.849. Adjusted R-squared is 0.835. Mean absolute error is 2.451. Mean squared error is 15.716. Root mean squared error is 3.964 (Table 5).

3.4. *Support vector machine*

Benefits of Support Vector Machine Regression include efficacy with high-dimensional domains, robustness to overfitting, versatility in kernel selection, and suitability for non-linear data.

DOI: 10.54254/2753-8818/52/2024CH0140

Figure 5. Actual prices compared to predicted prices (SVM egression).

Figure 5 shows the comparison of actual prices with prices predicted by Support Vector Machine Regression model. The scattered distribution of points in the graph indicates that the model has poor fitting and generalization capabilities.

Table 6. Model Evaluation (Support Vector Machine Regression).

Evaluation	Value
R-squared	0.590
Adjusted R-squared	0.551
MAE	3.756
MSE	42.811
RMSE	6.543

In fact, five evaluation index shown in table 6 reflects poor fitting and generalization capability of Support Vector Machine Regression model. R-squared is as low as 0.590. Adjusted R-squared is 0.551. Mean absolute error is 3.756. Mean squared error is as large as 42.811. Root mean squared error is 6.543.

3.5. Comparison of four models

In summary, table 7 gives a direct comparison of four regression models. Upon analyzing the performance of the four models, it is evident that Extreme Gradient Boosting Regression model demonstrates the highest R-squared and adjusted R-squared values, the lowest mean absolute error, mean squared error and root mean squared error. Therefore, Extreme Gradient Boosting Regression appears to be the most effective model among the four models.

Table 7. Comparison Table for Four Models.

	R^2	Adjusted R^2	MAE	MSE	RMSE
Multiple Linear Regression	0.712	0.685	3.867	30.068	5.483
Random Forest Regression	0.822	0.805	2.469	18.619	4.315
Extreme Gradient Boosting	0.849	0.835	2.451	15.716	3.964
Support Vector Machine	0.590	0.551	3.756	42.811	6.543

4. Conclusion

The study applied four different regression methods to the datasets derived from the StatLib library. Upon analyzing R-squared value, adjusted R-squared value, mean absolute error, mean squared error and root mean squared error, Extreme Gradient Boosting Regression was found to be the most effective model to predict Boston housing prices.

By the Boston housing prices prediction model using Extreme Gradient Boosting regression, the government can use the predictive results of the Boston housing price prediction model to formulate real estate policies such as rational planning of housing construction and controlling housing prices to promote sustainable urban development. For the general public and investors, the Boston housing price model can provide them with references when purchasing houses, helping them make wiser home purchasing and investment decisions. However, there are also some drawbacks. The model is based on historical data for prediction, and may not accurately predict the impact of future economic changes, policy adjustments, and other factors. And predictive results of the model are typically aimed at overall trends and may have errors in assessing the specific value of individual properties, unable to fully cover individual differences.

References

[1] Wu Z K, Tang W G and Wu B 2007 Using the Priority Factor Method to Analyze the Impact of House Price Factors on Buyers' Orientation. Journal of Tianjin University of Commerce, 27(3).
[2] Hu Q 2017 Analysis of housing price factors based on the SVAR model. Times Finance.
[3] Yang D X and Zhang Z M 2013 An empirical study on incorporating housing price factors into China's CPI. Statistics&Information Forum, 28(3).
[4] Li J, Chen Y and Guo W 2024 Will housing price fluctuations suppress total factor productivity growth? Empirical analysis based on 259 prefecture-level and above cities in China. Hainan Finance, 5, 3-21.
[5] Yu X F 2002 Potential risks facing the current residential housing industry in Zhejiang from the perspective of international experience: Lessons from the substantial fluctuations in housing prices in Los Angeles and Boston from 1983 to 1993. Journal of Zhejiang University of Technology, 3.
[6] Chen Z and Qing Q 2015 Factors affecting housing prices in different communities in Boston: Based on quantile regression analysis. Business, 30, 278-279.
[7] Wang Y 2019 A review of research on Boston housing prices based on classical and robust methods. Market Weekly, 3, 40-43.
[8] Tian R 2019 Boston housing price prediction based on multiple machine learning algorithms. China New Communication, 11, 228-230.
[9] Li H, Zhu G and Guo Z 2017 Statistical diagnosis of single index mode and its application in the analysis of Boston housing prices. Statistical Research and Management, 6, 1091-1105.
[10] Weng Y 2009 Semi-parametric varying coefficient quantile regression model and its two-stage estimation, Doctoral dissertation, Xiamen University.

Research on house price prediction in UK based on linear regression model

Mengxuan He

Shenzhen Experimental High School, Shenzhen, 518000, China

A02590504@lawsonstate.edu

Abstract. This paper aims to analyze the factors that affect housing prices and predict the trend of housing prices through a linear regression model. This article will use a linear regression model to analyze data on UK house prices from 1952 to 2022. And selected 12 different factors that affect housing prices, using a linear regression model to predict the fluctuations of housing prices. At the same time, a multiple prior regression model was used to explore the importance of each factor in affecting house prices. And the significance of these variables was also compared. The results show that the location of the house, per capita income, housing environment, immigration rate, actual deposit rate and actual loan rate have a significant linear relationship with house price changes, while per capita productivity and house type after the epidemic do not have a significant linear relationship with house price changes. Therefore, UK house prices can be predicted based on the above factors.

Keywords: House price prediction, influencing factors, multiple linear regression model.

1. Introduction

The increase in housing prices is an issue that every citizen has been concerned about [1]. Most people will understand the fluctuations in housing prices in real time and rent or buy a house at a time they think is appropriate, so that the house will be more cost-effective. There are many factors that can cause house prices to rise and fall, and ordinary citizens may not understand these hidden factors, so they cannot judge the price of a house well. Therefore, they cannot judge the price of a house well, so they may buy or rent a house at the highest point of the price or at an inappropriate time. This article will predict the fluctuations of house prices in the UK by studying these factors.

Given that each location and house size are different, the corresponding prices are also different [2, 3]. From 2004 to 2020, the average salary of British citizens fell by 0.97 times, which also caused the house price in the UK to fall to 0.96 times the original [4]. Incomes vary greatly in different parts of the UK. In order to cater to the economic strength of residents in different regions, the government will make different fine-tuning of housing prices in different regions. Which also leads to differences in house prices. However, the overall house prices do not differ too much, but the ratio of house prices to income varies a lot [5]. The COVID-19 outbreak in the UK in 2020 caused a drop in house prices at the time, which was undoubtedly a huge blow to the UK's economic development, which also led to a drop in house prices at the time. However, after the outbreak, the recovery of productivity prompted the recovery of citizens' personal economy, which in turn promoted, but after the outbreak, the recovery of productivity caused the rate of increase in UK house prices to gradually increase [6]. The type of housing in the UK also affects the fluctuations in house prices. For example, apartments are cheaper than single-family homes [7]. Given the surrounding environment, people may choose to move to improve the surrounding environment of their residence or increase the value of their house, which can also cause house prices to rise or fall. People prefer residences with convenient transportation and shopping [8]. The local immigration rate and emigration rate will also affect the housing price changes to a certain extent. At the same time, the actual interest rates of personal deposits and loans also lead to differences in personal economic capabilities [9, 10]. Through the data collected, it is found that the actual amount of money that homeowners can receive from selling their homes is gradually decreasing, which will cause most people to be unwilling to sell their homes, resulting in a shortage of housing [11].

In summary, this article will focus on the impact of these 12 factors on UK house prices, and select corresponding models to study the correlation between these factors and house prices.

2. Methodology

2.1. Data source

This study selected UK house price data from 1952 to 2022. The data in this article are all from Kaggle website (UK house prices), nationwide website and Bank of England website. The dataset from Kaggle is called UK Average House Prices and Salary (1975-2020) which is published in 2021. The images are all come from nationwide website and Bank of England website, which provide many research and interpretation of economic. Through the relevant information of the Economic Help website, which provide data for senior high school students to study, and the data of multiple websites are combined for analysis

2.2. Variable selection

House prices are a reflection of the local GDP. At the same time, house prices are also affected by many different factors, most of which are uncontrollable. And because there are too many factors, house prices fluctuate very frequently. However, the author can determine the approximate fluctuation trend based on certain factors, as the illustrated in Figure 1:

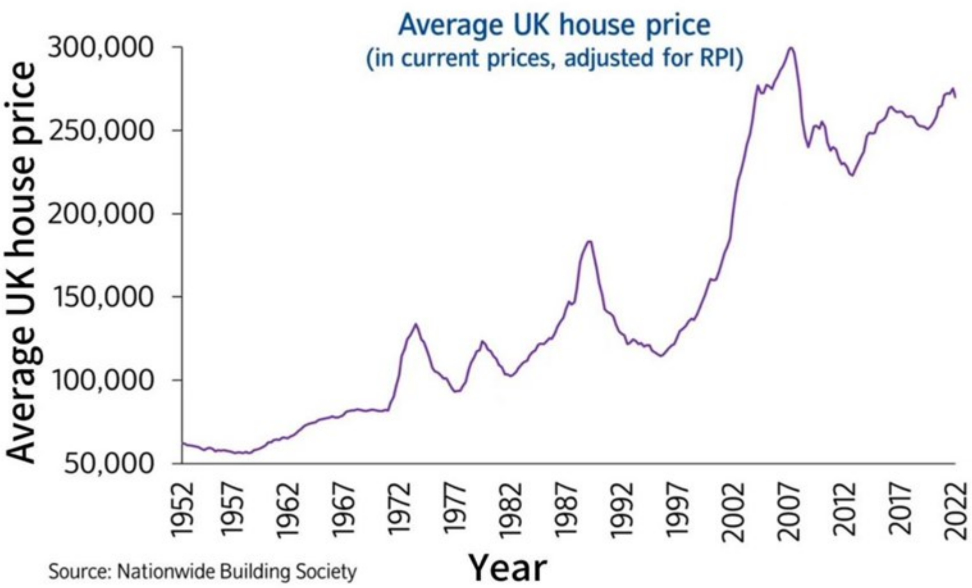


Figure 1. Average UK house price.

DOI: 10.54254/2753-8818/52/2024CH0146

As shown in Figure 1, before 1997, house prices rose steadily, experienced some ups and downs in the middle, and then experienced a significant rise and fall. The house price in 2022 is 1.1 times that of 2004, and the cycle is nearly 5 years. But overall, the average house price in the UK is still on an upward trend. As can be seen from Figure 1, UK house prices reached their highest point in 2007 and then began to plummet. Through the collection of data, the author can understand that the economic recession caused by the financial tsunami on August 9, 2007, investors lost confidence in the value of securities, triggered a liquidity crisis, and the world's developed economies (especially Europe, South America, North America) fell into absolute recession, which led to a sharp drop in house prices starting in 2007. Table 1 shows the variables used in this paper.

Table 1. List of variables.

Variable	Logogram	Meaning
Region	$x1$	Area where the house is located
Salary	$x2$	Average salary
Productivity	$x3$	Post-pandemic productivity per capita
Building Type	$x4$	Flats(1), Terraced(2), Semi(3), Detached(4)
Surroundings	$x5$	Criminal record of the surrounding area
Immigrate	$x6$	Local immigration rate
Deposit	$x7$	Actual deposit rate
Loan	$x8$	Actual loan rate
House Price	Y	Total Housing prices in UK

2.3. Method introduction

This paper uses a linear regression model to compare the impact of different factors on average house prices. Linear regression is a statistical model which estimates the linear relationship between a scalar response and one or more explanatory variables. In linear regression, the relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data.

Linear regression was the first type of regression analysis to be studied rigorously, and to be used extensively in practical applications. This is because models which depend linearly on their unknown parameters are easier to fit than models which are non-linearly related to their parameters and because the statistical properties of the resulting estimators are easier to determine.

3. Results and discussion

3.1. Data processing

From Figure 2 below, this paper can see that the growth of housing prices and dates is positively correlated, and there is a strong linear correlation. The R-square value of the best fit line in Figure 2 is 0.442. And the p-value is less than 0.05, so this regression model is meaningful.

DOI: 10.54254/2753-8818/52/2024CH0146

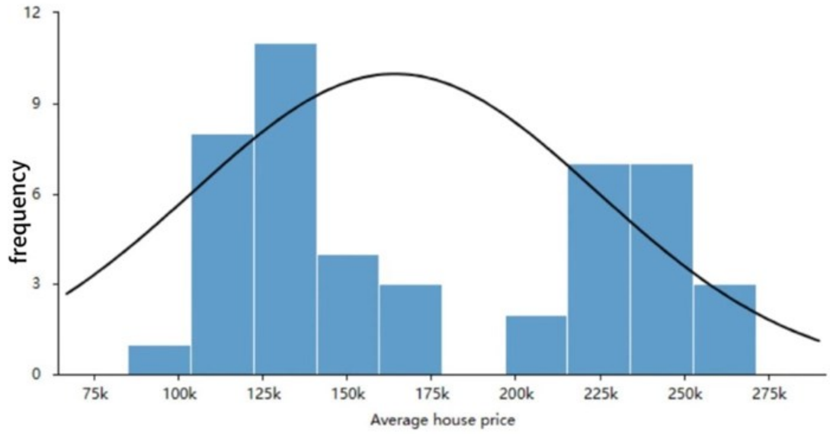


Figure 2. Scatter plot of date and average house price.

The linear fitting formula for scattered data in figure 2 is: $average\,price = -382091.090 + 0.000 * date$, R is 0.442. From table 2, the author can see that when linear regression analysis is performed with date as the independent variable and average price as the dependent variable, the model R-square value is 0.442, which means that date can explain 44.2% of the changes in average price.

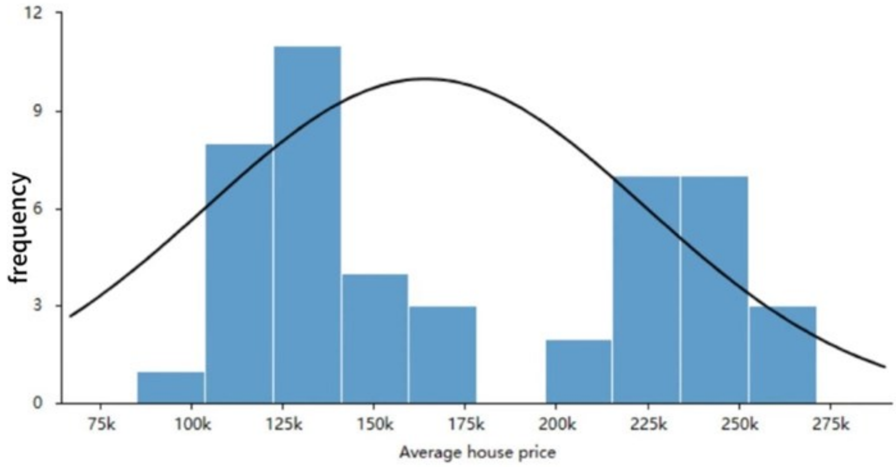


Figure 3. Frequency histogram of average house price.

As can be seen from Figure 3, the majority of UK house prices are between 125k and 137k, and the median of these data is around 162.8k. There are also many houses with prices between 214k and 252k, but most houses are still below 177k.

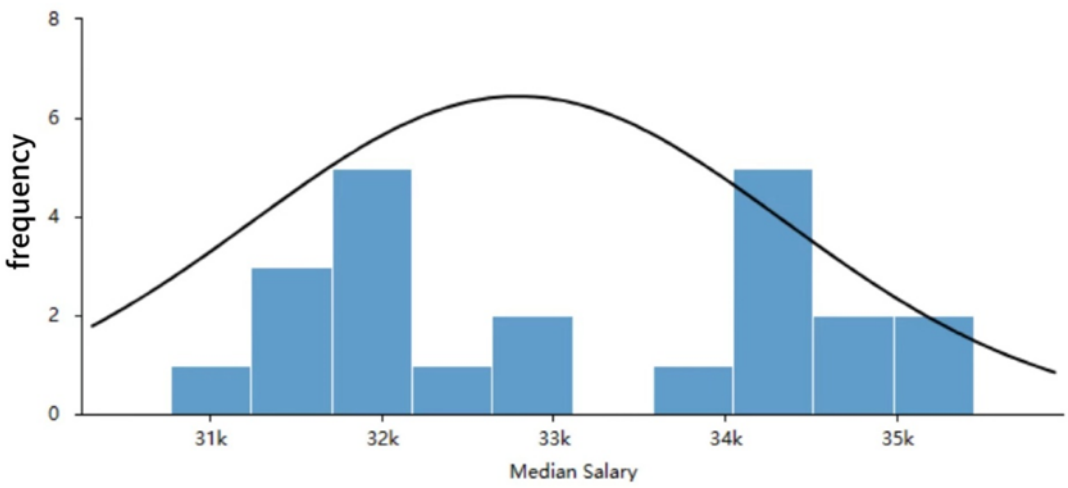


Figure 4. Frequency histogram of median salary.

Figure 4 shows the average salary of local people in the UK. It can be seen that the median per capita income is about 32.8k. In other words, a British citizen can afford an ordinary house with the salary he gets from working locally for 5 years.

3.2. Model results

Figure 5 shows the annual cumulative increase in housing prices. It can be seen that since 2019, housing prices have been on an upward trend, but suddenly fell in June 2020, but then surged. At the end of 2022, housing prices began to decrease, which also confirms the downward trend of housing prices at the end of Figure 1.

Figure 5. Cumulative house price increase by year.

From the data in Table 2, the figures show that x_7 and x_8 have the largest R values and are the main factors affecting housing prices, while x_6 also has a relatively large R value, but the other factors have less impact on housing prices, so loan, deposit and immigration rate are the main influencing factors.

DOI: 10.54254/2753-8818/52/2024CH0146

Table 2. Regression coefficient table.

	B	SE	Beta	T	significance	R
$x1$	-0.152	0.285	-0.152	-0.531	0.605	2.3%
$x2$	4.925	4.715	0.227	1.045	0.309	5.2%
$x5$	23.843	2.413	0.114	9.881	0.000	1.3%
$x6$	1977.371	1027.287	0.563	1.925	0.090	31.7%
$x7$	-7809.962	1785.435	-0.975	-4.374	0.143	95%
$x8$	-3312.253	695.290	-0.979	-4.764	0.132	95.8%

The data in Figure 6 shows that Wales has the largest increase among all regions in the UK, but it is also the region with the most severe decline in 2023. England is the region with the smallest increase in house prices since 2019. Although it has experienced a significant increase between 2021 and 2022, the growth rate is still the lowest among the four regions.

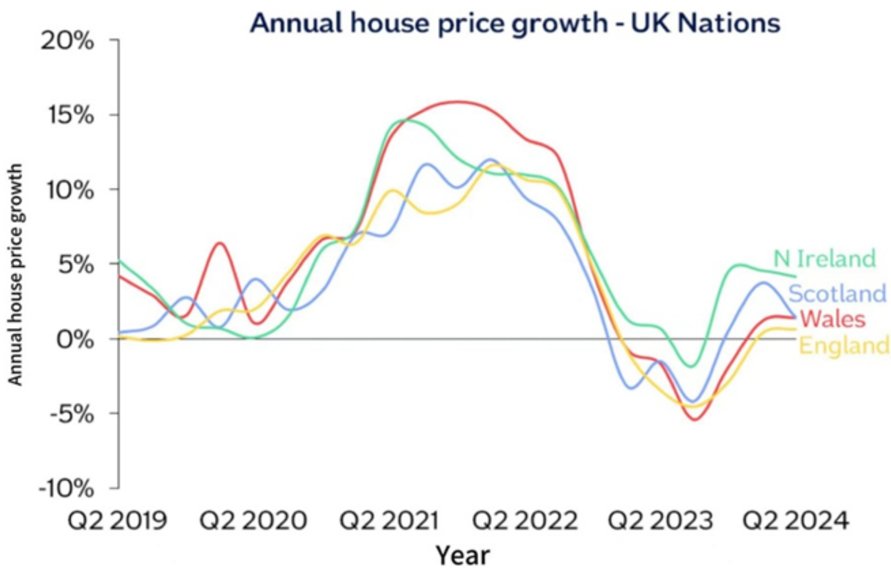


Figure 6. Annual house price growth in UK Nations.

The purpose of a multiple linear regression model is to construct a regression equation that uses multiple independent variables to estimate the dependent variable, thus explaining and predicting the value of the dependent variable. The dependent variables and most of the independent variables in the multiple linear regression model are quantitative values, and some qualitative indicators need to be converted to quantitative values before they can be applied to the regression equation. According to table 2, the author finds that the model can be written as:

$$y = -6451009.042 - 0.152x_1 + 4.952x_2 + 23.843x_5 + 1977.371x_6 - 7809.962x_7 - 3312.253x_8 \quad (1)$$

4. Conclusion

This article studies the UK house price data from 1952 to 2022, and the data contains 12 variables. From this study, it can be concluded that from 1952 to 2022, although the UK house price has fluctuated, it has mainly shown an upward trend in recent years. In addition, due to the impact of the epidemic in recent years, the global economy has been hit hard and is in a recovery period, so the UK house price will be more difficult to predict.

Through this study, people can consider their favorite houses from various angles and make rough house price predictions based on different factors. It is undeniable that given the limited number of samples in this article, the model may have errors, and there may be more different factors affecting

DOI: 10.54254/2753-8818/52/2024CH0146

the changes in house prices, which will affect the accuracy of house prices. But overall, it can still give the public different ideas to predict the development of house prices.

References

[1] Guirguis H, Mueller G, Dutra V, et al. 2024 Advances in Forecasting Home Prices. Comput. Econ.
[2] Chris H, John H and Bruce M 2018 Differing house price linkages across UK regions. Sage Publications, Ltd.
[3] Kuzlu M, Pipattanasomporn M and Rahman S 2015 Impacts of House Sizes, Appliance Ratings and Usage Patterns on Demand Response Applications: A Case-Based Study. Intelligent Industrial Systems.
[4] Luca E 2023 Corrigendum to Renewable energy consumption and per capita income: An empirical analysis in Finland. University of Salerno, 209, 558-568.
[5] Elias O, Steven C. Bourassa M H and Janne E 2023 Revisiting metropolitan house price-income relationships. Journal of Housing Economics.
[6] Robert C M, Beyer RC, Florian M, Claire L, Ezgi O O, and Lev R 2024 Monetary Policy Pass-Through to Interest Rates: Stylized Facts from 30 European Countries. IMF Working Papers.
[7] Wang Z Y, et al. 2023 Housing rental suggestion based on e-commerce data. Knowledge-Based Systems, 268.
[8] Inoue Y and Hashimoto M 2023 Significance of face-to-face service quality in last mile delivery for e-commerce platforms. Transportation Research Interdisciplinary Perspectives.
[9] Alam I A, Mansor H I and Ramzan Y 2023 Impact of House Price on Economic Stability: Some Lessons from OECD Countries. The Journal of Real Estate Finance and Economics.
[10] Mehmet B, Gizem U, Festus Victor B and Mark E W 2023 Housing price uncertainty and housing prices in the UK in a time-varying environment. Empirica.
[11] Omar H M N 2020 An Intra-City Analysis of House Price Convergence and Spatial Dependence. The Journal of Real Estate Finance and Economics.

DOI: 10.54254/2753-8818/52/2024CH0160

Research on application method of intelligent driving technology based on monocular vision sensor

Zeyu Zhang

School of Mechanical and Electrical Engineering, Chongqing Jiaotong University, Chongqing, China

200165@yzpc.edu.cn

Abstract. With the development of driverless cars, intelligent driving technology is increasingly used in the automotive industry, monocular vision sensor plays an indispensable role in intelligent driving technology because of its simple structure, low cost and abundant information. This paper discusses and optimizes the application of the monocular vision sensor in intelligent driving. The basic principle and key technologies of the monocular vision sensor are described in detail. In the specific application of the monocular vision sensor, this paper focuses on the monocular vision sensor's depth learning network, multi-information fusion technology, improved target detection and tracking algorithm. Through in-depth research and analysis, a series of optimization strategies based on the monocular vision sensor, such as the FAST Region Convolutional Neural Network (FAST–RCNN) vehicle target detection method and improved Scale-Invariant Feature Transform (SIFT) algorithm, are proposed. Finally, this paper summarizes the intelligent driving technology based on the monocular vision sensor and suggests that the monocular vision sensor will play a more important role in intelligent driving technology. Future research shall focus on improving the accuracy of the algorithm, such as the development of end-to-end convolutional neural network fusion methods, the proposed depth multi-modal sensor fusion network, and so on.

Keywords: Monocular vision sensor, intelligent driving, SLAM, deep learning network, multi-information fusion technology.

1. Introduction

Intelligent driving technology has become a research hotspot in the field of modern transportation. As an important part of intelligent driving systems, vision sensor plays a vital role. This paper focuses on the application of the single-binocular vision sensor in intelligent driving technology, aiming at discussing its working principle, application method, and optimization strategy, and analyzing its future development trend.

By capturing the image information of the surrounding environment, the vision sensor can provide abundant spatial and temporal data support for the intelligent driving system. Monocular vision sensor is widely used in intelligent driving because of its simple structure and low cost.

The applications of monocular vision sensors in intelligent driving include environment perception and target detection, ranging and positioning, multi-task fusion and optimization, etc.

Monocular vision is widely used in the environment perception of autonomous vehicles because of its simple structure and low computational cost. Monocular vision target recognition technology based on depth learning, such as Faster R-CNN and its improved version, can effectively improve the speed and accuracy of target recognition [1]. The simulation results show that the relative error is 2.71% in indoor environment and 3.81% in outdoor environment. Compared with the traditional algorithm, the proposed method can effectively improve the detection and ranging accuracy of the obstacles ahead, and its practicability and effectiveness are verified. This has important implications for the navigation control of autonomous vehicles under complex road conditions [2].

For the process of feature extraction and matching, a clustering obstacle detection method combining Scale-Invariant Feature Transform (SIFT) feature points is proposed. Firstly, the region of interest (ROI) of the image region is extracted, and the horizontal and vertical edge detection methods are used to identify the possible obstacles in the image region. Then, the following step is to calculate whether each ROI region contains obstacles, and remove the region that does not detect obstacles from the image. The remaining regions were used for the extraction of SIFT feature points. Then, K-means clustering is applied to these feature points to locate the obstacles accurately.

To reduce the influence of different size forward-looking targets on the depth of field estimation, a pyramid structure is adopted to preprocess the input image. During the follow-up training, the depth estimation problem was transformed into an image reconstruction problem, and a new loss function was designed, which could replace the true depth label [3].

Monocular vision simultaneous localization and map building (SLAM) provides a method to realize simultaneous localization and map building in an unknown environment. By relying only on monocular camera sensor data, this approach can achieve positioning and map construction. However, due to the relatively low computational complexity, monocular vision SLAM technology is widely used in small robots. Nevertheless, there are still some problems with this technique in terms of robustness and scale uncertainty [4].

Through the in-depth study of the application method of the monocular vision sensor in intelligent driving technology, this paper expects to further promote the development of intelligent driving technology, contribute to the construction of a safer, more intelligent and more efficient transportation system, and propose a series of optimization schemes. This paper not only sorts out the relevant theories in the field of intelligent driving but also provides a valuable reference for practical application.

2. Multitasking integration and optimization

2.1. Multi-tasking framework MGNet

MGNet, or Multiview Geometry Network, is a multi-task framework designed specifically for monocular Geometry scene understanding. It consists of two core tasks: panoramic segmentation and self-monitoring monocular depth estimation. In the task of panoramic segmentation, the model can segment the image and obtain the full semantic information and instance-level scene details, while in the aspect of self-supervised monocular depth estimation, the model automatically measures the monocular video depth using the constraints of the depth camera and the geometric environment [5].

2.2. Multi-sensor information fusion

In recent years, many researchers have begun to explore the combination of visual information and other sensor information to improve the accuracy and robustness of location. These fusion schemes usually use multiple sensors on the same hardware platform to collect data, to achieve higher positioning accuracy and better robustness. A fusion SLAM algorithm of monocular vision and inertia based on extended Kalman filter (EKF) and graph optimization (VISLAM) is proposed [4]. The algorithm improves the localization precision and the cost ratio of computation in three-dimensional space and realizes the localization with low cost and high precision. The EKF-based monocular vision and inertial fusion odometer are implemented to obtain motion estimation with less delay. The global

DOI: 10.54254/2753-8818/52/2024CH0160

map-aided EKF feedback mechanism is introduced, and the positioning accuracy is further improved by solving the system of linear equations.

3. Optimization strategy of monocular vision sensor in intelligent driving

In this chapter, the optimization strategy of depth learning framework, multi-information fusion technology, target detection and tracking algorithm and SLAM technology are presented, to improve the accuracy and robustness of the intelligent driving vehicle.

3.1. Optimization of the deep learning framework

Advanced depth learning framework such as depth residuals network is used to improve the accuracy of image recognition. These frameworks can effectively improve the performance of the model by increasing the depth of the network and introducing residual learning mechanism. At the same time, the use of very deep convolution networks can be considered, these networks improve the accuracy of recognition through the use of the small-scale convolution kernel.

3.2. Optimization of multi-information fusion technology

There is a certain risk of failure when using visual sensors to detect vehicles within the visual range. This risk is even greater in the non-line-of-sight condition. To solve this problem, it is suggested to construct a perceptual fusion framework based on deep learning [6]. The model combines complementary potential embedding with many advanced fusion strategies, and can effectively fuse image, radar, acoustic, seismic and other sensing modes. This model significantly surpasses the traditional single-peak detection method and greatly improves the accuracy of vehicle tracking and detection under non-line-of-sight conditions.

The validity of the model is verified on the multi-mode ESCAPE datasets. The experimental results show that the proposed fusion technique improves the vehicle detection performance by 33.16% on average compared with the visual detection method under the condition of 30-42% Natural Language Operating System (NLOS). On the more challenging multimodal NuScene datasets, the systems developed based on this model are on average 22% better than the competing approach. These results fully demonstrate the superiority and practicability of the model.

A functional model combining multi-vision information fusion is proposed. The feature fusion module is the core of the functional model, which combines the information of multiple visual sensors effectively, thus greatly improving the accuracy and robustness of detection and tracking [6]. The experimental results show that the proposed method is superior to the traditional single-vision sensor method in vehicle detection and tracking.

3.3. Optimization of target detection and tracking algorithm

An improved FAST Region Convolutional Neural Network (FAST-RCNN) vehicle target detection method based on binocular vision ranging is proposed. With the rapid development of intelligent driving technology, the demand for high-quality vehicle target detection algorithms is increasing. In this paper, a method of distance measurement based on binocular stereo vision is proposed to solve the limitation of traditional algorithms in speed and accuracy. Experimental results show that compared with FAST-RCNN, the proposed algorithm can significantly speed up the vehicle detection process and shorten the time by 42 ms. At the same time, the accuracy of the algorithm is only 2.4%, which shows high accuracy and good real-time performance. This innovative algorithm not only effectively improves the accuracy of vehicle detection, but also provides an advanced data processing platform for autonomous driving systems.

To improve the efficiency and reliability of the algorithm, an improved SIFT algorithm is proposed. Firstly, the feature descriptor is reduced from 128-dimension to 24-dimension, which makes the feature space closer to the image space. Then, the trilinear interpolation method is introduced to enhance the relationship between the feature points, thus significantly reducing the possibility of mismatching. Finally, a random sampling consistency algorithm is used to remove the mismatches.

DOI: 10.54254/2753-8818/52/2024CH0160

The experimental results show that compared with the original SIFT algorithm, the new algorithm improves the matching speed and accuracy significantly while keeping a higher number of feature points, especially, it performs well in the condition of angle change and illumination change, which can meet the requirement of real-time SLAM.

3.4. Real-time data processing and optimization

In order to improve the perception ability of intelligent vehicles to the vehicle information in the urban complex environment, a vehicle detection method combining radar and vision is proposed. By introducing the process of target tracking, the accuracy of vehicle position and velocity estimation is enhanced. An adaptive vehicle detection method based on target depth of field is designed, which takes into account the vehicle's driving speed, distance and visual angle, etc., the adaptive recognition of different types of vehicle information is realized. Finally, a target tracking method based on the combined framework of kernel correlation filter and extended Kalman filter (KCF-EKF) is proposed to track the vehicle accurately. The experimental results show that the proposed method shows good reliability and robustness in many traffic environments and weather conditions [7].

3.5. The latest improved SLAM algorithm based on monocular vision

In order to improve the accuracy and speed of feature matching, the traditional Oriented FAST and Rotated BRIEF (ORB) algorithm is improved, the scale space theory is introduced, and the possible range of feature points is predicted according to the prior information of robot motion, this method can effectively reduce the operation time and improve the matching accuracy [8].

To solve the problem of scale uncertainty in monocular visual SLAM, researchers begin to explore how to integrate other sensor data into monocular visual SLAM. For example, the methods of combining monocular images with wheel odometer data and inertial measurement unit (IMU) data can improve positioning accuracy and robustness to some extent.

4. Future trends in intelligent driving based on monocular vision

In this chapter, the development trend of monocular data forehead sensors in intelligent driving technology and the latest application methods, such as camera auto-exposure algorithm, end-to-end convolutional neural network fusion method, depth multi-mode sensor fusion network and hybrid multi-sensor fusion framework, are introduced.

4.1. Automatic camera exposure algorithm based on feature point detection

An innovative camera automatic exposure algorithm based on feature point detection is proposed, which aims at obtaining images with rich texture and clear details by adjusting exposure time [9]. In the current self-driving vehicles, most of the camera-based automatic exposure algorithms are only suitable for static scenes. However, in dynamic environments, these algorithms often fail to achieve the desired results. By detecting the feature points and analyzing their positions on the image, the camera shutter speed can be used to adjust the exposure parameters, to obtain clear and rich details of the image.

4.2. End-to-end convolutional neural network

In order to improve the precision of environment perception and depth estimation, adopts a binocular stereo matching algorithm to extract dense parallax information [9]. Because binocular vision system has significant uncertainty, and lidar system is relatively sparse, so it is necessary to combine the parallax images of the two systems for fusion processing. Based on this, a new end-to-end convolutional neural network is designed for the fusion of binocular parallax and lidar parallax. This method not only enhances the accuracy of environment perception, but also improves the accuracy of depth estimation.

DOI: 10.54254/2753-8818/52/2024CH0160

4.3. Depth multimodal sensor fusion network

A novel deep fusion network is proposed, which aims at robust fusion of sensor data under severe weather conditions without relying on a large amount of ground truth data [10]. The self-adaptive single-time model is used to measure entropy to drive feature fusion. After training the clean data set, the network model can demonstrate its effectiveness on a wide range of validation data sets. With this approach, it is expected that existing sensor data can be utilized, thereby significantly reducing the amount of marking required. This innovative framework is expected to be an important milestone in the development of data fusion technologies in intelligent transportation systems in the future.

4.4. Hybrid multi-sensor fusion framework

A novel hybrid multi-sensor fusion pipeline architecture is proposed to support autonomous vehicle environment perception in road segmentation, obstacle detection and tracking [11]. In this framework, a full-convolutional neural network (FCNX) and EKF are used to estimate the nonlinear state of the system. The goal is to create a more cost-effective, lightweight, modular, and robust fusion system solution that can maintain sensor performance even when it fails or underperforms. For road detection, the FCNX algorithm is used to improve accuracy. Compared with the traditional road detection based on a monocular vision sensor, the algorithm improves the detection accuracy and keeps the real-time efficiency on the embedded system. Test results on more than 3k road scenarios show that FCNX performs better in a variety of environment scenarios. In addition, the real-time performance of the algorithm is proved by experiments, and the real-time processing capability is further verified by the actual sensor data.

5. Conclusion

The application of the monocular vision sensor in intelligent driving technology is discussed in this paper. For an in-depth analysis, this research introduces the imaging principle, calibration method, ranging principle and application of vision sensor, and its combination with depth learning, and analyzes the key role and advantages of the vision sensor in intelligent driving.

The vision sensor provides a solid foundation for intelligent driving technology with its rich information acquisition ability and increasingly mature technology. Monocular vision sensor has a wide application potential in target detection and recognition, road scene understanding and so on. At the same time, the continuous progress of artificial intelligence technology provides strong support for the application of the vision sensor in intelligent driving. With the introduction of deep learning and machine learning, the vision sensor can identify the target accurately, estimate the depth and process the complex road scene in real time. The combination of these technologies not only improves the performance of intelligent driving system, but also opens up a new way for its future development. However, the vision sensor also faces some challenges and limitations in intelligent driving, such as low computing speed, non-dimensional invariance and data fusion with other sensors. Future research needs to explore these problems and seek effective solutions.

Looking into the future, with the development of technology and the increasing demand for intelligent driving, the application of the vision sensor in intelligent driving will be more extensive and in-depth. It is expected to see more innovative research and practice that will drive the development of intelligent driving technologies and bring more convenient and safe experiences for human travel.

In a word, the monocular vision sensor is an important part of intelligent driving technology, its development and application will provide strong support for the progress of the intelligent driving field. Through continuous exploration and innovation, it is believed that vision sensors will play an even more important role in the future development of intelligent driving technology.

References

[1] Chen Yupeng. Research on automatic driving monocular vision target recognition technology based on depth learning [D]. Jilin University, 2019.

[2] Xu Yang, Zhao Yanan, Gao Li et al.. Monocular vision-based vehicle detection and tracking [J]. Laser journal, 2020,41(05): 18-22.
[3] Ding Meng, Jiang Xinyan. Scene depth estimation method based on monocular vision in advanced driving assistance Système d' aide à la conduite, à l' exploitation et à la maintenance [J]. Journal of Optics, 2020,40(17): 137-145.
[4] Quan Meixiang. Research on SLAM algorithm based on multi-sensor information fusion [D]. Harbin Institute of Technology, 2021.
[5] Markus Schön, M. Buchholz et al. "MGNet: Monocular Geometric Scene Understanding for Autonomous Driving." IEEE International Conference on Computer Vision (2021). 15784-15795.
[6] Debashri Roy, Yuanyuan Li et al. "Multi-Modality Sensing and Data Fusion for Multi-Vehicle Detection." IEEE transactions on multimedia (2023). 2280-2295.
[7] Zhao Wangyu, Li Bijun, Shan yunxiao, etc.. Hybrid millimeter wave radar and monocular vision for front vehicle detection and tracking [J]. Journal of Wuhan University (Information Science Edition) , 2019,44(12) : 1832-1840.
[8] Hu Wei, Liu Xingyu. An improved SIFT algorithm for image matching in unidirectional SLAM [J]. Electro-optic and control, 2019,26(05): 7-13.
[9] Chen Kun. Road environment sensing technology based on binocular vision and lidar fusion [D]. Zhejiang University, 2020.
[10] Mario Bijelic, Tobias Gruber et al. "Seeing Through Fog Without Seeing Fog: Deep Multimodal Sensor Fusion in Unseen Adverse Weather." Computer Vision and Pattern Recognition (2019). 11679-11689.
[11] Babak Shahian Jahromi, Theja Tulabandhula et al. "Real-Time Hybrid Multi-Sensor Fusion Framework for Perception in Autonomous Vehicles." Italian National Conference on Sensors (2019).

Research on regional difference of economic development level in Zhejiang Province based on factor analysis

Yuao Feng[1], Ji Zhang[2,3]

[1]Tianfu Middle School, Chengdu, 610000, China
[2]School of Mathematics and Statistics, Huaibei Normal University, Huaibei, 235025, China

[3]20150725250@mail.sdufe.edu.cn

Abstract. Regional economic disparities are widespread, especially in China. This is reflected not only in the differences between the four regions, but also in the differences between provinces. At the same time, the regional economic differences within each province are also very obvious. This paper studies the gap between the economic development levels in Zhejiang Province, which contributes to the balanced development and sustainable development of the overall economy of Zhejiang Province, in order to intuitively see the gap between the cities. This paper uses SPSS to reduce the dimension of factor analysis and calculate the comprehensive score through 10 indicators reflecting the level of economic development in 11 cities. Finally, this paper finds that the economic development of each city in Zhejiang Province has great potential. But at the same time, it exposes the imbalance of economic development in the province. The economic development level of Hangzhou, Ningbo and Zhoushan is far higher than that of other cities in the province, and the economic development level of some cities is general. The development level of some cities is poor. According to the research results, the causes and problems of the current situation of economic development are analyzed. Reasonable suggestions are put forward to further promote the coordinated development of the economy.

Keywords: Economic development, factor analysis, Zhejiang province.

1. Introduction

As a major economic province, Zhejiang Province ranked fourth in the country in 2023, and its total economic output expanded steadily. The total economic output exceeded 8 trillion yuan. However, there are always economic differences among the cities in the province. The problem of regional economic differences refers to the per capita meaning between regions in a certain period of time under the influence of population, social resources and political factors. The overall level of economic development is not equal, the speed of development among regions is not equal, and some regions are more developed. Developed regions and relatively backward regions coexist [1].

For the study of the current situation of economic development and regional differences, Myrdal put forward the theory of circular accumulation causality in 1957.The elimination of "geographically dual economy" is explained by circular cumulative causality. Under the mutual influence of backflow effect and diffusion effect, a cumulative cycle development trend is formed. Albert Hirschman put forward the theory of unbalanced growth in 1985 through the recognition that major scarce resources should be fully utilized. It is believed that developing countries should invest selectively in the sector. Perroux proposed the growth pole theory in 1950 [2]. He believed that in real life, only the rapid growth of individual regional economy can drive the overall economic growth, and balanced growth only exists in the ideal [2].

At present, more and more scholars in China have invested in this kind of research. Li Yue used factor analysis and other research methods to study Guangdong. The economic differences between provinces and cities have been systematically studied [3]. Zhao studied the regional economic differences in the western region and measured the relative differences through Theil index and Gini coefficient. The Theil index from 1978 to 2021 is decomposed in one stage, and the results of intra-regional and inter-regional differences in the western region are obtained. It is found that the growth rate of per capita GDP in Southwest China is basically higher than that in Northwest China in the past ten years [4]. Wang used the entropy weight TOPSIS method to comprehensively evaluate and analyze the regional economic development of Guangxi. This paper explores the main factors affecting the differences of regional economic development in Guangxi. It is concluded that the level of Guangxi's total economic output is stable and progressive, but the economic growth rate is relatively low, and the per capita economic output is small. The industrial structure is unreasonable and the urbanization process is slow [5]. Fang selected 18 prefecture-level cities in Henan Province in 2020 as research samples. On the basis of literature review, this paper selects 14 indicators that directly reflect the development level of logistics industry. Through the method of factor analysis, the logistics development level of each prefecture-level city is analyzed, and the factor scores of each city are ranked [6]. Peng and Liu found that with the passage of time, the gap of economic development level among provinces in China has gradually widened [7]. In 2021, Liu analyzed the diversification of market economy development in rural areas. And put forward in the government policy guidance, strengthen the leading effect of investment projects, promote agricultural science and technology innovation, reasonable adjustment of rural industrial structure. Effective measures have been taken to alleviate the gap in the development of rural regional market economy. It effectively promotes the harmonious development of economy and society in rural areas [8]. When Yang studied the changing trend of regional economic disparities from 1952 to 1990, it is found that: China's economic development difference shows an inverted "U" change around 1978.Moreover, the economic development differences of different regional division systems have different changing trends [9]. According to the research of Liu and others in 2009, the regional disparity from 1978 to 2007 was analyzed by applying the Theil index. The results show that due to the unbalanced development within the region, the overall development differences in China are gradually emerging [10]. Although the contribution of the eastern region is outstanding, the contribution of the western region is small but shows an expanding trend. In addition, the study also classified China according to the latitude of north and south.

In a word, this paper is committed to in-depth analysis of the main factors affecting regional economic disparities in Zhejiang Province. And to explore the current situation of regional economic development in Zhejiang Province, so as to promote the balanced and coordinated development of the economy in Zhejiang Province.

2. Methodology

2.1. Data source

In this paper, 11 cities in Zhejiang Province are taken as the research objects. The data are selected from 10 indicators reflecting the level of economic development in the Statistical Yearbook of Zhejiang Province in 2023.The data are processed as follows: for some indicators of per capita development level, they are not given in the Statistical Yearbook. It can be obtained by dividing the total development level index by the resident population of each city at the end of the year, and there is no missing value in the data. Missing value processing is not required.

DOI: 10.54254/2753-8818/52/2024CH0124

2.2. Variable description

In this paper, the problems studied are classified into three levels from large to small, namely, the first level is the level of economic development, and the second level is the level of economic development. The second-level indicators are the total economic output and per capita economic level, while the third-level indicators are the 10 indicators this paper has selected (Table 1). Finally, the factor analysis process was carried out with 10 indicators in the three-level indicators.

Table 1. Variable description.

The first indicator	Second indicator	Third indicator
Economy Development Level	Total economic output	The scale is based on the total profit of public enterprises (100 million yuan) X2, the total fiscal revenue (100 million yuan) X6, and the total output value (100 million yuan) X10.
	Per capita economic level	Per capita GDP (yuan) X1, per capita total fiscal revenue (yuan) X4, per capita output value of agriculture, forestry, animal husbandry and fishery (yuan) X5, per capita general budget revenue of local finance (yuan) X7, per capita total import and export volume of customs (dollars) X8, per capita total retail sales of social consumer goods (yuan) X9, per capita disposable income of urban residents (yuan) X3

2.3. Method introduction

Factor analysis is a statistical method used to describe the relationship between variables in observed data. It also explains the latent factors (or latent variables, hidden variables), which can explain the correlation between variables. The goal of factor analysis is to explain most of the variability in the data by a few factors. It is usually used in psychology, sociology, market research and other fields to identify potential structures. Factor analysis was divided into Exploratory Factor Analysis, EFA and Confirmatory Factor Analysis, CFA. The EFA is used to explore the underlying structure of the data without a preset model, while the CFA is used to validate the preset model.

Factor analysis is generally divided into six steps, the first is data preparation, the collection of raw data and standardized data. The second is to select the number of factors, which can be judged by Kaiser criterion, and the third is to extract factors. Factor loading matrix is often obtained by principal component analysis, which can describe the loading of each variable on each factor. The fourth is factor rotation. It is used to simplify the explanation of factors and make the factor structure clearer. The commonly used rotation methods are orthogonal rotation and oblique rotation. The fifth is factor interpretation, which explains the factor loading matrix. Identify which variables each factor is primarily composed of, and thus name the factors. Finally, the factor score is calculated, and the factor score and the comprehensive score of each observation sample are calculated according to the factor analysis model.

3. Results and discussion

3.1. Descriptive statistics

Descriptive statistics of raw data can provide a basic understanding of the data set. Including the central trend and the degree of dispersion of the data, it can also help people get better standardized data. Lay a good foundation for the following research (Table 2).

DOI: 10.54254/2753-8818/52/2024CH0124

Table 2. Description analysis.

	Min	Max	Average	SD
Per capita GDP (yuan)	72812	167134	115270.18	35482.796
Total profit of the enterprise whose scale is based on public works (yuan)	85.67	1543.04	543.0927	498.22327
Per capita disposable income of urban residents (yuan)	55784	77043	70179.55	7160.611
Total fiscal revenue per capita (yuan)	9475	37089	19553.36	10829.834
Per capita output value of agriculture, forestry, animal husbandry and fishery (yuan)	2351	26113	7429.36	6446.501
Total fiscal revenue (100 million yuan)	277.88	4590.08	1260.3364	1391.60301
Per capita general budget revenue of local finance (yuan)	5929	19801	10587.73	4651.442
Total Customs Imports and Exports Per Capita (USD)	12906	289041	80498.27	77083.256
Total retail sales of consumer goods per capita (yuan)	32002	58933	43919.64	7732.553
Gross output value (100 million yuan)	1831	18753	7074.09	5524.269

From the preliminary analysis of the maximum value, minimum value and standard deviation in Table 2, this paper can see that there is a big gap between the economic development level indicators of each city. However, it is impossible to know the overall level of economic development of each city, so this paper proceeds to the next step of analysis.

Before factor analysis, the original data should be standardized. The purpose of standardization is to unify the units of variables involved in the actual problems studied. At the same time, it also paves the way for the calculation of the following factor scores.

3.2. KMO test

Using standardized data and SPSS software, the correlation matrix and KMO and Bartlett test were calculated (Table 3, 4). After observing the correlation matrix table, it is found that the correlation between most variables is above 0.5. It shows that the correlation between variables is relatively strong, and the preliminary judgment can be carried out by factor analysis, and then observe the KMO and Bartlett test table. It can be known that the KMO value is greater than 0.5, and the significance is less than 0.001, indicating that the selected indicators are completely suitable for factor analysis (Table 3).

Table 3. KMO and Bartlett test.

KMO sampling appropriateness measure	0.54
Approximate chi-square	169.169
Degree of freedom of Bartlett sphericity test	45
Significance	<.001

Table 4. Correlation results.

	x1	X2	X3	X4	X5	X6	X7	X8	X9	X10
X1	1	0.597	0.661	0.915	0.472	0.573	0.876	0.686	0.866	0.518
X2	0.597	1	0.627	0.658	-0.339	0.964	0.822	-0.013	0.724	0.968
X3	0.661	0.627	1	0.521	0	0.571	0.572	0.352	0.722	0.679
X4	0.915	0.658	0.521	1	0.432	0.712	0.953	0.641	0.859	0.588
X5	0.472	-0.339	0	0.432	1	-0.261	0.169	0.841	0.192	-0.375
X6	0.573	0.964	0.571	0.712	-0.261	1	0.847	0.037	0.75	0.966
X7	0.876	0.822	0.572	0.953	0.169	0.847	1	0.414	0.915	0.744
X8	0.686	-0.013	0.352	0.641	0.841	0.037	0.414	1	0.45	-0.036
X9	0.866	0.724	0.722	0.859	0.192	0.75	0.915	0.45	1	0.687
X10	0.518	0.968	0.679	0.588	-0.375	0.966	0.744	-0.036	0.687	1

3.3. Model results

The common factor variance (Table 5) represents the portion of the variance of each observed variable that can be explained by the common factor. This value reflects the dependence of each variable on the common factor, and the larger the value. It shows that more variance of the variable can be explained by the common factor. Through observation, this paper can see that the common degree of variable (per capita disposable income of urban residents) X3 is 56.5%. The information extracted from the rest of the variables is more than 80%, indicating that the extracted common factors are sufficient to explain most of the variables.

Table 5. Common Factor Variance.

	Initial	Extract
X1	1	0.937
X2	1	0.971
X3	1	0.565
X4	1	0.933
X5	1	0.947
X6	1	0.942
X7	1	0.931
X8	1	0.905
X9	1	0.884
X10	1	0.959

The selected common factor generally requires the characteristic value to be greater than 1. There are two factors whose eigenvalues are greater than 1, so it is preliminarily considered that there are two common factors. By observing the total variance interpretation table (Table 6), the contribution rate of the first factor is 59.368%, and the contribution rate of the second factor is 30.365%. Therefore, the cumulative contribution rate of these two common factors is 89.732% (> 85%). Most of the information of the 10 indicators that can reflect the level of economic development can be seen from the rotated load matrix table (Table 7).

The first public factor represents the per capita GDP X1, the total profit of industrial enterprises X2, and the per capita disposable income of urban residents X3.Per capita total fiscal revenue X4, total fiscal revenue X6, per capita local general budget revenue x7, per capita total retail sales of social consumer goods X9,Gross output value X10,Most of these indicators reflect the economic productivity, income level of residents, consumption capacity and government financial situation in the region. It represents

DOI: 10.54254/2753-8818/52/2024CH0124

the health and vitality of the overall economy. Therefore, this factor can be interpreted as "regional economic development and prosperity factor".

The second common factor represents the per capita output value of agriculture, forestry, animal husbandry and fishery X5 and the per capita total import and export volume of customs X8. They reflect the level of agricultural development and the strength of external economic ties in the region. It represents the breadth and depth of the region's economic activity in both domestic and international markets. It can be interpreted as "economic diversity and external economic openness factor".

Table 6. Explanation of Total Variance.

Ingredient	Initial eigenvalue			Extract the sum of the load squares		
	Total	Percent variance	Cumulative%	Total	Percent variance	Cumulative%
1	6.402	64.022	64.022	6.402	64.022	64.022
2	2.571	25.711	89.732	2.571	25.711	89.732
3	0.596	5.955	95.687			
4	0.207	2.072	97.759			
5	0.102	1.017	98.776			
6	0.08	0.799	99.575			
7	0.032	0.323	99.898			
8	0.01	0.097	99.994			
9	0	0.004	99.999			
10	0	0.001	100			

Table 7. Rotated Component Matrix.

	Ingredient	
	1	2
X1 (GDP per capita)	0.698	0.671
X2 (the scale is based on the total profit of industrial enterprises)	0.975	-0.143
X3 (per capita disposable income of urban residents)	0.722	0.21
X4 (total fiscal revenue per capita)	0.75	0.609
X5 (per capita output value of agriculture, forestry, animal husbandry and fishery)	-0.216	0.949
X6 (total fiscal revenue)	0.966	-0.09
X7 (per capita general budget revenue of local finance)	0.888	0.376
X8 (total customs import and export per capita)	0.119	0.944
X9 (per capita total retail sales of consumer goods)	0.844	0.414
X10 (GDP)	0.959	-0.198

The resulting two common factors are scored to obtain a component score coefficient matrix for each variable (Table 8).

DOI: 10.54254/2753-8818/52/2024CH0124

Table 8. Component Score Coefficient Matrix.

	Ingredient	
	w1	w2
X1 (GDP per capita)	0.078	0.189
X2 (the scale is based on the total profit of industrial enterprises)	0.191	-0.126
X3 (per capita disposable income of urban residents)	0.117	0.021
X4 (total fiscal revenue per capita)	0.092	0.163
X5 (per capita output value of agriculture, forestry, animal husbandry and fishery)	-0.112	0.359
X6 (total fiscal revenue)	0.185	-0.106
X7 (per capita general budget revenue of local finance)	0.135	0.068
X8 (total customs import and export per capita)	-0.05	0.331
X9 (per capita total retail sales of consumer goods)	0.124	0.085
X10 (GDP)	0.192	-0.144

The data from Table 8 is combined with the expression for the principal component:

$$F1 = 0.078ZX1 + 0.191ZX2 + 0.117ZX3 + 0.092ZX4 - 0.112ZX5 + 0.185ZX6 + 0.135ZX7 - 0.05ZX8 + 0.124ZX9 + 0.192ZX10$$

$$F2 = 0.189ZX1 - 0.126ZX2 + 0.021ZX3 + 0.163ZX4 + 0.359ZX5 - 0.106ZX6 + 0.068ZX7 + 0.331ZX8 + 0.085ZX9 - 0.144ZX10$$

Where ZX1 to ZX10 are the values in the normalization of the original data of the table, and they are substituted into the above equation. The score table of each factor can be obtained (Table 9).

Table 9. Scores of each factor.

FAC1_1	FAC2_1
2.15293	-0.29334
1.55949	0.2287
-0.32485	-0.91866
0.08777	-0.1866
-0.14713	0.02611
0.14068	0.04694
-0.47369	-0.40896
-0.90913	-0.38434
-0.48847	2.85885
-0.47281	-0.37758

After getting the score of each common factor, this paper can calculate the comprehensive score of the level of urban economic development. The comprehensive score of factor analysis is the sum of the scores of each common factor multiplied by their respective contribution rates. The comprehensive score formula in the factor analysis of this paper is:

$$F = F_1 \times 0.59 + F_2 \times 0.3 \tag{1}$$

Finally, by ranking the comprehensive scores, this paper can intuitively see the economic development level of each city (Table 10). According to Table 9, the city with the highest comprehensive score is Hangzhou, and the city with the lowest comprehensive score is Lishui.

DOI: 10.54254/2753-8818/52/2024CH0124

Table 10. Comprehensive Scores and Ranking of Economic Development Level of Cities in Zhejiang Province in 2022.

City	F1	F2	F	F1 rank	F2 rank	F Rank
Hangzhou	2.15293	-0.29334	1.18	1	6	1
Ningbo City	1.55949	0.2287	0.99	2	2	2
Zhoushan City	-0.48847	2.85885	0.57	9	1	3
Shaoxing City	0.14068	0.04694	0.1	3	3	4
Jiaxing City	0.08777	-0.1866	0	4	5	5
Huzhou City	-0.14713	0.02611	-0.08	5	4	6
Taizhou City	-0.47281	-0.37758	-0.39	7	7	7
Jinhua City	-0.47369	-0.40896	-0.4	8	9	8
Wenzhou City	-0.32485	-0.91866	-0.47	6	11	9
Quzhou City	-0.90913	-0.38434	-0.65	10	8	10
Lishui City	-1.1248	-0.59112	-0.84	11	10	11

3.4. Discussion

According to the comprehensive score, this paper divides the level of economic development into three echelons, the first echelon is the city with a comprehensive score of more than 0.5. Among them are Hangzhou, Ningbo and Zhoushan, which shows that the economic development level of these cities is better. The second echelon is a city with a comprehensive score between 0.5 and 0.5, including Shaoxing, Jiaxing, Huzhou and Taizhou. Jinhua City and Wenzhou City show that the economic development level of these cities is general, and the third echelon is the city with a comprehensive score below -0.5. Among them, Quzhou City and Lishui City show that the economic development level of these cities is poor.

Through the ranking table of two common factors, this paper finds that in the first common factor, Hangzhou, Ningbo, Shaoxing, Jiaxing ranks in the top four, indicating that these four cities have a high level of overall economic health and vitality. This paper argues that these cities are located in the Yangtze River Delta region along the southeastern coast of China, close to Shanghai, and have convenient transportation networks. This advantageous geographical location has promoted economic exchanges and trade within and outside the region. At the same time, the private economy in these cities is very active, entrepreneurship is strong, and many innovative enterprises and industrial clusters have emerged. It provides impetus for the sustained and healthy development of the economy. In the second public factor, Zhoushan, Ningbo, Shaoxing and Huzhou are the top four cities. It shows that the level of agricultural development and the intensity of foreign economic ties in these cities are relatively high. This paper argues that due to geographical advantages, abundant natural resources, policy support, agricultural modernization and superior transportation and logistics conditions, etc. The result of the interaction of heavy factors.

Comparing the ranking of the two public factors, this paper finds that Zhoushan ranks the bottom in the overall economy. However, the level of agricultural development and foreign economy is relatively high. It shows that Zhoushan's advantages in agriculture and foreign economy have not been fully transformed into the health and vitality of the overall economy. This suggests that in the future development, Zhoushan needs to pay attention to the diversification of economic structure and enhance the integrity of the industrial chain. Increase infrastructure construction and service level, enhance domestic demand market, and enhance support for high-end talents and innovation. Optimize government policies and support to achieve coordinated, healthy and sustainable development of the overall economy.

4. Conclusion

In general, by studying the regional difference of the economic development level of the cities in Zhejiang province, this paper can intuitively see the differences in the development of each city. This

DOI: 10.54254/2753-8818/52/2024CH0124

paper roughly divides the level of economic development into three echelons, which shows that the differences between regions are still obvious. Therefore, this paper puts forward the following suggestions:

First of all, the author needs to promote industrial transfer and upgrading, and encourage the transfer of high-end industries and technologies from developed areas to relatively backward areas. Promote regional integration of the industrial chain. Industrial cooperation zones can be established to promote cross-regional cooperation and share technology, resources and markets.

Second, it is necessary to optimize infrastructure construction and strengthen the construction of provincial transportation network. In particular, highways, railways and ports connecting relatively backward areas will enhance the level of regional connectivity. Facilitate the flow of logistics and personnel. At the same time, this paper should speed up the construction of digital infrastructure, improve network coverage and informatization level, and promote the development of smart cities and smart countryside.

Third, this paper will increase financial support and targeted poverty alleviation, and increase financial transfer payments to underdeveloped areas. This paper will continue to promote targeted poverty alleviation and ensure that the poor do not return to poverty. Through comprehensive measures such as industrial poverty alleviation, education poverty alleviation and health poverty alleviation, the self-development ability of poor areas and poor people will be enhanced.

Authors contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Li X H 2013 Research on the Regional Difference of Retail Business Performance and Its Influencing Factors. Chengdu: Southwestern University of Finance and Economics.
[2] Wu D T 2009 Regional Economics. Beijing: Science Press, 261-262.
[3] Li Y 2012 Research on Urban Economic Disparity in Guangdong Province. Chengdu: University of Electronic Science and Technology of China.
[4] Zhao C C 2023 Study on the Spatio-temporal Evolution and Influencing Factors of Regional Economic Disparity in Western China. Yan'an University.
[5] Wang D Q 2023 Spatio-temporal Evolution Analysis of Regional Economic Development Differences in Guangxi. Guangxi Normal University.
[6] Fang M Y 2021 Research on Regional Logistics Development in Henan Province Based on Factor Analysis and Cluster Analysis. Zhengzhou University.
[7] Peng W B and Liu Y J 2010 Spatio-temporal Evolution Characteristics of Regional Economic Disparity in Eastern, Central and Western China. Economic Geography, 574-578.
[8] Liu S Q 2021 Analysis and Countermeasures of Rural Regional Economic Development Differences. Shanxi Agricultural Economy, 52-53.
[9] Yang K Z 1994 Study on the Change of Regional Economic Disparity in China. Economic Research, 28-33.
[10] Liu J, et al. 2009 Historical Investigation and Empirical Study of Regional Disparity in China. Economic System Reform, 128-132.

DOI: 10.54254/2753-8818/52/2024CH0163

Analysis of customer purchasing behavior in E-commerce model

Wenxuan Ding

Guangdong Experiment High School, Guangzhou, 510000, China

kouyan@ldy.edu.rs

Abstract. Because of globalization and information technology (IT), e-commerce (EC) has changed business practices and had an impact on established market systems. Global EC sales are still increasing, which emphasizes how crucial it is to comprehend consumer purchasing behavior. Using a Kaggle dataset named "Exploring E-commerce Trends," which contains product data including price, rating, stock amount, and sales, this paper examines consumer purchasing behavior. The paper used Multiple Linear Regression and the Random Forest methods. The quantity of stock, price, and the number of reviews were found to be important predictors of sales using the Random Forest ensemble learning technique. However, the model exhibited overfitting, performing well on the training set but poorly on the testing set. Multiple Linear Regression results showed minimal explanatory power for sales, indicating limitations in the model's effectiveness. While the Random Forest model identified key factors influencing sales, its overfitting and the ineffectiveness of the Multiple Linear Regression model suggest the need for more robust methods to predict customer purchasing behavior in e-commerce accurately. Future research should integrate additional variables and employ more sophisticated models to enhance prediction accuracy and business decision-making.

Keywords: E-commerce, data analysis, purchasing behavior.

1. Introduction

E-commerce (EC), short for electronic commerce, known to the activity of purchasing or selling products online service or over the Internet. Jain states that business methods are transformed because of Information Technology (IT) and globalization, and EC takes a significant role in the electronic business, affecting the traditional market system [1]. Over the past few decades, sales on the global EC market have continued increasing. This tendency emphasizes how important it is for businesses to interact with customers on digital platform, as well as how consumer buying habits are changing. EC provides enormous market opportunities and advantages, but also new challenges, especially in analyzing and predicting [2].

The existing researches indicates that customer purchasing behavior is influenced by multiple factors. Kumar concluded that the following five elements have been found to be crucial to an e-commerce store's success: net benefit, trust, information quality, system quality, and service quality [3]. Besides, price sensitivity and promotional strategies contribute to consumers' decisions, while Hansen and Tambo showed that brand and product channels also affect e-commerce from the perspective of information system [4-6]. Social media is regarded as the major way of delivering information from

DOI: 10.54254/2753-8818/52/2024CH0163

sellers to customers and connecting the global community. Another definition of social media is a collection of web-based tools that expand on the technical underpinnings that facilitate the sharing and production of user-generated content [7]. Furthermore, individualized suggestion systems are essential for raising conversion rates and consumer satisfaction. Raji pointed out that Businesses must comprehend and take advantage of AI-driven tactics as they navigate the digital landscape in order to remain competitive and satisfy the changing demands of tech-savvy customers [8]. The study by Liao and Sundar investigated consumer preference by following the heuristic-systematic model in social psychology. The results indicated a tendency that people to prefer content-based filtering with higher matches to seek for cognition for the experience product, while collaborative filtering increases favorable reviews for search items by inducing the "bandwagon effect" [9].

The development of EC pointed out the problem of "information overload", and then discussed the personalized service content and personalized service mode to prove the relevance of the factors that affect customers' shopping behavior, but Liu also mentioned that the flexibility of the recommendation system is insufficient [10]. However, although these studies provide valuable insights into understanding customers' purchasing behavior, most studies do not cover a comprehensive range of variables or lack further research on multiple influences. For this reason, comprehensive insights into how these factors collectively impact purchasing decisions are sparse, so it is necessary to analyze it in a more integrated approach.

This study aims to address this gap by using the dataset available on Kaggle, named the "E-commerce dataset". For the businesses, the capacity of tracking customer consumption behavior over an extended period of time is helpful in obtaining specific measures for future product trends, customer preferences, and also the market rivalry.

In summary, E-commerce analysis and customer behavior have attracted numerous scholars, and this topic deserves in-depth study and further research. This article aims to analyze customer behavior based on the dataset about e-commerce trends to visualize the data using statistical models.

2. Methodology

2.1. Data source

The dataset used in this paper is obtained from the Kaggle website (Exploring E-commerce Trends). The dataset contains information on 1000 products across different categories ranging from electronics to commodities. A product's price, rating, quantity of stock, number of reviews, discounts, sales, and date of inventory addition are all related with it. The dataset was collected by Muhammad Roshan Riaz. The original dataset preserved in .csv format.

2.2. Variable selection

The original dataset included the variables of product ID, category, price, rating, number of reviews, stock quantity, discount, sales, and the date added the inventories (Table 1).

Table 1. List of variables

Variable	Logogram	Meaning
Product Name	x_1	The name of the product
Category	x_2	The type of the product
Price	x_3	The cost of the item
Rating	x_4	Grade according to quality
Number of reviews	x_5	Number of reviews consumer writing
Quantity of stock	x_6	Number of inventories
Discount	x_7	A deduction from the cost of the product
Sales	Y	Number of products that customers purchasing

DOI: 10.54254/2753-8818/52/2024CH0163

The product ID is not worthwhile in this study, so the list of product IDs would not be included in the research. Since the dataset is clean, without any null variables, all of the 1000 groups of data will be contained in the study. Eventually, the analysis contains 1000 groups of data. The data includes 7 variables: Product Name, category, price, rating, number of reviews, quantity of stock, and discount. The one dependent variable is sales. The specific description of those variables is demonstrated in Table 1.

2.3. Methods introduction
The paper uses the random forest model and multiple linear regression mainly. An approach to group learning is called random forest. In order to increase prediction accuracy and stability, it builds several decision trees during training and combines the outcomes. Each tree is generated by random sampling of the original data and at each split, only a random subset of features is considered. This approach helps to reduce overfitting and enhances the generalization capability of the model. Moreover, this model can handle non-linear relationships between features effectively, and also tolerate missing values better. It automatically assesses the importance of a feature, helping to identify which variables are most significant for prediction. A linear regression model containing several explanatory variables is called a multiple linear regression model. It serves as an explanation for the linear relationship that exists between the variable being explained and several additional explanatory factors. Based on the result of random forest, the paper would further explore the effects by multiple linear regression with that feature.

3. Results and discussion

3.1. Random forest analysis
The graph below lists the variety of products category that contained in the dataset.

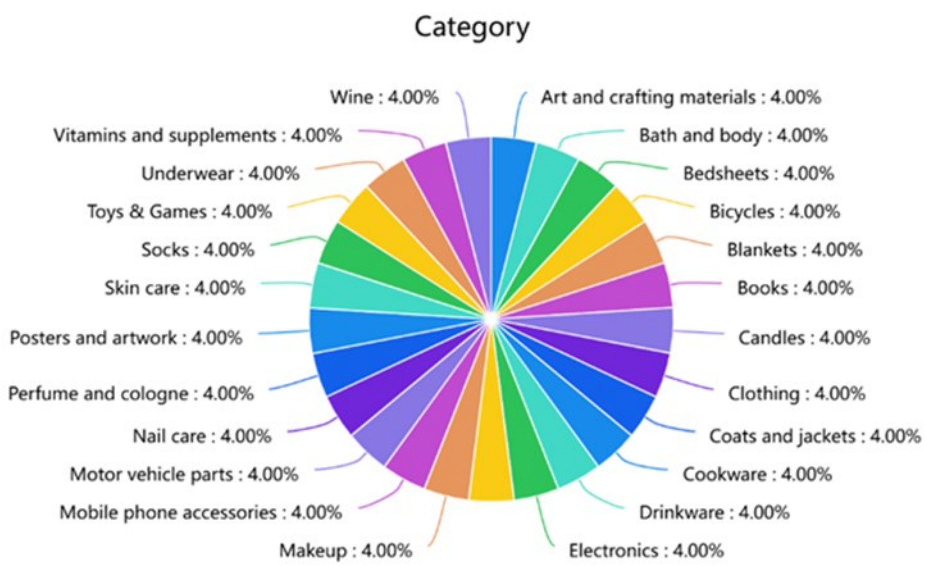


Figure 1. The Category of the dataset.

According to Figure 1, this dataset includes 25 different categories of products, ranging from electrics to blankets. Those products were mostly common in people's daily lives, which means the probability of purchasing those products is high. Therefore, the dataset provides the author the value to analysis the important factor that affects consuming. Each category contains 40 products.

Figure 2 run by the random forest model shows how much weight each variable currently holds.

DOI: 10.54254/2753-8818/52/2024CH0163

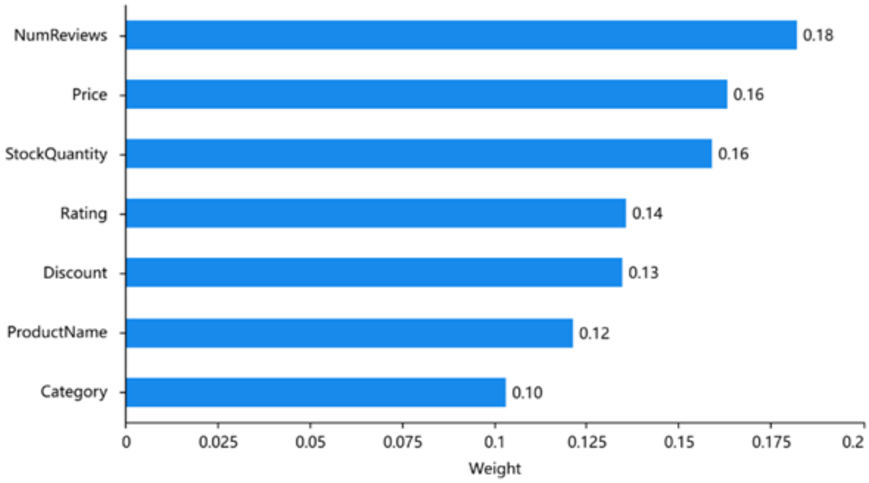


Figure 2. The varibles' weight.

Figure 2 reveals the feature weight shows the importance of each title in the model, and its sum is 1. Base on this graph, the author found that the variable of x5 (NumReviews) is the most significant variable that affect the sales, accounted for 18.23%. Then, Price accounts for 16.33%. StockQuantity accounts for 15.93%. Rating accounts for 13.57%. Discount accounts for 13.47%. The above five features account for 77.54% in total.

Table 2. Model evaluation results

Index	Training set	Testing set
R^2	0.854	-0.083
Mean Absolute Error (MAE)	190.669	508.060
Mean Square Error (MSE)	50165.560	348978.882
Root Mean Square Value (RMSE)	223.977	590.744
Median Absolute Deviation (MAD)	185.735	485.595
Mean Absolute Percentage Error (MAPE)	null	3.116
Explained Variance Score (EVS)	0.854	-0.080
Mean Squared Log Error (MSLE)	0.535	0.801

From Table 2, the random forest model performs well on the training set, exhibiting high explanatory power and low errors. However, its performance on the testing set is poor, with a negative R^2 value and significantly higher mean squared error, indicating overfitting and a failure to generalize to the new data.

3.2. Multiple linear regression

Base on the findings in random forest, this paper can learn that NumReviews, price, StockQuantity were the key factors on sales. For this reason, continuing using the multiple linear regression according to these three points (Table 3).

The aforementioned table illustrates how the linear regression analysis is carried out, using Sales as the dependent variable and NumReviews, Price, and StockQuantity as independent variables. As can be seen from the above table, the model formula is as follows: Sales=925.620 + 0.022*NumReviews + 0.117*Price-0.000*StockQuantity. The model R-square value is 0.004, indicating that the 0.4% change in Sales can be explained by the factors of Price, StockQuantity, and NumReviews. As a result, it is not possible to properly assess how the independent variable affects the dependent variable. The testing fails the F test (F=1.335, p=0.261>0.05). This implies that NumReviews, Price, and StockQuantity have no effect on Sales.

DOI: 10.54254/2753-8818/52/2024CH0163

Table 3. Parameter Estimates

	Unstandardized Coefficients		Standardized Coefficients	t	p	Multicollinearity	
	B	Std. Error	Beta			VIF	Tolerance
Constant	925.620	58.452	-	15.836	0.000**	-	-
NumReviews	0.022	0.013	0.056	1.777	0.076	1.001	0.999
Price	0.117	0.130	0.028	0.897	0.370	1.000	1.000
StockQuantity	-0.000	0.063	-0.000	-0.003	0.998	1.000	1.000
R 2	0.004						
Adj R 2	0.001						
F	F (3,996) =1.335, p=0.261						
D-W Value	1.995						

Notes: Dependent Variable=Sales
* p<0.05 ** p<0.01

4. Conclusion

This study examined consumer purchasing behavior using the Kaggle dataset "Exploring E-commerce Trends" through the application of Multiple Linear Regression and Random Forest models. The Random Forest analysis revealed that the number of reviews, price, and stock quantity are significant predictors of sales. However, the model overfitted, showing good performance on the training set but poor performance on the testing set. On the other hand, the Multiple Linear Regression model's limited efficacy was seen in its low explanatory power for sales. The aforementioned results indicate that although the Random Forest model was able to identify significant aspects impacting sales, its overfitting problem and the inadequacy of the Multiple Linear Regression model underscore the necessity for more resilient techniques to precisely forecast client buying patterns in e-commerce. Therefore, in order to improve prediction accuracy and commercial decision-making, it is advised to incorporate new factors and use more advanced models.

The insights gained from this research have significant implications for e-commerce businesses. Understanding the key factors that influence sales can help businesses optimize their pricing strategies, manage stock more effectively, and enhance their online presence through customer reviews. The identification of overfitting in the Random Forest model underscores the importance of model validation and the need for techniques that can generalize well to new data.

The limited explanatory power of multiple linear regression and overfitting in Random Forest models could be addressed in future research by introducing more variables and utilizing sophisticated modeling approaches. Enhancing understanding of e-commerce consumer behavior can result in more accurate predictive models for improved business strategies. Real-time data and external elements like seasonal trends and economic situations can be incorporated.

References

[1] Jain V, et al. 2021 An overview of electronic commerce (e-Commerce). J. Contemp. Issues Bus. Gov., 27(3), 665-670.
[2] Kedah Z 2023 Use of e-commerce in the world of business. Startupreneur Bus. Digit. (SABDA J.), 2(1), 51-60.
[3] Ayodeji O G 2021 E-retail factors for customer activation and retention: An empirical study from Indian e-commerce customers. J. Retail. Consum. Serv., 59, 102399.
[4] Hufnagel G, Schwaiger M and Weritz L 2022 Seeking the perfect price: Consumer responses to personalized price discrimination in e-commerce. J. Bus. Res., 143, 346-365.
[5] Garbarino E and Maxwell S 2010 Consumer response to norm-breaking pricing events in e-commerce. J. Bus. Res., 63(9), 1066-1072.

DOI: 10.54254/2753-8818/52/2024CH0163

[6] Hansen R and Tambo T 2011 Branding and Channel Issues in E-Commerce from an Information System's Perspective. Proc. IRIS, 200-215.

[7] Kaplan A M and Haenlein M 2010 Users of the world, unite! The challenges and opportunities of social media. Bus. Horiz., 53(1), 59-68.

[8] Raji M A, Olodo H B, Oke T T, Addy W A, Ofodile O C, Oyewole A T 2024 E-commerce and consumer behavior: A review of AI-powered personalization and market trends. GSC Adv. Res. Rev., 18(3), 66-77.

[9] Liao M, Sundar S S 2022 When e-commerce personalization systems show and tell: Investigating the relative persuasive appeal of content-based versus collaborative filtering. J. Advert., 51(2), 256-267.

[10] Liu L 2022 e-Commerce Personalized Recommendation Based on Machine Learning Technology. Mob. Inf. Syst., 1, 1761579.

DOI: 10.54254/2753-8818/52/2024CH0153

Time Series Analysis of the Connection between Exchange Rate of EUR/USD, Bond Yields, and Trade Growth

Yiwen Zhang

Graduate School of Art and Science, Columbia University, New York, 10027, The United States

yz4563@columbia.edu

Abstract. In the context of the growing importance of global trade, the US dollar plays a crucial role as the currency for international transactions. The exchange rate of the dollar and the monetary policies of other countries influence each other. This study delves into the connection among the exchange rate of Euro/United States dollar (EUR/USD), the 10-year government bond rates with the international trade growth rates of 20 European countries from January 1991 to December 2016. By employing descriptive statistics, correlation analysis, regression modeling, and time series analysis, this research aims to uncover the underlying dependencies and predictive insights among these financial variables. The results indicate that the EUR/USD exchange rate has an insignificant impact on bond yields ($R^2 = 0.003$), suggesting that other factors such as monetary policy, inflation expectations, and market liquidity play more significant roles. However, a weaker currency is linked to lower growth rates is shown by a strong negative correlation for the exchange rate and trade growth rates ($R^2 = 0.210$). The multivariate regression model shows that bond yields and trade growth rates together explain 44.5% of the exchange rate variations, highlighting the interconnections among these economic variables. These findings highlight the necessity for policymakers to take these interconnected factors into account when developing economic and monetary policies to ensure stable and sustainable economic growth.

Keywords: Time series, Exchange rate, Bond yield, International trade growth rate.

1. Introduction

In today's world, where global trade is becoming ever more crucial, the US dollar plays a pivotal role as the settlement currency for numerous international transactions. The exchange rates between United States dollar (USD) and the currencies of other nations, as other other nations' monetary policies, interact with one another [1]. For the European Economic Community, the dollar and the euro influence each other, and European economic and monetary policies are often affected by changes in the dollar-euro exchange rate. Government bond yields, particularly the 10-year benchmark yield, serve as important tools for adjusting the impact of euro-dollar exchange rate fluctuations, reflecting long-term interest rates and economic expectations [2, 3]. Although changes in bond yields are influenced by various factors such as economic expectations and price trends, exchange rates are also a significant factor [4]. Additionally, the international trade growth rate, which represents the growth of international trade, is directly affected by exchange rates, and there is a reciprocal influence between international trade

DOI: 10.54254/2753-8818/52/2024CH0153

growth rates and government bond yields [5]. In this context, studying the relationships among European and American exchange rates, European government bond yields, and European international trade growth rates has become a hot topic of research.

Researching the changes in exchange rates that lead to changes in government bond yields holds significant value for both investors and government departments. It helps investors achieve higher returns and allocate investment funds more reasonably, and it assists government departments in preparing measures to respond to exchange rate changes or understanding potential changes in government bond yields through exchange rate movements [6]. Additionally, understanding how exchange rates affect international trade growth rates can directly aid both sides of the trade in formulating trade plans and determining trade partners.

Existing research usually primarily the method of the individual effects of exchange rates on either bond yields or trade growth rates, often neglecting the comprehensive interactions among these variables. These studies have shown that while exchange rates can influence economic indicators, the degree and nature of these impacts varies greatly depending on the economic context with policy context [7, 8]. However, a significant limitation in past research is the insufficient exploration of the combined effects of exchange rates, bond yields, and trade growth rates. This gap in the literature highlights the need for a more integrated approach to understand the multifaceted relationships among these critical economic variables.

Moreover, the interplay between these economic variables can offer insights into broader economic conditions and trends. For instance, fluctuations in exchange rates may signal underlying shifts in economic strength, influencing investor confidence and market stability [9]. Similarly, variations in government bond yields can reflect changes in monetary policy or economic outlook, impacting fiscal strategies and investment decisions [10]. This multifaceted relationship underscores the complexity of global economic interactions and the importance of comprehensive analysis in understanding these dynamics.

This study aims to analyze the relationship among the monthly exchange rates of EUR/USD, the 10-year government bond yields with international trade growth rates of 20 European countries from January 1991 to December 2016. It also explores the correlation coefficients among the three variables and the pairwise correlation coefficients between each pair of variables. By utilizing descriptive statistics, correlation analysis, regression modeling, and time series analysis, this research seeks to uncover the underlying dependencies and predictive insights among these financial variables. This comprehensive approach is intended to address the gaps in current research and offer a more profound insight into the interconnectedness of these economic elements.

2. Methods

2.1. Data Source

In this research, used data includes the monthly exchange rates of EUR/USD (settlement data on the first of each month), average European government bond yields, and the international trade growth rates of 20 European countries from January 1, 1991, to December 1, 2015. The data of EUR/USD exchange rate is sourced from the Federal Reserve Economic Data (FRED), European government bond yield data is from the European Central Bank, and European international trade growth rate data is from Eurostat.

2.2. Variable Selection

The original data covers different time periods, ranging from January 1970 to June 2024. To specifically examine the relationship between the euro and the dollar exchange rates, European government bond yields, and international trade growth rates after the formation of the European Community, data from January 1991 to December 2015 was extracted using Python. Further, Python tools were used to isolate the columns representing time and the three research factors from the three datasets, forming a new file with four columns: Time, Exchange Rate, Bond Yield, and Growth Rate. Instead of fixing the dependent variable, this paper sets different dependent and independent variables sequentially and compares the resulting R-values to identify the most influential two or three factors.

2.3. Method Introduction

This paper uses time series combined with multiple linear regression equations for the study. To achieve the research objectives, time series were used to visualize the three datasets, followed by multiple linear regression analysis of the three research factors. The formulas used are:

$$\text{Exchange Rate} = \beta_0 + \beta_1 \cdot \text{Bond Yield} + \beta_2 \cdot \text{Growth Rate} + \varepsilon \tag{1}$$

$$\text{Exchange Rate} = \beta_0 + \beta_1 \cdot \text{Bond Yield} + \varepsilon \tag{2}$$

$$\text{Exchange Rate} = \beta_0 + \beta_1 \cdot \text{Growth Rate} + \varepsilon \tag{3}$$

The analysis aims to explore the mutual influence coefficients among the three factors and determine whether the mutual influence among the three factors is stronger than the influence between any two factors alone.

3. Results and Discussion

3.1. Descriptive Analysis

First, time series of Exchange Rate, Bond Yield, and Growth Rate data were visualized (Figure 1), showing different trends among the three. The Exchange Rate fluctuated continuously but with a small overall amplitude; the Bond Yield declined amid fluctuations; the Growth Rate rose amid fluctuations, with a high growth proportion.

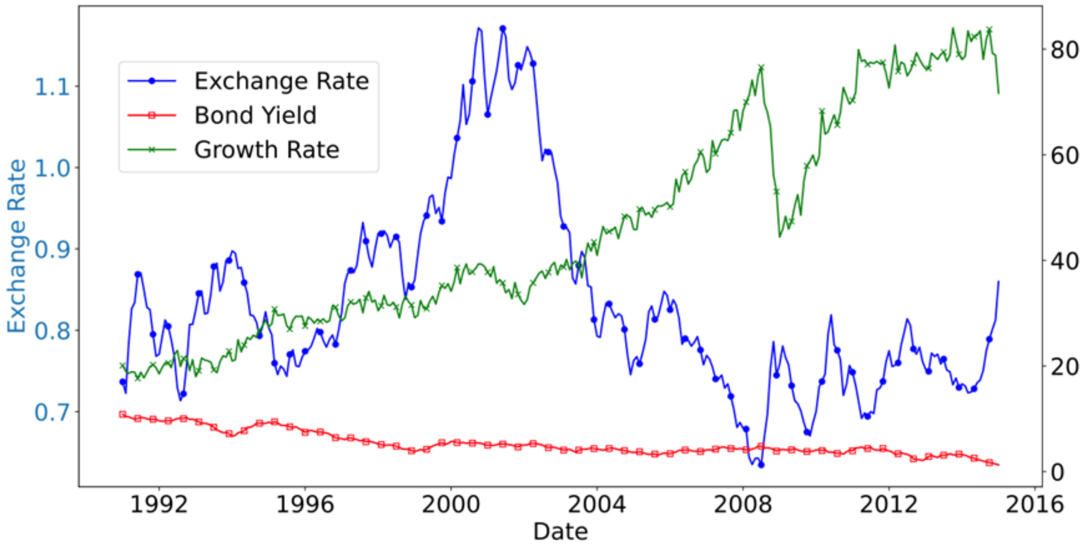


Figure 1. Time Series of Exchange Rate, Bond Yield, and Growth Rate

From Figures 2 and 3, there are correlations between the Exchange Rate and Bond Yield and Growth Rate, though the correlations are weak and negative. A detailed look at Figure 2 reveals a dense scatter plot of Exchange Rate and Bond Yield, but it does not exhibit a clear correlation. In contrast, Figure 3 shows a more dispersed scatter plot of Exchange Rate and Growth Rate, indicating a more evident negative correlation.

DOI: 10.54254/2753-8818/52/2024CH0153

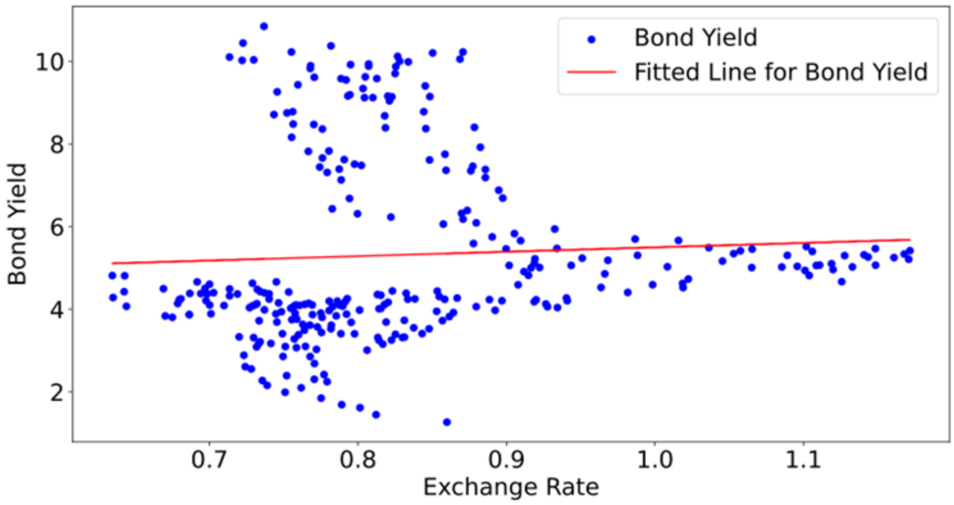


Figure 2. Regression Analysis: Exchange Rate Influencing Bond Yield

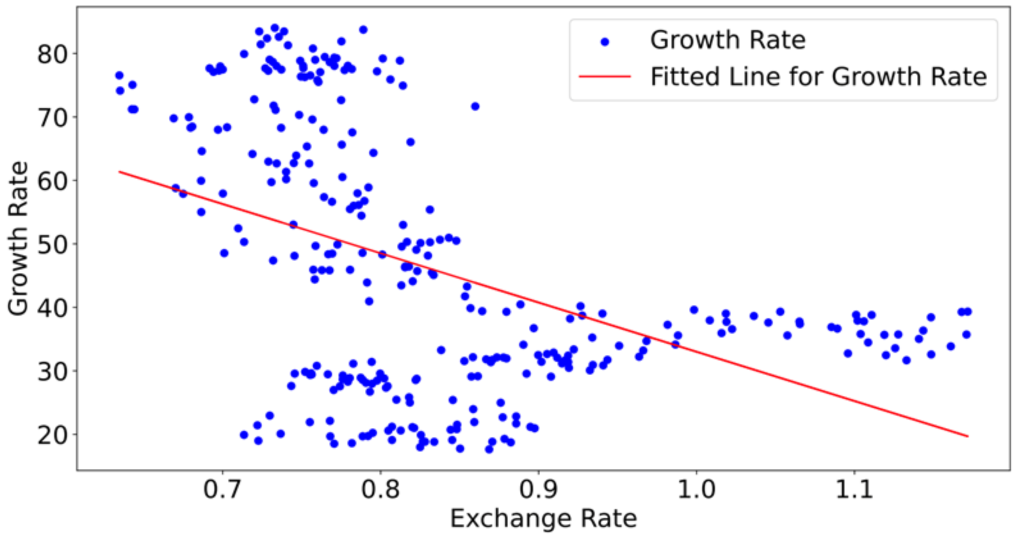


Figure 3. Regression Analysis: Exchange Rate Influencing Growth Rate

3.2. Model Results

The interaction of the Exchange Rate on two key variables: Bond Yield and Growth Rate, followed by understanding how Bond Yield and Growth Rate influence the Exchange Rate through multivariate regression analysis, is the main propose for this paper. The summary of the results is as follows: Multiple regression analysis reveals several key insights into the connection of Exchange Rate, Bond Yield, and Growth Rate. It shows that the interdependencies among these variables are complex and multifaceted. The results highlight the significance of considering multiple factors simultaneously rather than in isolation, as previous studies have often done. By incorporating both Bond Yield and Growth Rate into the analysis, this study provides a more holistic view of how these economic indicators interact and influence each other. This approach helps to uncover the underlying dynamics that drive exchange rate fluctuations and their broader economic implications. Moreover, the findings suggest that policy measures aimed at stabilizing exchange rates should take into account their potential impact on both bond yields and trade growth rates to achieve more effective economic outcomes.

Exchange Rate and Bond Yield (Table 1): The extremely low R-squared value (0.003) and insignificant coefficients indicate a fact: the interaction of Exchange Rate on Bond Yield is negligible. This suggests that other factors, such as monetary policy, inflation expectations, and market liquidity, have a greater influence in shaping bond yields. The bond market dynamics are intricate and can be affected by a range of economic indicators, both domestic and international.

Table 1. Linear Regression for Exchange Rate and Bond Yield

	Unstandardized Coefficients		Standardized Coefficients	t	p	Collinearity Diagnostics	
	B	SE	Beta			VIF	Tolerance
Constant	0.817	0.018	-	45.107	0.000**	-	-
Bond Yield	0.003	0.003	0.057	0.963	0.336	1.000	1.000
R^2	0.003						
Adj R^2	-0.000						
F	F (1,287)=0.927,p=0.336						
D-W value	0.030						

Dependent Variable: Exchange Rate

* p<0.05 ** p<0.01

Exchange Rate and Growth Rate (Table 2): The analysis shows a more significant relationship between Exchange Rate and Growth Rate, with an R-squared value of 0.210. The negative coefficient indicates that a higher exchange rate (implying a depreciated currency) is associated with a lower growth rate. This is consistent with economic theory, where a weaker currency makes imports more expensive, reducing consumer purchasing power and inhibiting economic growth. The significant influence from exchange rate to growth rate underscores the importance of maintaining currency stability for economic stability and growth.

Table 2. Linear Regression for Exchange Rate and Growth Rate

	Unstandardized Coefficients		Standardized Coefficients	t	p	Collinearity Diagnostics	
	B	SE	Beta			VIF	Tolerance
Constant	0.957	0.016	-	61.659	0.000**	-	-
Growth Rate	-0.003	0.000	-0.458	-8.730	0.000**	1.000	1.000
R^2	0.210						
Adj R^2	0.207						
F	F (1,287)=76.209,p=0.000						
D-W value	0.036						

Dependent Variable: Exchange Rate

* p<0.05 ** p<0.01

Multivariate Regression Analysis (Table 3): The multivariate regression model indicates that Bond Yield and Growth Rate collectively explain 44.5% of the variance in the Exchange Rate. Both factors have negative and significant coefficients, suggesting that higher bond yields and growth rates are associated with a stronger currency (lower exchange rate). This result highlights the interconnections among these economic variables. Higher bond yields can attract foreign investment, increasing demand for the domestic currency and thus enhancing its value. Similarly, robust economic growth can boost investor confidence and capital inflows, further supporting the currency's value.

DOI: 10.54254/2753-8818/52/2024CH0153

Table 3. Linear Regression for Exchange Rate, Growth Rate, Bond Yield

	Unstandardized Coefficients		Standardized Coefficients	t	p	Collinearity Diagnostics	
	B	Std. Error	Beta			VIF	Tolerance
Constant	1.344	0.037	-	35.889	0.000**	-	-
Bond Yield	-0.042	0.004	-0.779	-11.011	0.000**	2.582	0.387
Growth Rate	-0.006	0.000	-1.068	-15.090	0.000**	2.582	0.387
R 2	0.445						
Adj R 2	0.441						
F	F (2,286)=114.690,p=0.000						
D-W valuer	0.070						

Dependent Variable: Exchange Rate

* p<0.05 ** p<0.01

4. Conclusion

This research offers important understanding of the connections between three factors: Exchange Rate, Growth Rate, and Bond Yield. To summarize, the influence of Exchange Rate on Bond Yield is minimal, whereas its effect on Growth Rate is notably negative. When Growth Rate and Bond Yield act together on Exchange Rate, the negative effect is very significant and much greater than the sum of their individual effects on Exchange Rate. The output implies when the Exchange Rate falls, the Growth Rate rises; when both the Growth Rate and Bond Yield decline, the Exchange Rate increases significantly. Therefore, comparing univariate linear regression and multivariate linear regression analysis, the multivariate regression clearly provides better explanatory power, indicating that for models concerning Exchange Rate, Bond Yield, and Growth Rate, multivariate regression is more closely aligned with real data and has practical significance. These findings underscore the importance for policymakers to take these interconnected factors into account when devising economic and monetary policies to achieve stable and sustainable economic growth.

References

[1] Joscha, B. and Robert. C. (2017) Exchange rate expectations and economic policy uncertainty. European Journal of Political Economy, 47, 148-162.

[2] Degasperi, R., Hong. S. and Ricco, G. (2020) The Global Transmission of U.S. Monetary Policy. CEPR Discussion Paper, 14533.

[3] Riccardo, D., Seokki, Hong. and Giovanni, R. (2020) The Global Transmission of U.S. Monetary Policy. CEPR Discussion Paper, 14533.

[4] Valchev, R. (2020) Bond Convenience Yields and Exchange Rate Dynamics. American Economic Journal: Macroeconomics, 12(2), 124-66.

[5] Morina, F., Eglantina, H., Uğur, E., Mirela, P. and Marian, C.V. (2020) The Effect of Exchange Rate Volatility on Economic Growth: Case of the CEE Countries. Journal of Risk and Financial Management, 13, 177.

[6] Akalpler, E. and Duhok, D. (2018) Does monetary policy affect economic growth: evidence from Malaysia. Journal of Economic and Administrative Sciences, 34, 2-20.

[7] Suhadak and Dwi Suciany, A. (2020) Brief technical note: The influence of exchange rates on inflation, interest rates and the composite stock price index: Indonesia 2015 - 2018. Australasian Accounting, Business and Finance Journal, 14.

[8] Maganioti, Y. and Han, J. (2020) The Impact of Exchange Rate Fluctuation on Economic Growth – Empirical Studies Based on Different Countries. Proceedings of the 4th International Symposium on Business Corporation and Development in South-East and South Asia under B&R Initiative, Atlantis Press, 29-33.

DOI: 10.54254/2753-8818/52/2024CH0153

[9] Bhargava, V. and Konku, D. (2023) Impact of exchange rate fluctuations on US stock market returns. Managerial Finance, 49(10), 1535-1557.
[10] Akram, T. (2021) A Note Concerning the Dynamics of Government Bond Yields. The American Economist, 66(2), 323-339.

DOI: 10.54254/2753-8818/52/2024CH0135

Non-Contact Human Body Dimension Estimation Methods Based on Deep Learning: A Critical Analysis

Yichen Zhang

School of Artificial Intelligence, South China Normal University, Guangzhou, China

20212005177@m.scnu.edu.cn

Abstract. With the advancement of information technology, Human Body Dimension Estimation (HBDE) is moving towards digitalization and non-contact methods. This paper critically reviews non-contact HBDE methods based on deep learning, primarily covering image-based methods, the Adaptive Body Structure Segmentation (ABSS) algorithm with height and weight measurements, and 3D meshes or point cloud data methods. Convolutional Neural Networks (CNNs) excel in image measurement, while the ABSS algorithm with adaptive segmentation achieves high accuracy. Recent studies, such as the Grey Wolf Optimizer-Elman Neural Network (GBWO-ENN) model, have significantly improved prediction accuracy by optimizing neural network structures. Additionally, 3D point cloud data combined with the Skinned Multi-Person Linear (SMPL) model has proven effective in complex scenarios. This paper summarizes the performance of various methods, analyzes common datasets, and explores future development directions for non-contact HBDE. Despite significant progress, this field still faces challenges such as data quality, scarcity of extreme samples, and privacy issues. Future research should focus on building high-quality datasets and developing lightweight measurement networks to meet practical application needs.

Keywords: Non-contact measurement, human body dimension estimation, deep learning, convolutional neural networks.

1. Introduction

With the rapid development of information technology, Human Body Dimension Estimation (HBDE) is advancing towards digitalization and non-contact methods. Human body dimensions refer to measurements of various parts of the body such as height, weight, or more detailed metrics like shoulder width or chest circumference. Reliable methods for measuring the former, such as scales for weight, already exist. However, methods for measuring the latter often have significant errors in everyday scenarios despite being of greater reference value in primary applications such as clothing design.

Currently, non-contact HBDE methods mainly rely on 3D body scanning devices[1] or combine human depth data captured through RGB-D cameras. However, these methods face challenges such as high costs, stringent environmental requirements, and poor portability. Research on non-contact HBDE is still relatively scarce, and comprehensive reviews are particularly lacking. Therefore, this paper aims to provide a comprehensive, extensive, and in-depth summary of non-contact HBDE methods based on deep learning.

Starting from image-based measurement methods, methods based on the Adaptive Body Structure Segmentation (ABSS) algorithm with height, weight, and circumference values, and methods based on 3D meshes or 3D point cloud data, this paper reviews recent research achievements in the field of non-contact HBDE. It analyzes and compares the main models and performances of various methods and looks forward to the future development of HBDE.

2. Related Works

Non-contact HBDE relies on neural networks trained on a large number of human body dimension measurement data samples. This section introduces some key related works. Table 1 summarizes several models used for HBDE and their backbone networks.

Table 1. Network models and its backbone

Network Models	Backbone
Neural Anthropometer	CNN
Shapy	CNN
BMnet	CNN+GBWO-ENN
GBWO-ENN	ENN
IWOA-ENN-MC	ENN
UGA-BP-MC	BP
PC-BoDiEs	MLP
Conv-BoDiEs	CNN

Human Body Dimension Estimation: Non-contact methods for HBDE analyze images taken from different perspectives to measure various body dimensions without direct contact with the subject. In 2016, Dibra et al.[2] first attempted to use Convolutional Neural Networks (CNNs) to recover 3D human meshes and body measurements from silhouettes, training on synthetic silhouette images generated from the CAESAR dataset. In 2020, Hu et al.[3] proposed a human body multi-feature point extraction and dimension measurement algorithm based on ABSS, called Human pesm-abss, which accurately acquires dimension information of various body parts. In 2021, Hu et al.[4] introduced a model based on a BP network, integrating the Upgrade Genetic Algorithm(UGA) and Markov Chain(MC) method, resulting in a simpler neural network structure. In 2022, Nataniel Ruiz et al.[5] proposed using adversarial generative networks to increase the number of extreme samples and combine multi-view images, significantly reducing measurement errors. In 2024, Yang et al.[6] proposed a human body dimension prediction method based on an Elman Neural Network (ENN) network, which improved global stability and prediction accuracy compared to Hu's Upgrade Genetic Algorithm-Back Propagation-Markov Chain(UGA-BP-MC)[4] and the 2023 Improved Whale Optimization Algorithm-Elman Neural Network-Markov Chain(IWOA-ENN-MC)[7] neural network models.

Skinned Multi-Person Linear model(SMPL): SMPL[8] is a statistical model for 3D human shape and pose estimation based on 3D mesh vertices. It aims to capture the geometric shape and motion pose of the human body by training on large-scale 3D human scan data, and it is widely used in the field of computer vision.

Human pesm-abss algorithm (referred to as ABSS algorithm below): This is a method for extracting feature points and measuring human dimensions from 2D images based on Adaptive Body Structure Segmentation[3]. The method combines height proportion and image preprocessing techniques to address issues with traditional image detection models, such as long detection times and inaccurate key points. This method can obtain relatively accurate human height, gender, and circumference width and thickness results from 2D human images.

3. Summary of Non-Contact HBDE Methods

Generally, non-contact HBDE methods based on deep learning often rely on neural networks, with different methods employing different types of neural networks. Additionally, various non-contact

DOI: 10.54254/2753-8818/52/2024CH0135

HBDE methods often require different types of input data, which has significant implications for the choice of methods in different application scenarios. Therefore, this paper categorizes these methods based on the types of input data required during the prediction process into image-based human dimension estimation methods, estimation methods based on the ABSS algorithm[3] with height, weight, and circumference values, and estimation methods based on 3D meshes or 3D point cloud data. This section will introduce representative methods from these categories, and the next section will compare their performance based on prediction results from specified datasets.

3.1. Image-Based Human Dimension Estimation Methods

Classic image-based measurement methods use CNNs to extract features from RGB images. These extracted features are further processed to generate 3D human meshes based on these parameters or detect key points using specially trained neural networks. Traditional methods may also involve manual operations to obtain key point positions of the human body, including shoulders, elbows, wrists, etc. The key points obtained by the above methods are finally used to infer specific human measurement data through geometric calculations and regression models. This subsection will briefly introduce two representative methods and describe the improvements they have made based on the above basic implementation logic.

3.1.1. BMnet. This method[4] uses Generative Adversarial Networks (GANs) and CNNs for human dimension estimation, enhancing model robustness through adversarial samples. First, preprocessed front and side human images are input, and a unique adversarial body simulator is introduced to fill the gap of extreme individual data. After feature extraction by CNN, the extracted feature maps generate high-dimensional feature vectors through multiple layers of convolution and pooling operations. Subsequently, the feature vectors are input into a key point detection network to detect the coordinates of human key points, including shoulders, elbows, wrists, hips, knees, and ankles, and further infer specific human measurement data. Finally, the model outputs highly accurate human measurement results, demonstrating the effectiveness of using adversarial samples and multi-stage key point regression methods in human dimension estimation. Additionally, the model also includes tests on multi-view inputs and height and weight inputs, all achieving ideal results.

3.1.2. Shapy. This method[9] uses CNNs to extract human features from RGB images, which are used to regress the parameters of the human model. Shapy still employs the method of generating 3D human meshes and recognizing human key points to obtain measurement values. Additionally, Shapy introduces a method to enhance model training through crowd-rated language shape attributes. By collecting linguistic descriptions of human shapes, such as "tall," "slim," "muscular," etc., and pairing them with the obtained 3D shapes, Shapy introduces loss functions related to language shape attributes in training, optimizing the model multidimensionally in combination with measurement value loss functions.

3.2. Estimation Methods for Height, Weight, and Circumference Values Based on the ABSS Algorithm

Compared to the image-based human dimension estimation methods mentioned in the previous section, methods based on the ABSS algorithm for height, weight, and circumference values usually do not use CNNs as their core neural networks. Unlike traditional methods, these methods do not involve generating or using 3D body models but instead perform adaptive segmentation of key areas of the human structure based on orthogonal human images to obtain gender, height and weight values, and circumference values, and directly predict the outputs. Additionally, some optimization algorithms or stability and fluctuation handling methods are also included to optimize the neural network. This subsection will mention several representative methods and briefly introduce the neural networks and optimization methods they employ.

DOI: 10.54254/2753-8818/52/2024CH0135

3.2.1. GBWO-ENN. Grey Black Wolf Optimization–Elman Neural Network(GBWO-ENN) method[6] improves the traditional Grey Wolf Optimization Algorithm (GWO) by addressing its problems of easily falling into local optima and balancing global and local searches. The proposed GBWO algorithm uses a nonlinear decreasing method to reduce the convergence coefficient, thereby optimizing the algorithm's performance. GBWO is then used to optimize the ENN, using the weights and thresholds output by the trained GBWO model as the initial parameters for ENN training, achieving good accuracy and robustness in human dimension prediction.

3.2.2. A-BP-MC. This method[4] addresses the issues of the widely used BP neural network for human dimension prediction, such as easily falling into local optima and improper initial weight and threshold assignments. By optimizing the BP neural network with the improved UGA, and using the Markov Chain method to improve model accuracy and prediction stability in unstable environments. The main execution steps of the model include input image, body type classification, adaptive segmentation of key areas of the human structure based on the ABSS algorithm[3], feature dimension extraction, UGA-BP model training, and Markov correction model training, achieving good robustness and accuracy in human dimension prediction.

3.3. Human Dimension Estimation Methods Based on 3D Meshes or 3D Point Clouds

Human dimension estimation methods based on 3D meshes or 3D point clouds often face challenges such as high data acquisition difficulty, significant cost, and scarcity of datasets. Alldieck et al. [10-12] proposed using depth cameras to acquire human depth information, then reconstructing the human body through the SMPL model and the Poisson method[13] to obtain circumference information of various parts. In the methods introduced in this section, 3D point clouds are directly used as input data, and the SMPL model is used to generate 3D human data and obtain measurements of various parts, which are then used for subsequent model training and validation. This subsection will introduce two typical human dimension estimation methods.

3.3.1. Neural Anthropometer. This model[14] proposes a classic human dimension estimation method using data generated by the SMPL model to obtain measurements of various parts of the human body through 3D modeling. Combined with the Blender software package, it synthesizes 2D images to train CNN, enabling the estimation of human dimensions from input 2D images. This method overcomes the scarcity of traditional human dimension data while providing a reliable human dimension estimation method.

3.3.2. Point Cloud Body Dimensions Estimation (PC-BoDiEs). This model[15] uses stacked Multi-Layer Perceptron (MLP) convolution layers to extract global and local features from point cloud data and regress 16 predefined human dimensions. Each MLP layer is followed by a ReLU activation function to enhance the model's ability to handle nonlinear problems. The input to the PC-BoDiEs model is unorganized 3D body scan data merged from two viewpoints. This data is sampled using farthest point sampling to match a fixed number of points before being fed into the network, reducing the model's time and memory requirements. During training, the model parameters are optimized using the AMSGrad variant of the Adam optimizer. Throughout the training process, the learning rate is gradually reduced using a cosine decay strategy. The PC-BoDiEs model, by processing data from unorganized 3D point clouds, overcomes the cumbersome process of fitting human scan data to predefined body models, directly regressing body dimensions from point cloud data, effectively improving the accuracy and efficiency of human dimension estimation from 3D data.

DOI: 10.54254/2753-8818/52/2024CH0135

4. Major Datasets and Experimental Results Comparison

4.1. Major Datasets

Table 2 introduces commonly used datasets in HBDE tasks. According to the number of samples, the number of measurement items, and the types of data included in different datasets, they are suitable for various HBDE tasks.

Table 2. Commonly used datasets in HBDE tasks, including the datasets' name, number of samples, number of measurement items and the type of data used

Dataset Name	**Number of Samples (People)**:	**Number of Measurement Items**	**Data Type**
CAESAR	2400	40+	2D、3D、Measurement Values
MAD	4419	5	2D、Measurement Values
BodyM	2505	14	2D、3D、Measurement Values
ANSUR Ⅱ	13000+	100+	Measurement Values
Human3.6M	17	-	2D、3D

(1) CEASER (Civilian American and European Surface Anthropometry Resource): This dataset includes samples of adults from North America and Europe, providing detailed 3D body scans and body measurement data. It is widely used in generating 3D human mesh data using the SMPL model.
(2) MMTS (Model Measurements Test Set): This dataset comes from several modeling agency websites and includes high-resolution multi-view human RGB images and detailed body measurement data. It was used in the training of the Shapy model [10], where researchers added semantic labels.
(3) BodyM Dataset: This is an advanced dataset specifically for human pose estimation and 3D human reconstruction. It contains high-precision 3D body scans and pose annotations. The provided 3D human data includes 50 key points and was collected in a laboratory environment using different camera angles and lighting conditions to simulate outdoor image capture. Some subjects were photographed multiple times in different clothing to enhance the robustness of the dataset.
(4) ANSUR II (Anthropometric Survey of U.S. Army Personnel): This dataset is a body measurement survey conducted by the U.S. military, aiming to collect and analyze the anthropometric data of U.S. soldiers. It has a large number of samples, containing over 100 body measurement items, but it does not include images, and all subjects are U.S. military personnel.
(5) Human3.6M Dataset: This dataset contains approximately 3.6 million 3D human poses and corresponding image data from 11 professional actors (6 men and 5 women) performing in 17 different scenarios. The dataset includes pixel-level labels for 24 body parts of these actors and high-precision 3D joint data, which can be used to infer and calculate various human body dimensions.

4.2. Evaluation Metrics and Results

Among the methods introduced in Section 3, all methods used Mean Absolute Error (MAE) as one of the standards to evaluate model performance in the experiments. Most of the experiments recorded results for chest circumference, waist circumference, and hip circumference. Although these experiments were conducted on different datasets, they still allow for an objective evaluation of the models' performance to some extent. Table 3 shows the MAE results obtained from evaluating different models on different datasets.

DOI: 10.54254/2753-8818/52/2024CH0135

Table 3. Results of the models tested on different datasets, using MAE as evaluation metrics.

Model Name	Evaluation Metrics(MAE/mm)				Evaluation Dataset
	Chest	Hips	Waist	Overall	
Neural Anthropometer	25.22	27.53	25.85	26.20	SMPL Generated
Shapy	65.00	57.00	69.00	63.67	HBW(Human Body Wild)
	64.00	74.00	98.00	78.67	MMTS
BMnet(Single View)	33.95	31.03	31.93	32.30	BodyM
BMnet(Multi View+Height+Weight)	15.92	9.74	15.44	**13.70**	
GBWO-ENN	13.80	8.50	8.40	**10.23**	ANSURⅡ
IWOA-ENN-MC	15.91	17.16	17.85	16.97	CAESAR
UGA-BP-MC	29.10	23.00	20.00	24.03	ANSURⅡ
PC-BoDiEs	32.90	-	22.90	27.90	CAESAR
Conv-BoDiEs	25.70	-	16.50	33.95	CAESAR

It can be learnt from the table that PC-BoDiEs and Conv-BoDiEs (using grayscale images) have an overall result calculated as the average of only two metrics due to the missing hip test values. It is worth mentioning that the original test results included a larger number of metrics; for comparison purposes, only a subset is used here. The original MAE recorded for Conv-BoDiEs (using grayscale images) was 46.40mm, and for PC-BoDiEs, it was 49.50mm. Similarly, due to the large number of measurement metrics in the original experimental results, the Neural Anthropometer model had an overall MAE of 20.89mm when all measurement metrics were considered.

4.3. Evaluation Results Summary

As shown in Table 3, methods based on the ABSS algorithm for height, weight, and circumference values generally have higher accuracy. Among the image-based methods, both Neural Anthropometer and BMnet (Single View) exhibit good accuracy. It is worth noting that the latter shows a significant improvement in accuracy after incorporating multi-view images (not listed in the table), and its accuracy reaches an MAE of 13.70mm after including height and weight as metrics, almost comparable to the accuracy of GBWO-ENN, which does not rely on images. This highlights the important role of height and weight values and adversarial augmentation in improving the accuracy of human dimension estimation. It also provides a direction for future research: whether more accurate height and weight values can be obtained from images to assist in the measurement of more human dimension metrics, thereby enabling accurate measurement of multiple human dimensions using images alone.

5. Conclusion

As a research hotspot in the field of artificial intelligence in recent years, non-contact human body dimension estimation has significant practical implications in areas such as clothing design, health status estimation, and posture recognition. This paper categorizes human body dimension estimation methods based on the type of input data used: image-based human dimension estimation, estimation of height, weight, and circumference values using the ABSS algorithm, and estimation based on 3D human meshes or 3D point clouds. The paper summarizes the network models employed by these methods and their characteristics. It also introduces commonly used datasets and evaluation metrics in the field of human dimension estimation, analyzing the performance of different network models on various datasets.

Although significant research progress has been made in non-contact human body dimension estimation based on deep learning, several problems and challenges remain:

Quality of existing human image datasets needs improvement: The commonly used datasets contain some blurry images, and variations in gender and camera distance significantly impact the model's estimation accuracy. Additionally, some datasets used in research have a relatively homogeneous sample type, such as the ANSUR II dataset, which uses body measurement data from U.S. Army soldiers. These factors can affect the accuracy of the model's human dimension estimation results.

Scarcity of extreme samples and lack of representativeness: Current research on human dimension estimation based on 3D human meshes or 3D point clouds mainly relies on SMPL synthetic data, which may lead to insufficient generalization ability in practical applications. Most training datasets also lack samples of extreme body types, potentially reducing the model's prediction accuracy for these rare body types. Using adversarial Body Simulators (ABS) may significantly increase the time and cost of generating samples.

Occlusion issues and user privacy concerns: When the subject is wearing loose-fitting clothing, changes in body contour may affect the model's prediction results. Collecting 3D human data or precise 2D human images often requires subjects to wear tight-fitting clothing or no clothing, which raises potential privacy issues. Additionally, for methods based on the ABSS algorithm, optimizing the adaptive segmentation method to obtain accurate values is an important factor to consider.

Computational power requirements for model deployment: As the performance of human dimension estimation network models improves, the computational power required for model deployment and operation in real-life and production scenarios may not be met, preventing the models from achieving their expected prediction results.

In summary, addressing the above issues requires the establishment of comprehensive, well-classified human dimensions and image datasets with clear indicators. Improving the quality of datasets and synthetic data, and creating representative datasets, are the main directions for future dataset optimization. Current research methods are constrained by deployment costs and computational power requirements, making them difficult to apply in personal practical use. Therefore, the development of lightweight human dimension estimation networks will be a primary research direction in the future. Building models that can adapt to various input data in different application scenarios and addressing potential issues in obtaining input data are key to the widespread application of human dimension estimation networks.

References

[1] Song, Y., Wirta, J., Kämäräinen, J. K. (2020). Anthropometric clothing measurements from 3d body scans. *Machine Vision and Applications*, 31(4): 7.

[2] Dibra, E., Jain, H., Öztireli, C., Ziegler, R., Gross, M. (2016). Hs-nets: Estimating human body shape from silhouettes with convolutional neural networks. In *International Conference on 3D Vision (3DV)*, 108–117.

[3] Hu, X. R., Liu, J. W., Liu, J. P., et al. (2020). Multi-feature extraction and dimension measurement for dressed human bodies via adaptive body structure segmentation. *Computer Engineering*, 46(2): 238-246.

[4] Hu, X. R., Liu, J. W., Liu, J. P., et al. (2021). Prediction of 3D human dimensions based on improved GA-BP-MC neural network. *Computer Engineering and Science*, 43(08): 1443-1453.

[5] Ruiz, N., Bellver, M., Bolkart, T., Arora, A., Lin, M. C., Romero, J., Bala, R. (2022). Human body measurement estimation with adversarial augmentation. In *International Conference on 3D Vision (3DV)*, Prague, Czech Republic, September 12-16. IEEE: 219-230. DOI: 10.1109/3DV57758.2022.00031.

[6] Yang, X. W., Li, Y. T., Han, X., et al. (2024). Human dimension prediction model integrating GBWO and ENN. *Computer Technology and Development*, 34(06): 132-139. DOI: 10.20165/j.cnki.ISSN1673-629X.2024.0064.

[7] Hu, X. R., Liu, J. W. (2023). Prediction of human circumference dimensions based on improved WOA-ENN-MC model. *Computer Applications and Software*, 40(05): 190-199.

DOI: 10.54254/2753-8818/52/2024CH0135

[8] Bogo, A., Kanazawa, A., Lassner, P., et al. (2016). Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. *Proc of European Conference on Computer Vision (ECCV)*: 1.

[9] Choutas, V., Müller, L., Huang, C. H. P., Tang, S., Tzionas, D., Black, M. J. (2022). Accurate 3D body shape regression using metric and semantic attributes. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022)*: 2708-2718. DOI: 10.1109/ CVPR52688.2022.00274.

[10] Alldieck, T., Marcus, M., Xu, W. P., et al. (2018). Video-based reconstruction of 3D people models. *Proc of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 1.

[11] Alldieck, T., Marcus, M. (2019). Learning to reconstruct people in clothing from a single RGB camera. *Proc of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 1.

[12] Sun, Y. L., Miao, Y. W., Bao, C. (2019). Global registration cumulative error minimization based 3D human body reconstruction using RGB-D scanning data. *Journal of Computer-Aided Design & Computer Graphics*, 31(9): 1467-1476.

[13] Kazhdan, M., Hoppe, H. (2013). Screened Poisson surface reconstruction. *ACM Transactions on Graphics*, 32(3): Article No. 29.

[14] González, T. Y., Mayer, H. A. (2021). A neural anthropometer learning from body dimensions computed on human 3D meshes. *IEEE Symposium Series on Computational Intelligence (SSCI)*: 1-8.

[15] Škorvánková, D., Riečický, A., Madaras, M. (2022). Automatic estimation of anthropometric human body measurements. Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP): 537-544. DOI: 10.5220/0010878100003124.

DOI: 10.54254/2753-8818/52/2024CH0152

Study on the Influencing Factors of Heart Disease

Muyao Ding

School of Rosedale Global High School, Fuzhou, Fujian, 350005, China

sa@ossdedu.com

Abstract. With the continuous development of technology and scientific research, people are constantly exploring the mysteries of the universe and human. Since ancient times, most people have died under various incurable diseases, which has led to a need for more knowledge about diseases to help humans resist diseases and help people explore the world to the next level. According to years of data statistics, cardiovascular and cerebrovascular diseases are one of the diseases with the highest mortality rate, and heart disease is included in it. According to statistics, around 695000 people died from heart disease in 2021, which is an extremely large number, so this paper needs to conduct research on heart disease more. This experiment aims to identify and analyze data related to heart disease, as well as use correlation methods to discover the impact of different factors that contribute to heart disease. By doing so, experimental results can be obtained and conclusions can be drawn regarding the research topic. Through experimental results, it can help people understand the different degrees of influence of the causative factors of heart disease on heart disease, maintain good lifestyle habits in daily life, and provide some tips and assistance for medical treatment. And this experiment can provide some reference for future research on heart disease and contribute to the development of drugs and medical technology.

Keywords: Heart disease, pathogenic factors, correlation coefficient.

1. Introduction

With the development of society and technology, the global medical level is constantly improving, but it is not enough to support people's resistance to all diseases. According to the 2019 Global Health Assessment report, heart disease is one of the top ten diseases with the highest mortality rates, accounting for 16% of current deaths [1]. Since 2000, the number of deaths from heart disease has increased by over 2 million [2]. In the past two decades, the number of deaths due to heart disease has shown an overall trend of first decreasing and then increasing [3]. Moreover, the mortality rate of congenital heart disease is extremely high, because the organs in infants or children are not fully developed. Because of congenital heart disease, patients do not have high resistance, so the morality rate is high [4]. Therefore, people need to investigate and understand the different causes and potential triggering factors of heart disease and take corresponding treatment measures in a timely manner. This can help people reduce the probability of developing heart disease and take measures to cope with the disease, as well as alleviate the medical burden.

The main causes of heart disease can be divided into two categories. The first is congenital heart disease, which can be caused by genetic and non-genetic factors [5]. Congenital heart disease refers to abnormal cardiac performance caused by abnormal cardiovascular development during embryonic development in humans. Genetic factors, in short, refer to genes related to heart disease in family genes, while non genetic factors mainly include maternal factors, external environmental factors, etc. (pregnancy infections, gestational age, use or exposure to drugs, tobacco, alcohol, chemicals, working in a strong radiation environment, noisy working environment, exposure to toxic chemicals, passive smoking, heavy metal exposure, etc.), all of which can lead to heart disease in newborns [6]. Statistical and other techniques can be used to study the degree of relationship between these factors and heart disease.

Another type of heart disease is acquired heart disease, which refers to the heart disease that patients suffer from childbirth [7]. The main factor is that the body is induced to develop diseases by internal or external factors, and most of them are caused by poor lifestyle habits that lead to a decline in various aspects of the body's performance, thus affecting the adverse condition of the heart. The experimental results on the birth cohort effect of ischemic heart disease attributed to high salt diet in China from 1990 to 2019 showed that inducing heart disease due to high salt diet was the top factor among many, and the mortality rate caused by high salt diet in China increased with age, and the higher the age, the higher the Risk multiplier of mortality rate [8]. Not only high salt diet, there are more reason can cause acquired heart disease, for example: drinking, high MBI, low temperature, etc. [9]. Statistics show that the risk of mortality increases significantly among the population aged 60-64 to 80-84. Older people also have a weakened immune system, and most of the people who die from IHD are over 70 years old [10]. And according to the research data, women have a 3% higher rate of acquired heart disease than men, but death is lower than men. Therefore, both men and women should pay more attention to their daily routines and dietary habits to reduce the probability of developing heart disease. Therefore, to sum up, this article will go through various experimental investigations and analyse the data. The mortality rate of heart disease caused by different causes. Through this survey and research, people can know more about heart disease and give people some medical and health tips.

2. Methods

2.1. Data Source

The data for this literature comes from the Kaggle website, which conducted exploratory data analysis and machine learning modeling on heart disease prediction by Rashad Mammadov and was published and updated for one person a month ago.

Table 1. List of variables

variables	type	range
cholesterol	The total cholesterol content of various lipoproteins in the blood	150-349
blood pressure	The lateral pressure exerted on a unit area of the vessel wall by blood flowing within the vessel	90-179
heart rate	The number of beats per minute of the human heart	60-99
smoking	Smoking Status	Never/Currently/Other
alcohol intake	The energy of ingested alcohol	Heavy/None/Other
exercise time	Daily exercise volume	0-9
family history	Does the family have a history of heart disease	true/false

The data used in this article consists of a total of 1000 people, with half being male and half being female (Table 1). The age of the patients ranges from 25 to 79 years old. In this data, other eight variables (cholesterol, blood pressure, heart rate, smoking, alcohol intake, exercise time, family history, diabetes) are also included.

DOI: 10.54254/2753-8818/52/2024CH0152

2.2. Method Introduction

Taking heart disease as an example, in order to observe the relationship between the dependent variable and the independent variable and draw conclusions, this article needs to use a probability regression model, the probability of having heart disease (Y), eight independent variables (X), and observe the influence of X on Y through data analysis to determine the relationship between eight different factors and the probability of developing heart disease. Among them, logistic regression models were also applied [10].

3. Results and Discussion

3.1. Correlation Analysis

Correlation analysis refers to the correlation between a dependent variable and multiple independent variables. In this study experiment, the correlation between heart disease and multiple disease factors is investigated, as shown in figure 1, figure 2 and so on.

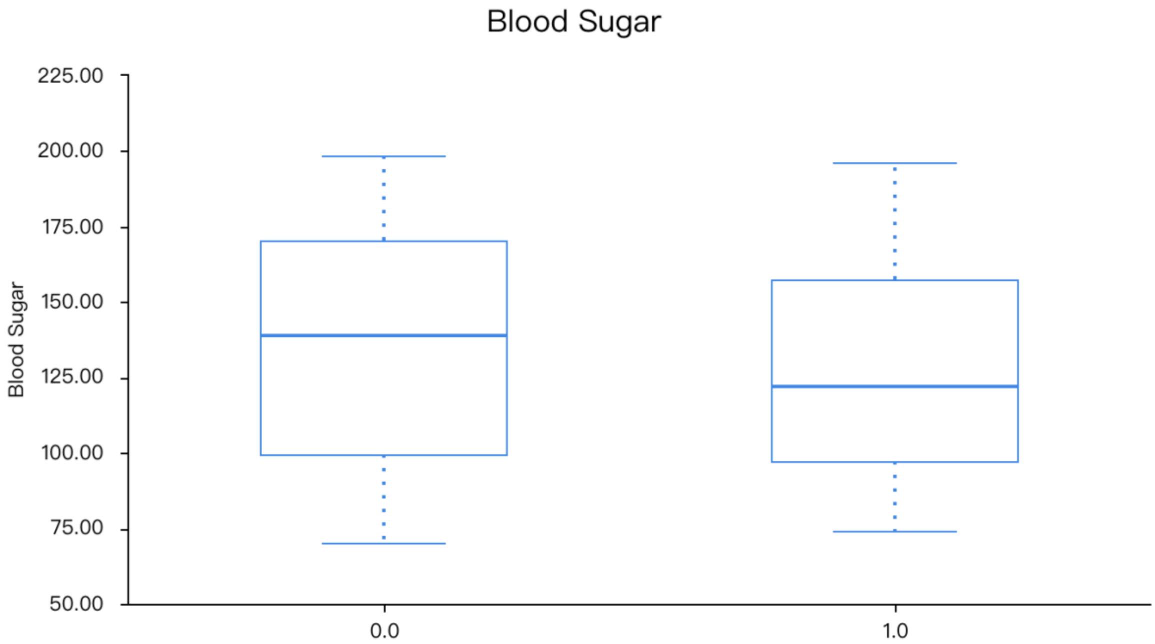


Figure 1. Box plot of blood glucose factors and heart disease causes

Figure 1 is a boxplot showing the relationship between blood glucose factors and heart disease. This data shows the maximum, minimum, median, and other values related to blood glucose factors and heart disease, such as a maximum value of 200.00 and a minimum value of approximately 75.00. By using this data, the author can determine the impact of blood sugar as a factor on the development of heart disease.

DOI: 10.54254/2753-8818/52/2024CH0152

Figure 2. The correlation between heart disease and various factors

The correlation coefficient between factors affecting heart disease and heart disease can be obtained from figure 2. From research data, it has been found that factors such as gender, exercise-induced angina, blood glucose, alcohol intake, exercise duration, and blood pressure are positively correlated with heart disease. Other factors are negatively related to heart disease, such as grade, family history, diabetes, stress level, etc. Among them, smoking is not related to heart disease. In summary, all factors can affect whether one develops heart disease, except for smoking.

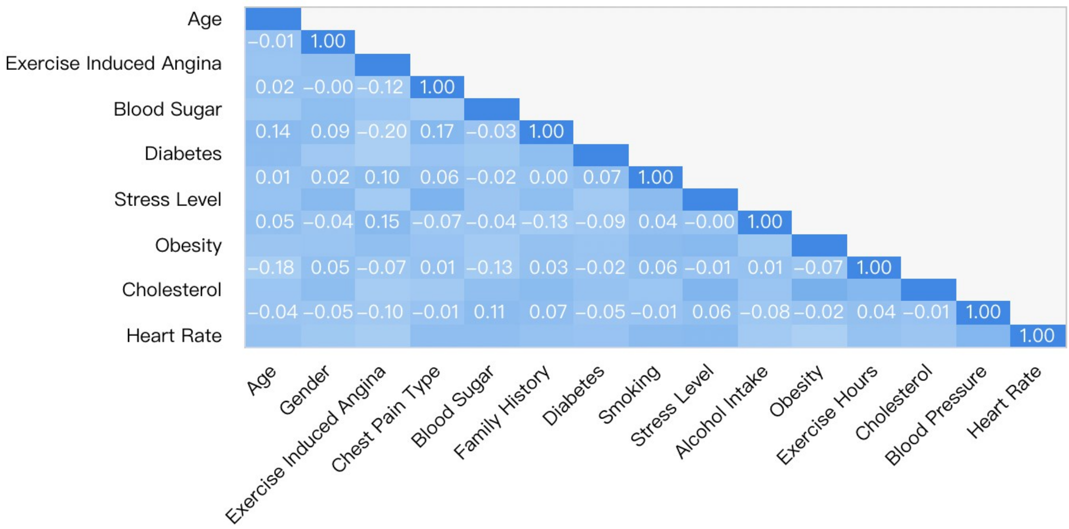


Figure 3. The correlation between various factors

In figure 3, the data shows the correlation between various factors. From this, the degree of influence between a certain factor and another factor can be determined, such as the negative correlation between age and heart rate, and the positive correlation between obesity, etc., from which the correlation between different factors can be obtained.

3.2. Logistic Model Results

From table 2, it can be seen that age and cholesterol have a significant positive impact on heart disease, while other factors do not have an impact on heart disease. From the chart, it can be seen that the

DOI: 10.54254/2753-8818/52/2024CH0152

regression coefficient of age is 0.212 and shows a significance level of 0.01, indicating that age has a significant positive impact on heart disease. Its odds ratio is 1.236, which means that when age increases by one unit, the change in heart disease is 1.236 times. The regression coefficient value of cholesterol is 0.037, and it shows significance at the 0.01 level, indicating a positive relationship between cholesterol and heart disease. Its odds ratio is 1.038, which means that when cholesterol increases by one unit, the change in heart disease is 1.038 times.

Table 2. Binary logistic regression model of heart disease and various factors

Items	Coefficient	S.E.	z value	Wald	p
Gender	0.200	0.222	0.900	0.810	0.368
Exercise Hours	0.013	0.038	0.343	0.118	0.732
Family history	0.247	0.223	1.109	1.230	0.267
Exercise Induced Angina	0.090	0.226	0.401	0.161	0.688
Chest Pain Type	0.056	0.102	0.551	0.304	0.582
Blood Presure	0.000	0.004	0.021	0.000	0.983
Age	0.212	0.014	14.728	216.901	0.000
Cholesterel	0.037	0.003	12.402	153.809	0.000
Diabetes	0.238	0.225	1.057	1.116	0.291
Blood Sugar	0.001	0.003	0.185	0.034	0.853
Smoking	-0.019	0.135	-0.144	0.021	0.885
Obsesity	-0.058	0.222	-0.261	0.068	.0794
Alcohol Intake	-0.073	0.137	-0.535	0.286	0.593
Heart Rate	0.004	0.010	0.405	0.164	0.685
Stress Level	0.004	0.040	0.105	0.011	0.916
Constant	-23.092	2.109	-10.951	119.931	0.000

4. Conclusion

By analyzing and studying various data and using multiple methods (such as studying the relationship between dependent and independent variables, the relationship between various dependent variables, etc.), as well as inferring the correlation between various data, the results of the topic that needs to be studied were finally obtained: the influence of different factors that cause diseases on the development of heart disease.

It must be noted that due to the large amount of data on factors that can cause heart disease and the limited sample size of the population that can be investigated, and the fact that the sample data I studied cannot cover all disease factors and different types of populations, there is a certain degree of bias and low accuracy in the experimental conclusions. But it must be acknowledged that research on the impact of different factors that trigger heart disease on the development of heart disease is a very valuable study. Firstly, through the icons and data presented in this study, the degree of correlation between various factors and heart disease can be intuitively perceived, allowing for a clear understanding of the probability of different causes leading to heart disease. This enables the study to present its findings in a straightforward manner. The second point is that this study used various methods, such as box plots, correlation data, etc., which can use the presentation of data to present the research results more clearly and intuitively. Third, this experimental result can help people understand knowledge about heart disease, such as how the onset of heart disease is related to various causes (cholesterol, blood sugar, blood pressure, family history, etc.). This result can provide assistance in medical treatment for heart disease and also help people who have not suffered from heart disease understand knowledge about heart disease, and take preventive and coping measures. Finally, the data from this experiment can assist the latter in referencing and analyzing relevant experiments, contributing to future experiments on the topic of heart disease and advancing the development of disease-related reports and research.

DOI: 10.54254/2753-8818/52/2024CH0152

References

[1] Zhao, M.R. and Yu, C.C. (2014) Tittle: The trend of ischemic heart disease burden attributed to high salt foods in China. School of Public Health, Wuhan University, Hubei, Wuhan.
[2] Li, K.L. and Jiang, L.L. (2020) Research report on incidence rate of congenital heart disease at home and abroad. Yan'an Hospital affiliated to Kunming Medical University, Yunnan, Kunming.
[3] Huang, L.L. and Huang, X.D. (2024) Analysis of influencing factors of heart disease based on XGBoost and SHAP. Minnan Institute of Technology.
[4] Wang, Z.H., et al. (2019) Screening of key genes for diagnosis of Parkinson's disease based on bioinformatics and machine learning algorithm. Neurological Diseases and Mental Health, 23(12), 837-847.
[5] Liu, C.R. (2024) Parkinson's disease with high disease burden will have the first rehabilitation "group standard". China Youth Daily.
[6] Wen, P.C. (2023) Research on intelligent diagnosis model of Parkinson's disease based on speech signal. Hubei University for Nationalities, China.
[7] Tuncer, T., Dogan, S. and Acharya, U.R. (2019) Automated detection of Parkinson's disease using minimum average maximum tree and singular value decomposition method with vowels. Biocybernetics and Biomedical Engineering, 40(1), 211-220.
[8] Lin, Z., et al. (2019) Progress in the treatment of bladder dysfunction in Parkinson's disease. Nerve Injury and Functional Reconstruction, 19(01), 51-54.
[9] Xu, B.L., et al. (2019) Influence factors of Parkinson's disease combined with depression and changes of glucose metabolism in brain. Journal of Clinical Neurology, 35(05), 324-328.
[10] Liu, A.Y. (2023) Research and system implementation of speech detection algorithm for Parkinson's disease. Shandong Agricultural University.

DOI: 10.54254/2753-8818/52/2024CH0137

Public concerns about terrorist threats amid Vietnam's high-tech boom

Kexin Lan

Department of Security and Crime Science, University College London, London, WC1E 6AE, UK

zctzkla@ucl.ac.uk

Abstract. This study analyzed 1200 samples from the 2020 World Values Survey in Vietnam. Ordinal logistic regression and multinomial logistic regression were used to explore the impact of seven sociodemographic variables on concerns about terrorist attacks, including sex, age, education level, occupation group, social class, income level and religion. Additionally, the Spearman rank correlation test was employed to examine whether there is a significant relationship between concerns about terrorist attacks and perceptions of technological development. The results indicated that different gender, educational level, occupation, income, and religious beliefs significantly influence concerns about terrorism. Furthermore, there is a significant correlation between those who do not believe that technological development will lead to a better life and heightened concerns about terrorist attacks. Thus, when implementing policies to reduce the insecurity of Vietnam citizens, in addition to taking into account relevant demographic factors, it would be advisable to consider incorporating strategies that enhance positive perceptions of high technology.

Keywords: Terrorist attack, technology development, insecurity, ordinal logistic regression.

1. Introduction

Terrorism is the use of violence to create fear (terror, mental fear) for political, religious, or ideological reasons. It deliberately targets non-combatant targets to gain maximum publicity for a group, cause, or individual [1]. Klaus Schwab, founder and chairman of the World Economic Forum, has stated that scientific breakthroughs and new technologies driving the Fourth Industrial Revolution could be misused for malicious purposes, capable of causing large-scale harm. This latest vulnerability will lead to new types of terrorist attacks [2].

Since implementing economic reforms in 1986, Vietnam has achieved sustained economic growth rates. A study revealed that the growth rates were 6.8% in 2017, 7.1% in 2018, and 7% in 2019. Despite the substantial effects of the COVID-19 pandemic in 2020, the growth rate of the country still managed to stay between 2% and 2.5% [3]. This brisk economic expansion has positively influenced numerous areas, notably in the fields of science and technology research. Studies showed that Vietnam's science and technology have grown astonishingly [4].

However, the growth of high technology is a double-edged sword. A 2019 study indicated that the advanced information technology (IT: internet, media, etc.) in recent years has provided terrorists with unprecedented opportunities for mass self-propagation [5]. Additionally, in the field of emerging technologies, there are clear indications that drones, the dark web, and malware are being exploited, heightening the threat of chemical, biological, radiological, and nuclear (CBRN) terrorism. Additionally, synthetic biology and 3D printing present long-term dangers [6]. Take malware as an example, Vietnam has already dealt with cyberspace terrorism attacks from both internal and external territories [7].

The increase in various terrorist attacks due to technological advancements naturally raises public concerns. The apprehension regarding terrorist attacks is linked to social trust [8]. Social trust plays a vital role in the smooth and effective operation of society. In nations where trust is deficient, economic growth can decelerate, political structures may be unstable, and overall societal welfare might suffer. Conversely, countries with high levels of social trust or "general" trust typically experience faster economic growth, more stable political and institutional systems, less corruption and conflict, and higher levels of individual life satisfaction, better health, and greater happiness [9]. Reducing public fear of terrorist attacks is always essential.

Many studies have analyzed the fear of terrorism and demographic characteristics, but most have focused on Europe and the United States. A recent Swedish study found that women are more concerned about terrorist attacks than men [10]. A 2018 U.S. study using Bivariate analyses found that respondents who used social media more and older respondents had higher levels of fear of new types of cyber terrorism. This study included 13 demographic independent variables, with no significant findings for the other variables [11]. Additionally, a study based on Kent State University fitted a multiple linear regression model on personal factors such as age, gender, racial status, religious views, political inclination, and economic status and their fear of terrorist attacks. The results showed that, in addition to the mentioned age and gender-related factors, Catholics' fear of terrorist attacks is particularly significant [12]. Due to different political backgrounds and social developments, the significant concerns of different demographic characteristics also vary. The rapidly developing Southeast Asian country of Vietnam has not yet been the focus of attention. Most current studies focused more on the growth of high-tech terrorism. Therefore, the first question this paper explores is, in Vietnam, which demographic groups are particularly concerned about terrorist attacks. The aim is to provide insights for the Vietnamese government to formulate more precise prevention policies. Additionally, since the emergence of new types of terrorist attacks in Vietnam has occurred in the context of rapid technological development, is the level of concern about terrorist attacks related to views on technological development. The author predicts that those more concerned about terrorist attacks may not be optimistic about technological development. This is the second question this paper explores. If the hypothesis is confirmed, it can provide a new perspective for government prevention policies, such as improving citizens' positive views on technological development to reduce the fear of terrorist attacks.

In summary, this paper aims to provide Vietnamese policymakers with more precise directions for formulating policies to reduce the fear of terrorism and to try to achieve better results through indirect guidance, especially in the context of rapid technological development.

2. Methodology

2.1. Data source

The World Values Survey (WVS) collects public opinion on various social and political issues in different countries and regions. The survey covers a wide range of topics such as religion, economy, education, development, and security [13, 14]. This study uses data from the 2020 Vietnam WVS (1200 samples). All variables with missing data (NA values) were removed, retaining 1181 usable samples for this study. The original dataset remained in .csv format.

2.2. Variable selection

Previous research has explored various causes of terrorism, including social, economic, educational, and religious reasons [15-17]. Considering these influencing factors, this study incorporates the following demographic variables of Table 1.

Table 1. Demographic characteristic variables.

Variables	Form
Sex	(1): Male (2): Female
Age	(1): 16-29 years (2): 30-49 years (3): Over 50
Education level	(0): ISCED 0 (1): ISCED 1 (2): ISCED 2 (3): ISCED 3 (4): ISCED 4 (5): ISCED 5 (6): ISCED 6 (7): ISCED 7
Occupation group	(0): Never had a job (1): Professional and technical (2): Higher Administrative (3): Clerical (4): Sales (5): Service (6): Skilled worker (7): Semi-skilled worker (8): Unskilled worker (9): Farmworker (10): Farm owner
Social class	(1): Lower class (2): Working class (3): Lower middle class (4): Upper middle class (5): Upper class
Income level	Level from 1 to 10
Religion	(1): A religious person (2): Not a religious person (3): An atheist

The variable of "Worries of terrorist attack", see as Figure 1, demonstrate the level of worry about the occurrence of a terrorist attack (1-Not at all, 2-Not much, 3-A good deal, 4-Very much). More than half of them (58.3%) are very much worried about the occurrence of terrorist attacks.

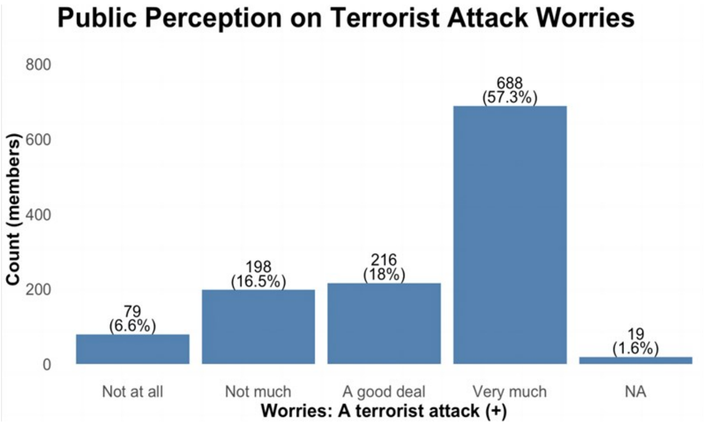


Figure 1. The worries of terrorist attack.

For Figure 2, "Views on technological developments" variable is a discrete variable from 1-10, with 1 being Completely disagree and 10 being Completely agree. 32.8% of the respondents agreed that technological developments have brought about a better life. And the majority of people's opinion is favourable to the development of technology (choose a number greater than 5).

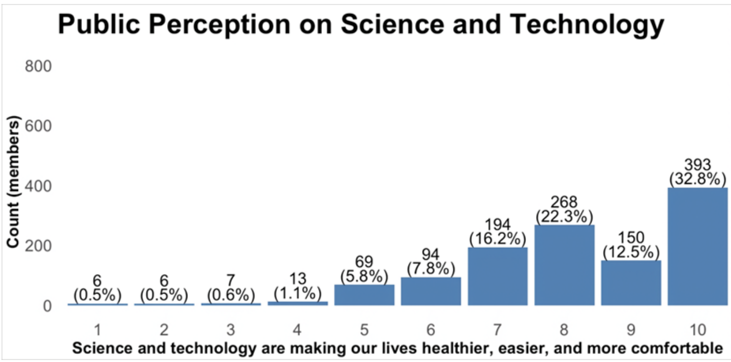


Figure 2. The views on technological developments.

DOI: 10.54254/2753-8818/52/2024CH0137

2.3. Method introduction

Ordinal logistic regression was applied to predict an ordinal dependent variable based on one or more independent variables, which could be continuous, categorical, or ordinal. Each independent variable has a coefficient β that indicates how these factors influence the dependent variable. Positive coefficients indicate that an increase in the independent variable correlates with a likely increase in the dependent variable's categories. Conversely, negative coefficients suggest that as the independent variable rises, the dependent variable is inclined to decline into lower categories. In this study, fear of terrorist attacks served as the dependent variable and different demographic characteristics served as the independent variables.

Multinomial logistic regression was used to predict a nominal dependent variable from one or more independent variables, which can also be continuous, categorical, or ordinal. A nominal variable usually has three or more categories. The same variables were applied to both multinomial logistic regression, where the dependent variable is considered nominal, and ordinal logistic regression, where the dependent variable is treated as ordinal. This approach was used to evaluate and select the more appropriate model.

Spearman's Rank correlation coefficient was utilized to examine the relationship between views on technology development and worries about terrorist attacks (Figure 3). If the P-value is less than 0.05, it can reject the Null hypothesis and prove that their relationship is statistically significant. This method assesses the strength and direction (negative or positive) of the relationship between two ordinal variables, with results always ranging between 1 and -1.

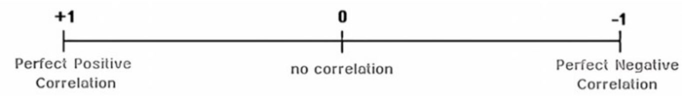


Figure 3. The 0-1 nature of Spearman's rank correlation.

3. Results and discussion

3.1. Correlation analysis

Creating pairwise variable relationship plots (Figure 4) before fitting the model helps in understanding the relationships and characteristics of the independent variables better. This includes examining correlation coefficients, assessing the significance of relationships, determining whether relationships are linear or nonlinear, and understanding the distribution of the variables. Since such relationship plots are not suitable for categorical nominal variables, the paper excluded the variables for Occupation group and Religions from this analysis.

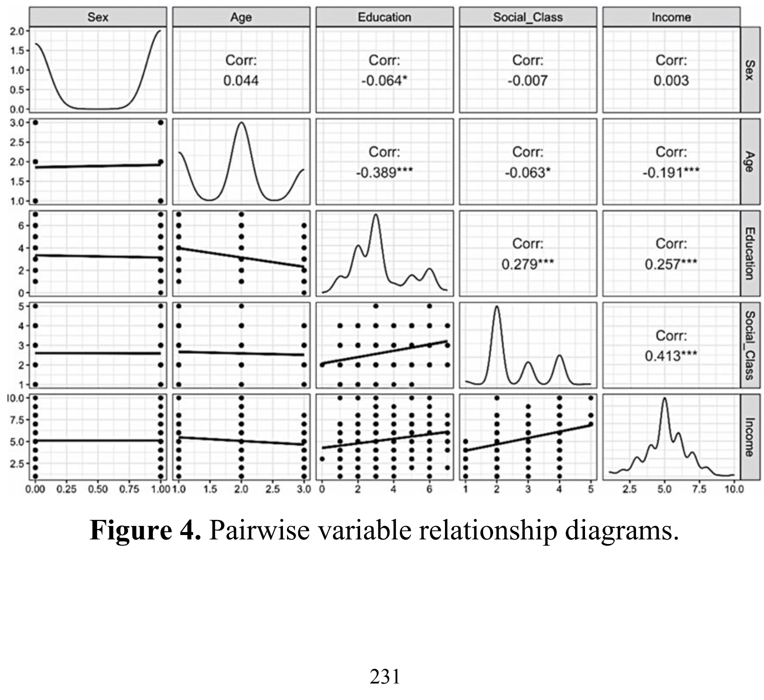


Figure 4. Pairwise variable relationship diagrams.

Focus on the correlation coefficients in the upper right corner of Figure 4. Among the relatively strong correlations are a negative correlation between Age and Education (-0.389), and a positive correlation between Income and Social Class (0.413). While many variables show significant relationships, the correlations are not particularly high (not close to 1 or -1). Therefore, all variables will be retained for model fitting at this stage. After fitting the model, check the VIF to accurately assess multicollinearity. Remove the variables with excessively high VIF values and then refit the model.

3.2. Model comparison

From Table 2, the results of the two models showed differences in the significance of various variables. For instance, Model 1 indicates that gender differences are significant in explaining the fear of terrorist attacks, while Model 2 finds this relationship insignificant. Another notable difference is that Model 2 specifies the levels at which variables are significant. For example, both models show that the Clerical occupation group is substantial. However, Model 2 highlights that Clerical is particularly important at the 'Not much' level (compared to the baseline 'Not at all').

Given these differences, it is crucial to identify a better model to explain the results accurately. Comparing the models, Model 1 and Model 2 have the same Akaike information criterion (AIC). However, Model 1 had a lower Bayesian Information Criterion (BIC) (2671) than Model 2 (2853), indicating a better fit. Moreover, the issue of multicollinearity should be considered. Ensure that multicollinearity does not adversely affect the accuracy of the model by assessing the variance inflation factor (VIF) value.

Since the variables included categorical ones, VIF was used to detect multicollinearity issues. A VIF value greater than 10 is considered to indicate severe multicollinearity, while a value less than 5 suggests that multicollinearity is not a significant problem. As shown in Table 3, the VIF values for Model 1 are all around 1, indicating no variables need to be removed. In contrast, all variables in Model 2 exhibit severe multicollinearity, with values far exceeding 10.

Overall, it is obvious that Model 1 is superior to Model 2. This paper used the more robust ordinal logistic regression model (Model 1) to interpret the results. To enhance the explanatory power and practical significance of the model, the research also included Odds ratios to help interpret the model, as shown in Table 2. This approach not only indicated the direction of the variables' impact but also showed the magnitude of the effect.

Table 2. Results of logistic regression (Model 1 & Model 2).

Variables	Model 1 (Ordinal)		Model 2 (Multinomial)		
	Estimate	Odds ratio	Not much	A good deal	Very much
Sex	0.394 [0.001]	1.483 [1.173, 1.877]	-0.085 [0.924]	0.428 [0.652]	0.473 [0.431]
Age (16-29) 30-49	0.013 [0.905]	1.014 [0.769, 1.333]	-0.121 [0.879]	0.028 [0.973]	-0.015 [0.979]
50 and over	-0.215 [0.249]	0.807 [0.558, 1.167]	-0.289 [0.311]	-0.710 [0.381]	-0.486 [0.335]
Educational level	-0.114 [0.029]	0.892 [0.805, 0.989]	-0.233 [0.410]	-0.265 [0.732]	-0.290 [0.628]
Occupation (No job) Professional	-0.257 [0.454]	0.774 [0.391, 1.544]	0.567 [0.028]	-1.409 [0.018]	-0.361 [0.000]
Higher administrative	-1.288 [0.003]	0.276 [0.116, 0.658]	1.151 [0.001]	-0.524 [0.346]	-1.141 [0.057]
Clerical	-0.834 [0.007]	0.434 [0.235, 0.801]	1.580 [0.000]	0.674 [0.181]	-0.106 [0.584]
Sales	-0.236 [0.309]	0.789 [0.494, 1.247]	0.607 [0.053]	-0.407 [0.245]	-0.144 [0.459]

DOI: 10.54254/2753-8818/52/2024CH0137

Table 2. (continued).

Service	-0.554 [0.131]	0.575 [0.280, 1.201]	19.147 [0.000]	17.461 [0.000]	17.573 [0.000]
Skilled worker	-0.427 [0.082]	0.652 [0.400, 1.054]	0.199 [0.645]	-0.168 [0.547]	-0.433 [0.000]
Semi-skilled worker	-0.030 [0.902]	0.970 [0.538, 1.761]	1.484 [0.000]	0.979 [0.112]	0.974 [0.000]
Unskilled worker	-0.256 [0.345]	0.774 [0.451, 1.325]	0.698 [0.000]	0.609 [0.279]	0.196 [0.033]
Farm worker	-0.820 [0.001]	0.441 [0.267, 0.719]	0.034 [0.775]	-0.521 [0.312]	-1.041 [0.007]
Farm owner/ Farm manager	0.369 [0.728]	1.446 [0.219, 28.63]	10.854 [0.000]	-10.346 [0.000]	9.880 [0.000]
Social status	0.063 [0.403]	1.065 [0.916, 1.240]	0.277 [0.714]	0.118 [0.895]	0.242 [0.491]
Income level	0.092 [0.039]	1.096 [1.004, 1.196]	0.207 [0.805]	0.317 [0.707]	0.283 [0.574][
Religion (A religious person) Not a religions	-0.023 [0.882]	0.977 [0.679, 1.397]	-0.038 [0.954]	0.469 [0.524]	0.034 [0.947]
An atheist	0.603 [0.003]	1.828 [1.221, 2.726]	0.335 [0.732]	1.287 [0.057]	1.196 [0.009]
AIC	2564		2564		
BIC	2671		2853		

Table 3. VIF of Models.

Variables	Model 1	Model 2
Sex	1.034	130032.628
Age	1.085	1742.618
Education level	1.357	NaN
Occupation	1.037	359.794
Social status	1.141	253013.423
Income level	1.140	88779.553
Religion	1.018	1964.836

3.3. Discussion

According to Model 1, controlled for other variables, males are 48.3% less likely than females to worry about terrorist attacks, which is statistically significant at the 0.1% level. Women exhibited a higher fear of crime despite statistically being less likely to become victims compared to men-a phenomenon known as the "fear-victimization paradox" [18]. A similar paradox exists for fear of terrorism, and this result is consistent with that theory [19].

Furthermore, each additional unit of education increases the likelihood of concern about terrorist attacks by 10.8%, which is significant at the 5% level. Higher educational attainment might correlate with a greater awareness of potential threats. A cross-national study found that in countries with the lowest levels of education, an increase in years of schooling is significantly associated with the intensity of terrorist attacks [20]. This result may suggest that more educated individuals are more sensitive or more aware of terrorism-related issues.

Among occupational categories, three occupations are statistically significant at the 1% level. Those in Higher administrative positions are 72.4% more likely to worry about terrorist attacks compared to the unemployed, Clerical workers are 56.6% more likely, and Farm workers are 55.9% more likely.

DOI: 10.54254/2753-8818/52/2024CH0137

People in Higher administrative roles usually have higher socioeconomic status and more access to information, which might make them more aware of potential threats and consequences of terrorist attacks. Clerical workers and Farm workers, although lower in socioeconomic status compared to administrative personnel, may have specific sensitivities to social events due to the nature of their work. For instance, clerical workers might be more concerned about social security issues due to their frequent interaction with information, while farm workers might be more directly concerned about specific threats like bioterrorism, which targets agricultural and food supply systems. Agriterrorism, for example, has not been well-prevented in Vietnam, particularly during wartime [21].

Income level is also significant at the 5% level, with each additional unit of income decreasing the likelihood of concern about terrorist attacks by 9.6%. This aligns with the finding that lower-income individuals tend to have a lower sense of security [22].

Finally, atheists are 82.8% less likely than religious individuals to worry about terrorist attacks, which is significant at the 1% level. However, some scholars argued in 2015 that religion might have a protective effect on how people cope with terrorism in their daily lives [23]. Contrarily, a study in 2015 found that the more religious the respondents were, the more likely they were to express concerns about terrorism [24]. Additionally, a 2019 study in predominantly Christian countries found that non-religious or self-identified atheists were less worried about terrorist attacks compared to those who considered themselves religious, partially supporting the findings of this study [25].

For the second theme of the paper, regarding the relationship between the fear of terrorist attacks and views on technological development, the Spearman correlation coefficient is 0.103 with a p-value of 0.000414 (<0.05). Although this coefficient indicates a positive correlation, it is not strong (far less than 1). However, the null hypothesis still could be rejected, confirming a significant relationship between negative views on scientific progress and heightened concern about terrorism. As shown in Figure 5, people with more positive views on technological development (closer to 10 on the Y-axis) are slightly more likely to be unconcerned about the occurrence of terrorist attacks (closer to 4 on the X-axis).

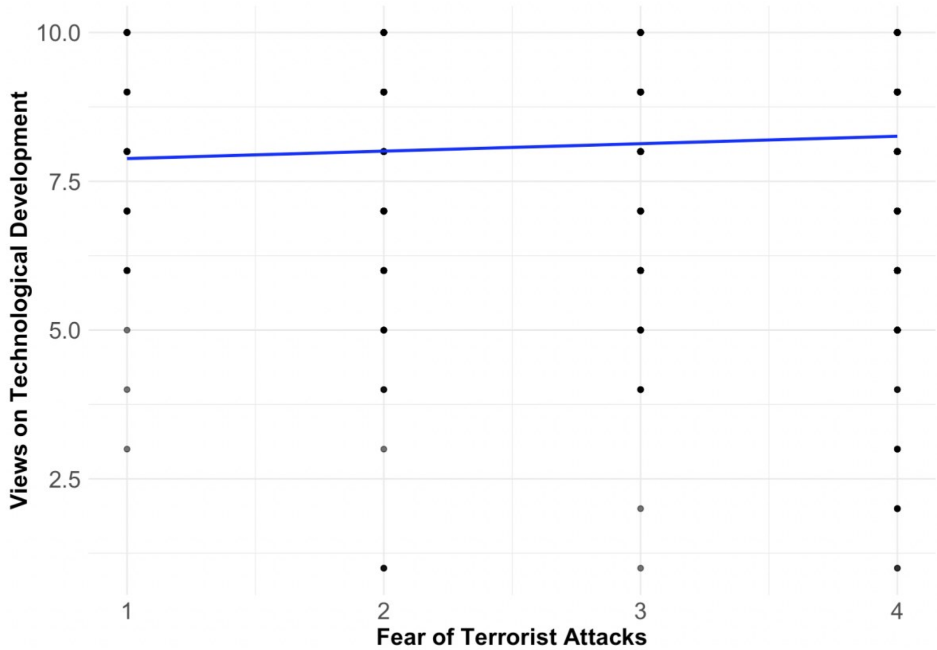


Figure 5. Views on technology development and Fear of Terrorist Attacks relationship diagrams.

Therefore, Vietnam's policy of helping people eliminate their fear of terrorist attacks can be accompanied by a policy of raising the public's positive perception of science and technology, which has the potential to enhance citizens' sense of security. This is particularly important in the context of rapid technological development.

DOI: 10.54254/2753-8818/52/2024CH0137

4. Conclusion

This study explored various social and individual factors that provided insights into mitigating fears of terrorist attacks through targeted social policies and public education for different groups. It also highlighted the correlation between rapid technological advancement and public concerns about terrorism and provided new perspectives for policy development. However, this study has limitations and areas for improvement, as the current analysis focuses more on descriptive and correlational aspects without establishing causality. In addition, the study is only based on 2020 data and some uncertainties, such as socio-economic and lifestyle changes brought about by COVID-19, may affect the results. Future research with a longitudinal design and implementation of targeted policy experiments would help explore trends in technology-related terrorism issues and test the effectiveness of policies. In a country like Vietnam, which is rapidly developing and dependent on technology for its development, understanding the root causes of public fear is crucial to building a safer and more harmonious society.

References

[1] Schmid A P 2011 The definition of terrorism. London: Routledge, 39-157.
[2] Schwab K 2017 The Fourth Industrial Revolution (New York: Crown Business) pp 104–105
[3] Thuy N 2020 Vietnam gov't targets 2.5% GDP growth for 2020. Hanoi Times.
[4] Vuong Q H, et al. 2021 Mirror, mirror on the wall: Is economics the fairest of them all? An investigation into the social sciences and humanities in Vietnam. Research Evaluation, 57-72.
[5] Mahmood R and Jetter M 2020 Communications Technology and Terrorism. Conflict Resolution, 64, 127-166.
[6] Koblentz G D 2020 Emerging Technologies and the Future of CBRN Terrorism. The Washington Quarterly, 43, 177-196.
[7] Dai C T 2017 Cybersecurity governance framework in Vietnam: State of play, progress and future prospects. Asian Research Policy, 8, 86-97.
[8] Godefroidt A and Langer A 2018 How Fear Drives Us Apart: Explaining the Relationship between Terrorism and Social Trust. Terrorism and Political Violence, 32, 1482-1505.
[9] Jan D and Georgi D 2015 Happier Together: Social Cohesion and Subjective Well-Being in Europe. International Journal of Psychology, 51, 163-176.
[10] Ekholm S 2020 Swedish Mothers' and Fathers' Worries about Climate Change: A Gendered Story. Journal of Risk Research, 23, 288-296.
[11] Onat I, Bastug M F, Guler A and Kula S 2022 Fears of cyberterrorism, terrorism, and terrorist attacks: an empirical comparison. Behavioral Sciences of Terrorism and Political Aggression, 16, 149-165
[12] Elmas M S 2021 Perceived risk of terrorism, indirect victimization, and individual-level determinants of fear of terrorism. Security Journal, 34, 498-524.
[13] Loll D and Hall K S 2018 Differences in abortion attitudes by policy context and between men and women in the World Values Survey. Women & Health, 59, 465-480.
[14] Mink C 2015 It's about the group, not God: Social causes and cures for terrorism. Journal for Deradicalization, 63-91.
[15] Caruso R and Locatelli A 2014 Understanding Terrorism: A Socio-Economic Perspective. Bingley: Emerald Group Publishing.
[16] Kluch S P and Vaux A 2017 Culture and Terrorism: The Role of Cultural Factors in Worldwide Terrorism (1970–2013). Terrorism and Political Violence, 29, 323-341.
[17] Elmas M S 2021 Perceived risk of terrorism, indirect victimization, and individual-level determinants of fear of terrorism (Security Journal vol 34) pp 498–524
[18] Ferraro K F 1996 Women's Fear of Victimization: Shadow of Sexual Assault? Social Forces, 75, 667–690.
[19] Shechory-Bitton M and Cohen-Louck K 2018 Does Fear of Terrorism Differ from Fear of Crime and Sexual Assault: A Question of Geographical Location and Residential Area. International Journal of Offender Therapy and Comparative Criminology, 62, 806-826.

[20] Korotayev A, Vaskin I and Tsirel S 2019 Economic Growth, Education, and Terrorism: A Re-Analysis. Terrorism and Political Violence, 33, 572-595.
[21] Mukherjee T 2021 Agroterrorism: A Less Discussed Yet Potential Threat to Agronomy. Science and Culture.
[22] Hecht K, Burchardt T and Davis A 2024 Richness, insecurity and the welfare state. Journal of Social Policy, 53, 573-594.
[23] Kupor D M, Laurin K and Levav J 2015 Anticipating divine protection? Reminders of God can increase nonmoral risk taking. Psychological Science, 26, 374-384.
[24] Adamczyk A and LaFree G 2015 Religiosity and reactions to terrorism. Social Science Research, 51, 17-29.
[25] Leite Â, Ramires A, Dinis M A P and Sousa H F P 2019 Who Is Concerned about Terrorist Attacks? A Religious Profile. Social Sciences, 8, 316.

DOI: 10.54254/2753-8818/52/2024CH0151

Study on the Influencing Factors of Adolescent Mental Health

Aoyu Wang[1], Yangyang Xue [2,3]

[1]Shanghai Ulink Bilingual School, No. 559 Lai Tingnai Road, Shanghai, 200000, China
[2]Nanjing University of Information Science and Technology, Nanjing, 210000, China

[3]bx805033@student.reading.ac.uk

Abstract. Adolescent mental health is a popular topic of general concern to the global public, and there are many researchers already finding out various factors that lead to psychological problems represented by depression. However, usually one study only focuses on the analysis of one factor, causing it being lack in a comprehensive consideration of the personal life experience of adolescents. Therefore, this paper collects the questionnaire dataset on Kaggle website and conducts correlation and stepwise regression analysis to explore the factors leading to adolescent depression through the score evaluation on the questionnaire. The results show that typical factors such as teacher-student relationship, academic performance, and safety have a negative impact on adolescent depression, while factors like study load, extracurricular activities, noise level have a positive impact on adolescent depression. Based on this conclusion, this paper suggests that schools should pay more attention to adolescent mental health to ensure that they thrive.

Keywords: Adolescent depression, study load, teacher-student relationship, mental health, regression analysis.

1. Introduction

Recently, there has been growing awareness about adolescent mental health, with a specific emphasis on depression due to its widespread occurrence and profound effects on students' academic achievements and overall well-being. Depression in adolescence is often influenced by a complex interplay of social, environmental, and personal factors, making it a critical area of study [1]. This research aims to explore these factors through a detailed analysis of their correlation with depression among high school students.

Adolescence is a crucial stage marked by significant physical, emotional, and social transformations, which increase the likelihood of experiencing mental health challenges, including depression [2]. The World Health Organization reports that depression ranks among the primary causes of illness and disability worldwide for adolescents [3]. Early identification and intervention are crucial, as untreated depression during adolescence can lead to severe consequences, including academic decline, substance abuse, and even suicidal behavior [4].

A significant body of research has highlighted the detrimental impact of bullying on adolescent mental health. Victims of bullying are at a higher risk of developing depression due to the chronic stress and emotional trauma inflicted by repeated bullying incidents [5]. Studies have shown that both

DOI: 10.54254/2753-8818/52/2024CH0151

traditional bullying and cyberbullying significantly contribute to depressive symptoms among adolescents [6]. Furthermore, the relationship between bullying and depression is bidirectional; while bullying can lead to depression, adolescents with depressive symptoms are also more likely to become targets of bullying [7].

Family environment is another critical factor influencing adolescent depression. Dysfunctional family dynamics, including parental conflict, lack of emotional support, and negative parenting styles, have been associated with higher levels of depression in adolescents [8]. Conversely, a supportive family environment characterized by open communication, emotional warmth, and positive parental involvement can act as a protective factor against depression [9].

The school environment also plays a pivotal role in adolescent mental health. Academic pressure, teacher-student relationships, and the overall school climate can either exacerbate or mitigate depressive symptoms [10]. Schools that foster a positive and inclusive atmosphere, where students feel safe and supported, are likely to have lower incidences of depression among their student population [11].

Previous studies have utilized various methodologies to investigate the factors contributing to adolescent depression. For instance, structural equation modeling has been employed to examine the complex interactions between different variables such as family environment, peer relationships, and individual psychological traits [12]. Longitudinal studies have also provided valuable insights into how these factors evolve over time and their long-term impact on mental health [13].

In this study, the author employed a multiple regression analysis to identify the key predictors of depression among high school students. By focusing solely on depression as the dependent variable, this research aims to provide a clearer understanding of the most significant factors contributing to adolescent depression. The study's outcomes provide valuable insights for crafting specialized interventions and preventive measures tailored to the mental health requirements of adolescents.

2. Methods

2.1. Data Source and Description

The data for this research were obtained from the Kaggle platform, comprising 14 distinct variables: environmental noise, sleep quality, respiratory health, living conditions, safety concerns, fulfillment of basic needs, academic outcomes, study load, dynamics of teacher-student interactions, career-related anxieties, levels of social support, peer pressure, participation in extracurricular activities, and experiences with bullying. The adolescents made score evaluation of the factors, depression and anxiety. The total point of each factor's score evaluation was 5 and total point of depression and anxiety score is 30. According to Kaggle, all the data comes directly from online questionnaires.

2.2. Indicator Selection

For the above data, the paper removed missing and outliers from the collected data and carried out analysis. According to the preprocessed dataset, it included one dependent variables, which is depression, as well as 8 independent variables (chosen from totally 14 factors). The research set the independent variables as X1-8. The basic overview of quantitative variable is shown in Table 1.

Table 1. Independent variables

Symbol	Meaning	Range
X1	Teacher student relationship	1~5
X2	Extracurricular activities	1~5
X3	Future career concerns	1~5
X4	Peer pressure	1~5
X5	Sleep quality	1~5
X6	Study load	1~5
X7	Headache	1~5
X8	Safety	1~5

DOI: 10.54254/2753-8818/52/2024CH0151

2.3. Method Introduction

This paper aims to use linear regression to study the influencing factors of adolescent mental health. The research took factors as the independent variables and depression as the dependent variable, and then it analyses the data. Finally, the research used the data to establish a stepwise regression formula to show the relationship between different influencing factors and the mental health problem, to determine whether the influencing factors can affect mental health or not, and if there is a relationship, to further determine whether the influence is large or small. The formula was:

3. Results and Discussion

3.1. Descriptive Analysis

The box plot analysis for study load (Figure 1) revealed a widespread, indicating significant variability in academic pressure among students. Higher levels of study load were associated with increased depression levels, as evidenced by the broader interquartile range and the presence of several outliers. This suggests that students experiencing higher academic pressure tend to report higher depression levels, highlighting the impact of study load on mental health. Addressing this variability could be crucial in reducing depression rates among students by implementing effective stress management and academic support strategies.

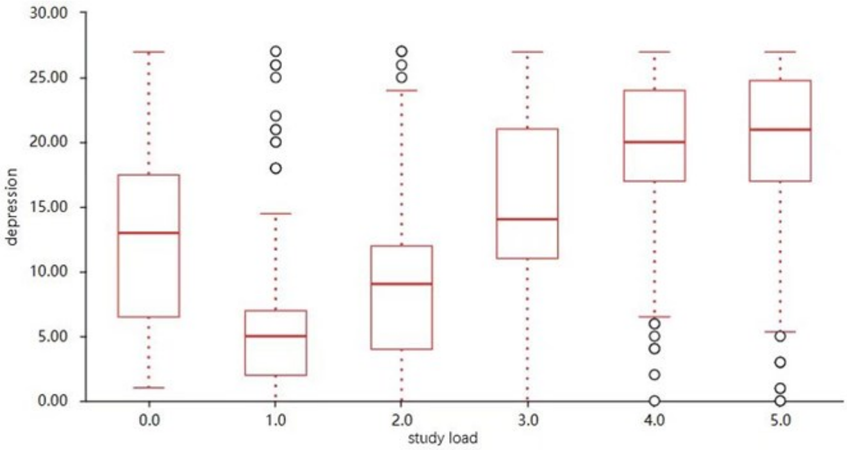


Figure 1. Box Plot of Study Load and Depression.

The box plot (Figure 2) for teacher-student relationships demonstrated a considerable range, showing that teacher-student relationship varies dramatically. Students with better teacher-student relationships reported lower depression levels, indicated by the lower median and narrower interquartile range. This underscores the importance of fostering positive interactions between teachers and students to mitigate depression. Enhancing teacher-student relationships through training programs and supportive school environments could play a vital role in improving student mental health.

DOI: 10.54254/2753-8818/52/2024CH0151

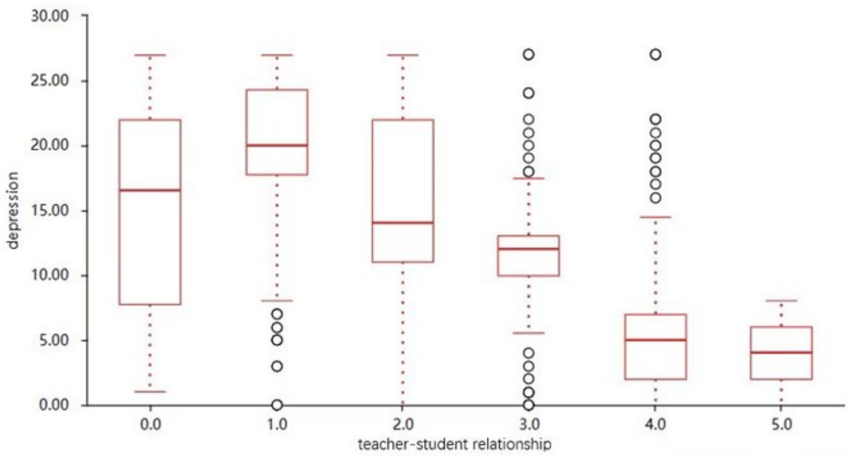


Figure 2. Box Plot of Teacher-Student Relationship and Depression.

3.2. *Correlation Analysis*

The correlation analysis provided further insights into the relationships between the variables (Figure 3). Teacher-student relationships had a strong negative correlation with depression (r = -0.674), indicating that improved relationships with teachers are associated with lower depression levels. Study load showed a strong positive correlation with depression (r = 0.602), suggesting that increased academic pressure correlates with higher depression levels. Bullying and living conditions also exhibited significant correlations, with bullying positively correlated with depression (r = 0.666) and living conditions negatively correlated with depression (r = -0.690). These findings emphasize the multifaceted nature of depression and the need for comprehensive strategies addressing various contributing factors.

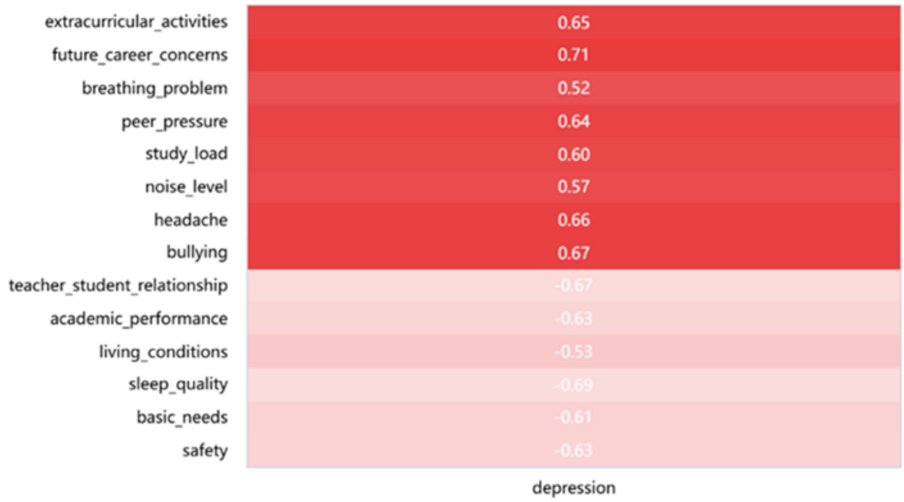


Figure 3. Pearson Correlation Heatmap of Independent Variables and depression.

3.3. *Stepwise Regression Analysis*

The stepwise regression analysis (Table 2) identified the most significant predictors of depression, resulting in the following regression equation:

$$depression = 8.973 - 0.874 * teacher\ student\ relationship\ + \cdots - 0.319 * safety \quad (1)$$

This model explained 65.7% of the variance in depression ($R^2 = 0.657$) and passed the F-test (F = 261.077, $p = 0.000 < 0.05$), indicating the model's overall significance. The VIF values were all below 5, suggesting no multicollinearity issues, and the Durbin-Watson statistic (2.109) confirmed no autocorrelation in the residuals.

DOI: 10.54254/2753-8818/52/2024CH0151

The regression analysis revealed that better teacher-student relationships and improved sleep quality significantly reduced depression levels, while increased study load and future career concerns significantly elevated depression levels. The exclusion of bullying from the stepwise regression model, despite its high correlation with depression, suggests that its unique contribution is less significant when other predictors are considered, possibly due to overlapping effects with variables like peer pressure and teacher-student relationships.

Table 2. Stepwise Regression Analysis Results for Depression.

	Nonnormalized Coefficient		Standard Coefficient	t	p	Collinearity diagnostics	
	B	Std. Error	Beta			VIF	Tolerability
Constant	8.973	1.028	-	8.730	0.000**	-	-
X1	-0.874	0.156	-0.157	5.613	0.000**	2.475	0.404
X2	0.695	0.144	0.127	4.837	0.000**	2.208	0.453
X3	0.805	0.155	0.159	5.186	0.000**	2.998	0.334
X4	0.383	0.146	0.071	2.629	0.009**	2.302	0.434
X5	-0.803	0.146	-0.161	5.513	0.000**	2.71	0.369
X6	0.820	0.138	0.140	5.951	0.000**	1.751	0.571
X7	0.660	0.148	0.120	4.449	0.000**	2.328	0.430
X8	-0.319	0.148	-0.058	-2.16	0.031*	2.299	0.435
R 2	0.657						
Adj R 2	0.654						
F	F (8,1091) =261.077, p=0.000						
D-W	2.109						
Dependent Variable: depression							
* p<0.05 ** p<0.01							

Overall, this comprehensive analysis highlights the significant factors influencing depression among high school students, emphasizing the need for interventions focusing on improving teacher-student relationships, managing academic pressure, and ensuring better sleep quality to mitigate depression. The result of this study gives a strong foundation for researches of this topic in the future and targeted strategies for relieving mental health problems.

3.4. Discussion

According to the data analysis and charts, there is a significant negative relationship between teacher-student relationship and adolescent depression, which means that in general, the better the teacher-student relationship is, the lower the severity of depression is. Conversely, there is a significant positive relationship between study load and adolescent depression, meaning that in most cases, the greater the study load is, the higher the severity of depression is. This paper tries its best to complement the previous research on the relationship between personal experience and depression. It explores the impact of different life experiences on depression in adolescents' study and relationship with others (elders).

4. Conclusion

This study found out that various aspects of adolescent life experiences have an impact on the severity of depression, and typical factors included teacher-student relationship and study load. This is generally because, teenagers, whether day students or boarding students, usually spend much more time in school than at home, having longer contact time with teachers. Therefore, how strict teachers are with adolescent students, or how strict teachers are about the management and control of bad peer relationships such as school bullying, will have an impact on adolescent psychology. Because of longer time in school as well, school study load, such as the amount of homework and the difficulty of exams, is easy to affect the pressure of teenagers. As a result, the greater the study load is, the less time for entertainment and rest is, and students' spirit cannot be relaxed, causing it more likely to give rise to depression or deepen the severity of depression. This research considers as many aspects of the influencing factors of adolescent depression as it can, which is relatively comprehensive, and conducts an in-depth analysis of the more typical factors, which is conducive to more all-sided research on this topic in the future.

Finally, the dataset collected on Kaggle takes the form of scores for the degree of depression and the degree of influencing factors. Strictly speaking, the scores of 1-5 or 1-30 are integers, which is difficult to accurately describe the real situation of adolescents. In the future, more accurate descriptions of adolescent depression and various influencing factors can be found in the dataset collection, so as to achieve a more accurate analysis.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Smith, A. and Jones, B. (2019) Adolescent Mental Health: A Comprehensive Overview. Journal of Adolescent Health, 64(2), 123-135.
[2] Brown, C. and Green, D. (2020) The Impact of Social Changes on Adolescent Mental Health. Youth Studies, 15(3), 245-260.
[3] World Health Organization. (2021) Adolescent Mental Health Fact Sheet. Retrieved from https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health.
[4] Johnson, M. and Thompson, P. (2018) Early Intervention in Adolescent Depression: Strategies and Outcomes. Clinical Child Psychology, 23(1), 89-104.
[5] Kowalski, R.M., Limber, S.P. and Agatston, P.W. (2019) Cyberbullying: Bullying in the Digital Age. Blackwell Publishing.
[6] Tokunaga, R.S. (2018) Following you home from school: A critical review and synthesis of research on cyberbullying victimization. Computers in Human Behavior, 26(3), 277-287.
[7] Vaillancourt, T., Hymel, S. and McDougall, P. (2019) The Bi-Directional Pathways Between Peer Victimization and Depression in Adolescence: A Review. Developmental Psychology, 45(3), 74-88.
[8] McLeod, B.D. and Weisz, J.R. (2018) The Family Environment and Adolescent Depression. Journal of Family Psychology, 32(4), 402-417.
[9] Parra, A., Oliva, A. and Sánchez-Queija, I. (2020) Family relationships and adolescents' psychological well-being: the mediating role of self-esteem. Journal of Adolescence, 33(2), 251-260.
[10] Roeser, R.W., Eccles, J.S. and Sameroff, A.J. (2019) School as a Context of Early Adolescents' Academic and Social-Emotional Development: A Summary of Research Findings. The Elementary School Journal, 100(5), 443-471.
[11] Wang, M.T. and Degol, J.L. (2021) School Climate: A Review of the Construct, Measurement, and Impact on Student Outcomes. Educational Psychology Review, 28(2), 315-352.
[12] Byrne, B.M. (2018) Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming. Routledge.
[13] Fergusson, D.M. and Woodward, L.J. (2019) Mental health, educational, and social role outcomes of adolescents with depression. Archives of General Psychiatry, 59(3), 225-231.

Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations
DOI: 10.54254/2753-8818/52/2024CH0149

The Effect of Dataset Limitation on the Diabetes Risk Factors

Yueyang Ma

School of Mathmatics, Harbin Institute of Technology, Harbin, 150000, China

2022110634@stu.hit.edu.cn

Abstract. Diabetes research is an important topic in the medical domain. However, the research on diabetes risk factors always restricts objects to a certain category. This research wants to figure out the effect of such limitations. The result of this study could give guidance in this field. A dataset built by the Center of Disease Control and Prevention, which has been cleaned is used in this study. The binomial logistic regression analysis is applied to this research and the decision tree model is used for testing. The results suggest that many environmental factors are associated with diabetes. However, taking these risk factors together did not accurately predict diabetes. By analyzing and comparing, limitations on the dataset could increase the accuracy of prediction. The result states that limitations during the study on diabetes risk factors could lead to high-accuracy correlation results. This study provides a good guide for other research on diabetes risk factors to investigate people with more common points in their daily lifestyle.

Keywords: Diabetes, Limitation, Logistic Regression, Decision Tree model.

1. Introduction

In recent years, people with diabetes are increasing rapidly, and the patients are of all ages. Diabetes can be classified into many types, the most common 3 groups are Type 1 diabetes, Type 2 diabetes, and Gestational diabetes [1]. Diabetes not only changes the lifestyle of patients but also leads to mental disease [2]. Although Diabetes is a chronic disease, it has already become one of the ten main causes of death [3]. Risks that accompany diabetes leads to public attention to this disease.

Diabetes is mainly caused by two types of factors, the first is genetic factors and the second is environmental factors [4, 5]. Research in the diabetes field is rich and diverse, which covers therapy methods, risk factors, life influence, and so on. Many research studies focus on the risk factors of diabetes, which makes this field a good academic development prospect. A study in 1997 found that men have an increased chance of developing clinical diabetes when they become weight, increase fat distribution, and become obese [6]. Some researchers studied the relationship between diet, lifestyle, and other physiological factors and diabetes in women [7]. Jing made an effort to verify which factor has a relationship with diabetes by Logistic Regression method [8]. Oppermann and Spritzer conducted a study on Brazilian women and focused on some women-only factors, which gain early age at menarche, and systolic BP were highly and independently related to diabetes [9]. A published paper focused on factors affecting diabetes in the liver transplant population [10]. There was also a study that tried to find the corresponding factors of Type 1 diabetes among children and adolescents [11].

Among all reference studies, there are some common points among them. First of all, every study limits the object of study to a class of people, such as women, men, or the place objects live. Moreover,

DOI: 10.54254/2753-8818/52/2024CH0149

some studies even focus on typical types of diabetes. All such restrictions will lead to more accurate results, but such restrictions also cause the limitations of research conclusions. The study that used the method of Logistic Regression only gained a fuzzy relationship between diabetes and the risk factors.

The reason why a large number of studies abandoned the general applicability of the findings, but focused on small and specific samples is worth studying. Therefore, this article will investigate whether additional restrictions are necessary for diabetes risk factors. Datasets that are not restricted in all aspects will be used, and the logistic regression method will be used for analysis to determine whether a conclusion with high precision and wide application can be obtained. This paper will help other researchers who want to study diabetes to conduct better sample collection.

2. Methods

2.1. Data Source

The data to be used in this paper is from the Kaggle website, and the dataset contains 17 risk factors and detailed information was collected from 70693 people. The strength of this dataset is it contains a large amount of data, and the author who built this dataset balanced all its classes. This dataset is actually from The Behavioral Risk Factor Surveillance System (BRFSS), which is a health-related telephone survey made by the Center of Disease Control and Prevention (CDC) annually. Each year, the survey collects responses from over 400,000 Americans on health-related risk behaviors, chronic health conditions, and the use of preventative services. It has been conducted every year since 1984. This original dataset contains responses from 441,455 individuals and has 330 features. These features are either questions directly asked of participants, or calculated variables based on individual participant responses.

2.2. Variable Introduction

The variables used in this paper are as table 1 shows. The specific situation of the various variables is shown in Table 1 below, and the introduction of some of the abbreviations and criteria are described as follows. High cholesterol (High Chol), Cholesterol check (Chol Check), Body Mass Index (BMI), Coronary heart disease or myocardial infarction (Heart Diseaseor Attack), Physical activity (Phys Activity), Drinks per week (Hvy Alcohol Consump), General health (Gen Hlth), Poor mental health (Ment Hlth), Physical illness or injury (Phys Hlth), Difficulty walking (Diff Walk).

The special standards are following. Cholesterol check is whether cholesterol check in 5 years. 1 of smoker means smoked at least 100 cigarettes in their entire life.1 for physical activity means having physical activity in the past 30 days not including job. The value of Mental health relates to days of poor mental health scale of 1-30 days. Physical illness or injury value indicates physical illness or injury days in the past 30 days scale 1-30.

Table 1. Variable introduction.

Variable	Types	Range/Meaning
Age(x1)	Categorical	13-level age category
Sex(x2)	Categorical	0=female,1=male
High Chol(x3)	Categorical	0 = no,1 =yes
Chol Check(x4)	Categorical	0 = no,1 =yes
BMI(x5)	Numeric	12 to 98
Smoker(x6)	Categorical	0 = no,1 =yes
Heart Diseaseor Attack(x7)	Categorical	0 = no,1 =yes
Phys Activity(x8)	Categorical	0 = no,1 =yes
Fruits(x9)	Categorical	0 = no,1 =yes
Veggies(x10)	Categorical	0 = no,1 =yes
Hvy Alcohol Consump(x11)	Categorical	0 = no,1 =yes

DOI: 10.54254/2753-8818/52/2024CH0149

Table 1. (continued).

Gen Hlth(x12)	Categorical	1 = excellent 2 = very good 3 = good 4 = fair 5 = poor
Ment Hlth(x13)	Numeric	1 to 30
Phys Hlth(x14)	Numeric	1 to 30
Diff Walk(x15)	Categorical	0 = no,1 =yes
Stroke(x16)	Categorical	0 = no,1 =yes
Diabetes(x17)	Categorical	0 = no,1 =yes

2.3. Method Introduction

2.3.1. Logistic Regression

The Logistic function is also called the Sigmoid function. This function has the following form: Logistic Regression is a classical classification method, which is especially good for binary classification problems. It is a classification algorithm based on a generalized linear model. Logistic regression can easily solve the classification problem only by using a logistic function based on the general linear regression model. The Logistic function is also called the Sigmoid function. This function has the following form:

$$y = \frac{1}{1+e^{-z}} \tag{1}$$

By expressing z as $z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots$ (2), substituting z back to function (1), then having logarithmic operation on both sides simultaneously. This paper can get the logistic regression mode:

$$\ln\left(\frac{y}{1-y}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_p x_p \tag{3}$$

Based on the data type of Diabetes, is classified into whether have diabetes or not. Binomial logistic regression analysis is applied. The paper uses SPSS to receive the previous information and gain the result of the regression function and the evaluation of fitting results.

2.3.2. Decision Tree Model

Decision Tree model is a model that displays decision rules and classification results with a tree-like structure. As a type of inductive learning algorithm, the core of the decision tree model lies in transforming seemingly disordered and chaotic known data into a model resembling a tree structure and then using certain technical means to predict unknown data.

3. Results and Discussion

3.1. Descriptive Analysis

In the process of dealing with practical problems, this paper randomly extracts the data, retain its original distribution characteristics, and uses the random 10,246 data for analysis. According to statistics and analysis, age and whether have diabetes were taken as examples. As shown in Figure 1 and Figure 2, the sample size this paper extracted also maintained the same balance as the overall sample and could be analyzed. After a similar analysis, the other variables maintained the same balance. This result shows the randomly selected sample could be used for analysis.

DOI: 10.54254/2753-8818/52/2024CH0149

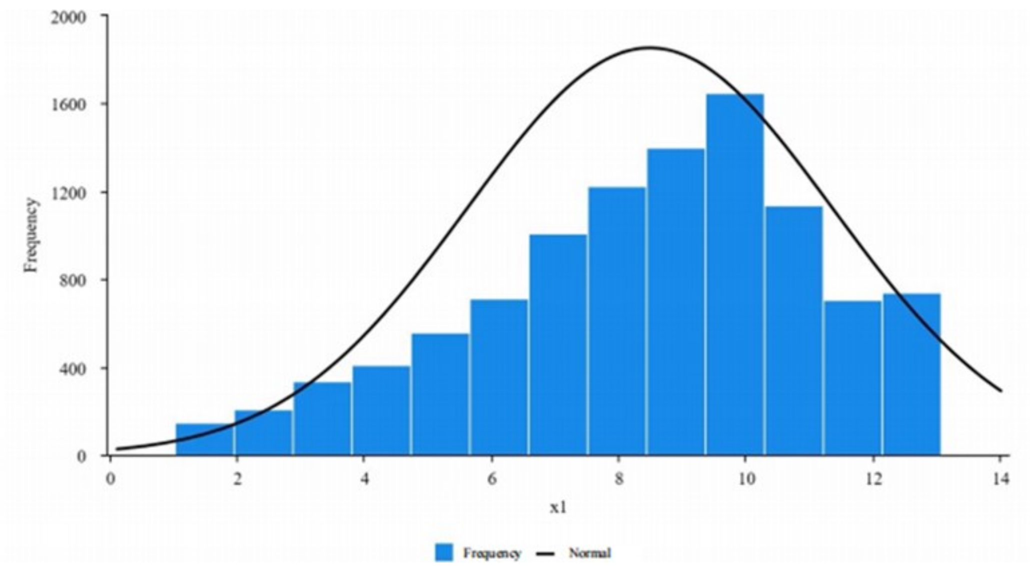


Figure 1. Age distribution in sample.

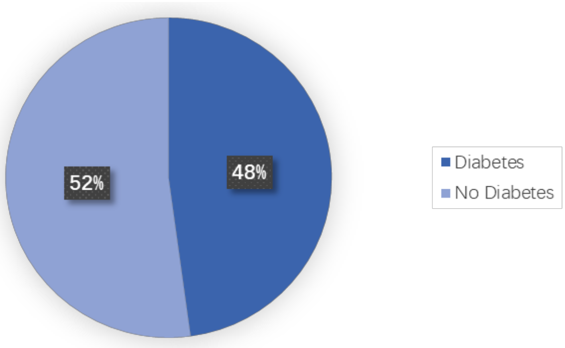


Figure 2. The proportion of diabetes in sample.

3.2. Correlation Analysis

By using the tool SPSS, the correlation between all risk factors and diabetes is shown in Figure 3. By analyzing the result, even though the correlation exists and is remarkable, the responding degree is low. There are 4 negative correlation factors and 13 positive correlation factors. The selection of relevant factors is confirmed by the analysis result. All risk factors could be chosen and put into binomial logistic regression analysis. The next step is to check the colinearity relation between risk factors.

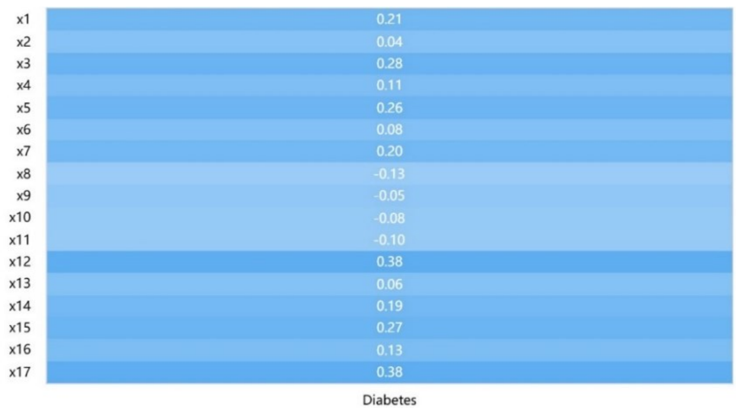


Figure 3. The correlation between dependent and independent variables.

Based on the Kendall corresponding analysis the relationship between the two factors is diverse, including positive relation, negative relation, and irrelevancy (Figure 4). However, for each risk factor, there are no colinearity or multicollinearity risk factors. This implies that the binomial logistic regression hypothesis is satisfied. The binomial logistic regression analysis could be used now.

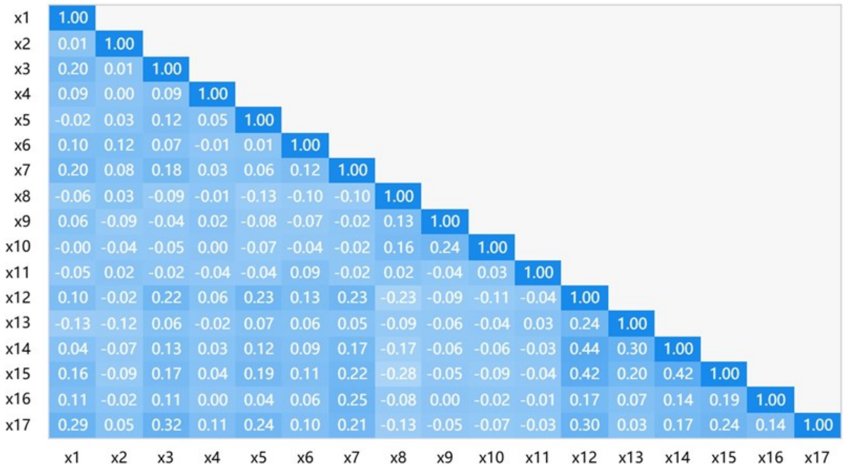


Figure 4. The correlation between independent variables.

3.3. Model Results

The results of binomial logistic regression are shown in Table 2. According to the analysis of p-values less than or equal to 0.05, Sex, High Chol, Chol Check, BMI, Heart Diseaseor Attack, Gen Hlth, Diff Walk, Stroke, Veggies, Hvy Alcohol Consump, PhysHlth produce significant positive effects on Diabetes. Veggies, Hvy alcohol consump, Phys hlth produce obvious negative effects on Diabetes. However, Smoker, Phys Activity, Fruits, and Ment Hlth do not have any effect on Diabetes.

Table 2. Binomial logistic regression analysis result summary.

Term	β	SD	z	Wald $\chi 2$	p	OR	OR 95% CI
x1	0.138	0.010	13.916	193.643	0.000	1.148	1.126 ~ 1.170
x2	0.227	0.049	4.647	21.590	0.000	1.255	1.140 ~ 1.381
x3	0.565	0.049	11.509	132.467	0.000	1.760	1.598 ~ 1.937
x4	1.207	0.200	6.044	36.531	0.000	3.344	2.261 ~ 4.947
x5	0.066	0.004	16.228	263.350	0.000	1.068	1.060 ~ 1.077
x6	-0.011	0.049	-0.215	0.046	0.830	0.990	0.899 ~ 1.089
x7	0.209	0.076	2.746	7.542	0.006	1.233	1.062 ~ 1.432
x8	0.054	0.056	0.971	0.943	0.332	1.055	0.947 ~ 1.177
x9	-0.020	0.051	-0.392	0.153	0.695	0.980	0.887 ~ 1.083
x10	-0.147	0.061	-2.427	5.889	0.015	0.863	0.766 ~ 0.972
x11	-1.030	0.133	-7.748	60.024	0.000	0.357	0.275 ~ 0.463
x12	0.619	0.029	21.537	463.863	0.000	1.857	1.755 ~ 1.965
x13	-0.000	0.003	-0.139	0.019	0.889	1.000	0.993 ~ 1.006
x14	-0.008	0.003	-2.723	7.412	0.006	0.992	0.986 ~ 0.998
x15	0.170	0.067	2.547	6.488	0.011	1.185	1.040 ~ 1.350
x16	0.179	0.108	1.653	2.731	0.098	1.196	0.967 ~ 1.479
x17	0.734	0.052	14.220	202.208	0.000	2.084	1.883 ~ 2.306
Intercept	-6.895	0.273	-25.218	635.937	0.000	0.001	0.001 ~ 0.002

DOI: 10.54254/2753-8818/52/2024CH0149

The gained formula of binomial logistic regression analysis result is:

$$\ln\left(\frac{y}{1-y}\right) = -6.895 + 0.138x_1 + 0.227x_2 + \cdots + 0.179x_{16} + 0.734x_{17} \quad (4)$$

Table 3. Binomial logistic regression analysis accuracy.

		Prediction 0	Prediction 1	Accuracy	Error
True value	0	2516	1782	58.54%	41.46%
	1	1070	5488	83.68%	16.32%
Collect				73.73%	26.27%

Although the results show that about 12 factors relate to diabetes. The accuracy of the model is only about 70%. As shown in Table 4, the prediction accuracy obtained by binary logistic regression is relatively low, to make sure that the output of this model is correct, the decision tree model is used to test twice. The outcome is shown in Table 4, which shows that the accuracy is similar to the result by using binary logistic regression. The accuracy of the decision tree model is 70%, which is close to the accuracy of the binomial logistic regression method.

Table 4. Decision tree model accuracy.

Term	Accuracy	Recall rate	f1-score	N of samples
0	0.74	0.70	0.72	1100
1	0.67	0.71	0.69	951
Accuracy			0.70	2051
Mean value	0.70	0.70	0.70	2051
Mean value(synthesis)	0.70	0.70	0.70	2051

To find the effect of the limitation, the data who maintain the same age stage and gender are chosen randomly. As shown in Table 5, by again using binomial logistic regression, the result this paper gain has a higher accuracy, which improves to 81.10%. Meanwhile, the number of risk factors that may have a significant influence on diabetes is decreased.

Table 5. Binomial logistic regression analysis accuracy with limitation.

		Prediction 0	Prediction 1	Accuracy	Error
True value	0	1262	138	90.10%	9.90%
	1	235	341	59.20%	40.80%
Collect				81.10%	18.90%

3.4. Discussion

In this study, the binomial logistic regression model was used to analyze the effects of different risk factors on diabetes. The decision tree model is used for testing. The accuracy of the binomial logistic regression model is about 70% even changing the method to the decision tree model, the problem of low accuracy cannot be solved. By adding the limitation, the model accuracy under data with limitation is higher than data without it, which shows limitation is a point to the study on risk factors.

Even with the limitation, the accuracy is still only about 80%. The reason why accuracy is low could be analyzed from three aspects. First, the absence of restrictions leads to more and larger differences between individuals, which causes many unrelated factors to be taken into account with large volume. These unrelated, contingent factors showed an association with diabetes and affected the final regression results. Second, the data has no classification on the type of diabetes. Since the causes of different types of diabetes are different, now the data analysis may be influenced. The type and proportion of diabetes patients currently under consideration are unknown. Both type and proportion affect the accuracy of regression results. Third, the risk factors this paper considered only contain the environmental factors. There is a possibility that the genetic factor has a higher relation with diabetes.

4. Conclusion

This paper shows that with the limitation on data, the binomial logistic regression could have higher accuracy and the correlated risk factors of diabetes will be less. This indicates that data with more common points, such as age stage, and gender, even with smaller amounts of data, the result could be simpler and more accurate. Research that based on limitation, could gain better results with a smaller workload.

The results of the study show that in the process of research, it is really necessary to limit the study object so that the study of diabetes risk factors can get better accuracy in correlation analysis. It is possible to include geographical restrictions, gender restrictions, diabetes type restrictions, and restrictions in different diabetes periods to better obtain fitting results. The researcher recommends that freshmen should limit their subjects to risk factors if they want to minimize the influence of unrelated factors and get better correlation results.

Using the results of this paper as a starting point, how much different restrictions can lead to improvements in the accuracy of regression could be investigated. The results trend under the change of number of limitations could be studied to find the most suitable number of limitations. All the branches of this study will make guidance on other studies to focus on the analysis of how environmental factors influence diabetes. The thought of this study could also be applied to the study of other types of disease, which is useful to research pathogenic factors in the medical field.

References

[1] Ondiagnosis, T.E.C. (1997) The Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care, 22.

[2] Richardson, L.K., Egede, L.E. and Mueller, M. (2008) Effect of race/ethnicity and persistent recognition of depression on mortality in elderly men with type 2 diabetes and depression. Diabetes Care, 31(5), 880.

[3] D'Souza, M.J. , Li, R.C., Gannon, M.L., et al. (2019) Leading Causes of Death Information Due to Diabetes, Neoplasms, and Diseases of the Circulatory System, Issues Cautionary Weight-Related Lesson to the US Population at Large. IEEE network.

[4] Green, A. (1990) The role of genetic factors in the development of insulin-dependent diabetes mellitus. Current Topics in Microbiology & Immunology, 164, 3-16.

[5] Longnecker, M.P., et al. (2001). Environmental contaminants as etiologic factors for diabetes. Environmental Health Perspectives Supplements.

[6] Chan, J.M., Rimm, E.B., Colditz, G.A., et al. (1994) Obesity, Fat Distribution, and Weight Gain as Risk Factors for Clinical Diabetes in Men. Diabetes Care, 17(9), 961-969.

[7] Lako, J.V. and Nguyen, V.C. (2015) Dietary patterns and risk factors of diabetes mellitus among urban indigenous women in Fiji.. Asia Pacific Journal of Clinical Nutrition, 10(3), 188-193.

[8] Jing, Z. (2006) Logistic Regression Analysis of Diabetes Pathogenic Factors. Medicine and Society.

[9] Oppermann, K. and Spritzer, P.M. (2024) Prevalence and risk factors associated with diabetes mellitus among middle-aged women in southern Brazil: a population-based study. Menopause, 31(3), 225-230.

[10] Campos, et al. (2022) Risk factors associated with diabetes after liver transplant. Archives of endocrinology and metabolism 66, 2, 182-190.

[11] Andrea, L., Varga, B., et al. (2014) Factors influencing the diabetes-specific health-related quality of life in children and adolescents with type 1 diabetes mellitus. Journal of Child Health Care for Professionals Working with Children in the Hospital & Community, 18(3), 253-60.