

# ***Data Analysis Transformation: Analysis of the Impact of ChatGPT on Various Industry Application***

**Liujunyang Wang<sup>1,a,\*</sup>**

*<sup>1</sup>Business School, University of Sydney, Sydney, Australia  
a. lwan0410@uni.sydney.edu.au*

*\*corresponding author*

**Abstract:** This research investigates Chat Generative Pre-Trained Transformer (ChatGPT)'s potential to significantly transform data analysis across various industries. As an advanced artificial intelligence language model, ChatGPT offers innovative solutions for automating and refining data analysis processes. Its applications are extensive, encompassing fields such as education, commerce, healthcare, and the military. The study provides a thorough assessment of ChatGPT's capabilities, highlighting its strengths in generating actionable insights and enhancing decision-making efficiency. Additionally, it evaluates the obstacles and constraints linked to artificial intelligence, encompassing apprehensions regarding data privacy, security, biases, and the precision of outcomes. By conducting detailed analyses of practical implementations and comparing them with previous research, this study emphasizes the considerable impact of ChatGPT while also underscoring the need for cautious and responsible use. The findings indicate that while ChatGPT holds significant promise for transforming data analytics, additional research and development are essential to enhance its performance and address potential risks. The study calls for ongoing investigation to address these challenges and maximize the benefits of ChatGPT in the evolving landscape of data analysis.

**Keywords:** ChatGPT, Data Analytics, Artificial Intelligence.

## **1. Introduction**

With the proliferation of data in today's world, information analysis has become critical to many academic fields. The tremendous progress in artificial intelligence and linguistic technology has created more advanced and efficient language models.

Chat Generative Pre-Trained Transformer (ChatGPT), an artificial intelligence (AI) model created by OpenAI, Inc., is specifically designed to produce text closely resembling human language. ChatGPT, a language model trained on an extensive textual dataset, can generate material that closely resembles the data it was trained on. It has made substantial progress in Netherlands (NL), which is concerned with enabling computer systems to comprehend, interpret, and produce human language [1]. ChatGPT, built around the GPT-3 model released in 2020, enables interactive discussions and demonstrates sophisticated skills in comprehending and producing text that resembles human speech. The system can address subsequent inquiries, rectify mistakes, examine erroneous assumptions, and decline unsuitable demands [2].

AI has made remarkable progress and is now widely used across various sectors, including healthcare and education. By analyzing vast amounts of data, AI systems can mimic brain functions and perform routine tasks. In healthcare, AI supports practitioners by integrating patient records, analyzing diagnostic images, and identifying health conditions [3]. ChatGPT, through Natural Language Processing (NLP) and sentiment analysis, helps businesses understand consumer behavior and emotions, leading to better marketing strategies and improved customer service [4]. In finance, AI examines transactions, analyzes market trends, offers compliance guidance, assesses credit risk, and generates reports, aiding decision-making processes. Additionally, data analysis in healthcare enhances patient care by improving diagnostic accuracy and treatment options. AI applications in education aim to streamline administrative tasks and offer academic support. ChatGPT has gained global recognition for its logical and insightful responses [5]. Research shows that AI can pass multiple exams, suggesting it could achieve a college degree, despite varied performance ratings. However, the success of AI also raises new challenges, such as AI-assisted cheating, leading some educational institutions to ban its use [6].

This article explores ChatGPT's potential to revolutionize data analytics across industries. By assessing its strengths and limitations, the paper encourages further research into this innovative tool. It thoroughly examines ChatGPT's role in data analysis, its challenges, and real-world applications, outlining a methodology that includes data collection, processing, and analysis. Moreover, a comparative analysis of existing research is conducted to highlight the uniqueness and significance of the study's findings. The conclusion offers a summary of key insights and suggests directions for future research, with the goal of enhancing ChatGPT's capabilities in data analysis. The study is structured into five sections: Introduction, Methodology, Results and Discussion, Conclusion, and References.

## **2. Methodology**

### **2.1. Dataset Description and Preprocessing**

This study utilizes a dataset predominantly comprising secondary data from academic publications, industry papers, and case studies. These sources describe the development and deployment of ChatGPT across many disciplines. The data sources encompass scholarly publications, conference proceedings, white papers, and technical reports, guaranteeing a thorough examination of ChatGPT's evolution and influence. Data extraction involves systematically extracting the essential attributes of research projects in a structured and standardized manner, using information from journal articles and reports [7]. By categorizing the data, this study can analyze ChatGPT's historical milestones, sectoral uses, and social and economic repercussions in detail, making it easier to do qualitative analysis.

### **2.2. Proposed Approach**

ChatGPT has seen widespread adoption across various industries, reaching over 100 million monthly active users by January 2023 [8], making it the fastest-growing application in terms of user engagement. It enhances human tasks by understanding goals and generating content, marking a shift from merely analyzing data to actively creating it. Its powerful text generation also helps create synthetic data, addressing the issue of limited data in AI projects and promoting broader AI adoption. This article examines ChatGPT's impact on data analytics across industries, including education, trade, and the military, and discusses the challenges of its implementation [9]. The study's pipeline is shown in Figure 1.

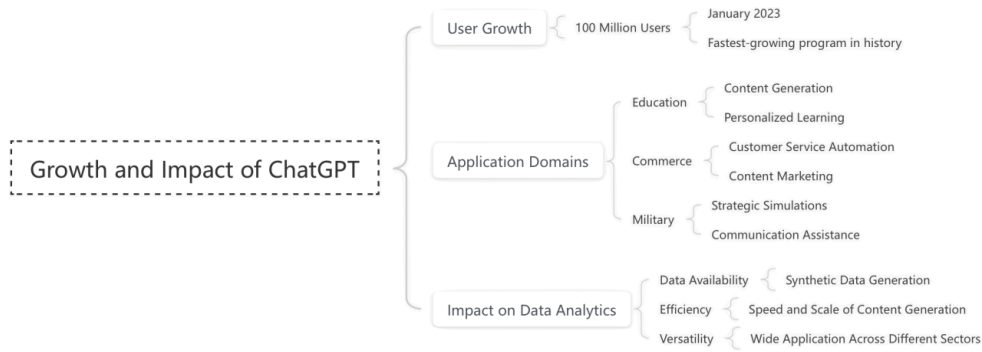


Figure 1: Research details.

### 2.2.1. ChatGPT in Education

ChatGPT is an advanced iteration of the GPT model, evolving from GPT-1 to GPT-4. GPT-1, introduced in 2018, was the first generative language model based on the Transformer architecture, using unsupervised learning with fine-tuning. In 2019, GPT-2 improved on this by incorporating multi-task learning with more parameters and data, achieving broader task performance with minimal fine-tuning. GPT-3, with over 100 billion parameters, introduced meta-learning and contextual learning, improving generalization even with little or no training data [10]. The ChatGPT pilot, or InstructGPT, used Reinforcement Learning with Human Feedback (RLHF) to better align with user intent. GPT-4, released in 2023, is a multimodal model capable of analyzing both images and text, achieving human-level performance in professional and academic tests.

ChatGPT utilizes data analytics in education across various domains. These include creating and automating course materials, instructing in statistical programming, supporting data analytics projects, and providing post-course exercises and self-assessment opportunities. ChatGPT can produce course materials relevant to data analytics, including practice questions, test questions, and instructional aids. The course contents can be tailored to suit the particular requirements of the course, hence assisting instructors [11]. Instructors can utilize ChatGPT to produce questions about hypothesis testing and sample size calculations for a statistics course. The generated content can be modified as necessary. Furthermore, ChatGPT can aid students in comprehending intricate statistical concepts while working on data analysis projects. It can offer suggestions and code examples for analysis, assist in identifying and resolving programming errors, and provide explanations and solutions.

ChatGPT helps students learn multilingual programming and supports post-class activities by generating questions and answers, offering instant feedback to identify weaknesses and improve learning. However, it also poses risks of academic dishonesty by enabling easy plagiarism, undermining fairness in evaluations. Over-reliance on ChatGPT can hinder independent thinking and problem-solving skills. Additionally, ChatGPT's potential for errors can confuse students, especially in areas with limited data. Teachers may struggle to assess student learning accurately if students submit high-quality AI-generated work instead of their own. Educational institutions must update policies and provide better training to address these challenges and maximize AI's benefits.

### 2.2.2. ChatGPT in the Business Domain

ChatGPT fulfills several functions in data analysis within the business sector, particularly in customer care and support, e-commerce, digital marketing, the financial industry, and data analysis and decision support. ChatGPT utilizes advanced natural language processing and machine learning techniques to handle customer service inquiries, leading to quicker response times and cost reductions. It is utilized in e-commerce to automate customer care and assist consumers in discovering suitable

products through personalized recommendations and services. This ultimately enhances customer happiness and stimulates sales development. Furthermore, ChatGPT can efficiently manage a high volume of client inquiries, particularly during busy shopping seasons, successfully addressing the issue of limited workforce availability [4]. ChatGPT in digital marketing effectively minimizes labor expenses by efficiently managing a high volume of repeated duties, including customer inquiries and order processing [12]. The powerful analytics engine of this system offers valuable data insights that assist organizations in optimizing their marketing efforts and eventually enhancing their return on investment. ChatGPT enhances business efficiency by automating processes and lowering the time and expenses of resolving frequent issues. ChatGPT enhances corporate productivity by automating tasks and lowering the time and expenses of addressing frequent issues.

ChatGPT uses natural language processing to quickly address client needs, gather precise data, and provide customized financial guidance by analyzing consumer requirements. It manages customer inquiries, account information, performs risk assessments, and detects fraud, enhancing decision-making and operational efficiency. By analyzing extensive text data, ChatGPT aids in strategic business decisions, optimizing customer experience and providing a competitive edge.

However, ChatGPT also poses risks. Data privacy and security concerns can lead to breaches or unauthorized use, harming businesses legally and reputationally. Its reliance on training data may introduce bias, leading to inaccurate analysis and poor decisions. While ChatGPT improves customer service, it cannot fully understand or address emotional needs, potentially lowering customer satisfaction. Over-reliance on ChatGPT might also reduce staff proficiency in data analysis and customer service, hindering innovation. Therefore, organizations must carefully manage data privacy, ensure accurate analysis, and balance automation with human input while being mindful of potential biases.

### 2.2.3. ChatGPT in the Military Domain

The potential utilization of ChatGPT in the study of military data is highly evident. It efficiently handles vast volumes of text data, including intelligence reports. The depth of analysis has been enhanced, resulting in improved effectiveness in decision-making. This is accomplished by creating summaries and extracting essential information, finding patterns seen, analyzing feelings expressed, and recognizing entities within the data. ChatGPT facilitates the automated identification of targets. Targets can be identified by analyzing sensor logs, examining satellite photos, and reviewing surveillance footage, and their movements can be predicted.

Additionally, ChatGPT improves the performance of military robots and autonomous systems, processes natural language commands, generates situational awareness responses, and conducts predictive data maintenance [1]. Within cybersecurity and cryptography, ChatGPT can spot irregularities in network traffic and recognize potential cyber hazards. In addition, it may assist in the domains of sophisticated steganography and cryptography. By incorporating neural networks, ChatGPT can enhance autonomous control and decision-making functions in missile guidance systems, enhancing the effectiveness and precision of target identification and strikes [1]. ChatGPT's capacity to process and incorporate data from various sources enables it to offer timely situational awareness, aiding military troops in making well-informed decisions during operations and generating a comprehensive overview of the battlefield environment by merging data. Utilizing ChatGPT's AI technology can improve situational awareness of individual soldier wearable systems by offering real-time suggestions or notifications. It can also assist in assessing the battlefield environment, including terrain analysis and weather prediction [1]. Consequently, it plays a crucial role in planning and executing military operations.

Issues such as bias, security, and legal accountability constrain the deployment of artificial intelligence in military contexts and necessitate meticulous attention throughout the implementation

process. Currently, the utilization of ChatGPT in military robotics is still at the research stage and has not yet achieved complete development or deployment [1]. Furthermore, the precision of ChatGPT analysis is contingent upon the caliber and volume of its training data [13]. The results produced by ChatGPT depend on the quality of the training data from which they are derived. If the training data contains biases or erroneous information, it might lead to biased analyses and misleading results. Such conclusions and assessments could potentially adversely affect the precision of military judgments. Moreover, employing ChatGPT to manage classified military information entails potential hazards, including malevolent interferences or abuses, which could result in data breaches or even facilitate cyber-attacks. These hazards are substantial and provide a worrisome threat to national security. Hence, it is imperative to meticulously evaluate these potential hazards of employing ChatGPT in military contexts. Furthermore, to safeguard military operations and national security from any detrimental impact caused by the implementation of AI technology, it is imperative to employ suitable measures to neutralize these potential dangers.

### 3. Result and Discussion

#### 3.1. Result

Using ChatGPT for data analysis offers significant benefits but also presents challenges. Its strong natural language processing skills make data analysis more intuitive, allowing users to ask questions without needing complex programming knowledge. ChatGPT can quickly generate initial analytics like descriptive statistics and trend analysis, automate repetitive tasks such as data cleansing, and integrate multiple data sources to create informative reports. This makes it useful in fields like finance, healthcare, and market research.

However, there are limitations. ChatGPT's results can be imprecise, especially in complex or specialized domains, and it cannot replace the expertise of domain specialists. Data privacy and security are critical concerns, and the quality of its analyses depends on the training data. While proficient in many tasks, ChatGPT often needs integration with other tools for more complex analyses.

ChatGPT has significantly impacted various fields, as shown in the Table 1. In education, it offers personalized learning recommendations and helps teachers with instructional materials and grading. In finance, it analyzes market trends and assists with investment decisions. In customer service, ChatGPT engages with clients using natural language, providing a seamless experience. It also analyzes social media data and consumer feedback to enhance marketing strategies, improve ad copy, and create natural language product descriptions. Overall, the utilization of ChatGPT is broadening and intensifying, resulting in significant effects on productivity and effectiveness across diverse industries. Nevertheless, as these technologies progress, it is crucial to consistently focus on properly tackling the obstacles they provide, including data privacy, bias, and ethical issues arising from technology.

Table 1: Performance of different algorithms.

Domain	Application of ChatGPT	Existing Issues
Education	Personalizes learning, aids in material creation, automates grading.	May produce inaccurate results, struggles with complex or specialized data, and lacks deep domain expertise.
Business	Analyzes trends, assesses risks, aids decisions, enhances marketing, and provides customer service.	Data privacy and security concerns; analysis limited by training data; complex tasks require other tools.
Military	Strategic analysis, intelligence aggregation, automation, etc	Critical data privacy and security concerns; limited by training data; may not handle specialized tasks well.



### 3.2. Problem and Solution

When integrating ChatGPT into data analytics, addressing data privacy and security is crucial. Given its handling of sensitive information, strict security measures like data encryption, anonymization, access controls, and locally installed models are essential. Additionally, ChatGPT's outputs should be treated as preliminary and validated with expert opinions and other tools, especially in complex domains like medicine or law. Regular updates to the model's training data and algorithms are necessary to improve accuracy.

ChatGPT also faces challenges with bias and fairness due to potential biases in training data. Periodic reviews, de-biasing procedures, and diversity evaluations can help mitigate these issues. Users should provide clear instructions and engage in iterative communication to avoid misinterpretations, as ChatGPT may struggle with context and intent. It also has limitations in real-time data processing, which can be addressed by combining it with specialized platforms. By identifying and addressing these challenges, organizations can optimize ChatGPT's effectiveness in data analytics while minimizing risks.

### 4. Conclusion

ChatGPT, as a sophisticated tool for natural language processing, has demonstrated remarkable potential across various domains. It not only streamlines data analysis by providing preliminary insights but also enhances users' understanding and utilization of data through its powerful text-generating capabilities. Despite its advantages, implementing ChatGPT presents certain challenges. The accuracy of its outputs heavily relies on the quality of the training data, which can lead to erroneous conclusions, particularly in complex or specialized fields. Additionally, deploying ChatGPT requires stringent security measures to address data privacy and protection concerns. Nonetheless, ChatGPT remains a versatile tool with significant applications in sectors such as education, finance, and marketing. Integrating ChatGPT with other data analysis tools and expert inputs is expected to enhance its effectiveness. Future research should focus on expanding ChatGPT's analytical capabilities, reducing model bias, and improving real-time data processing performance. These improvements will enhance ChatGPT's role in data analytics and drive industries toward more intelligent, data-driven decision-making.

In conclusion, ChatGPT has not only entered but significantly transformed the field of data analytics. Its widespread adoption requires careful consideration to balance the convenience it offers with the potential risks and challenges. As technological advancements progress, ChatGPT is expected to assume an increasingly central role in data analysis, enhancing precision and efficiency in decision-making processes across various industries. Its transformative impact highlights the importance of harnessing its potential thoughtfully to advance the field of data analytics.

### References

- [1] Biswas, S.S. (2023) *Potential use of chat gpt in global warming. Annals of biomedical engineering*, 51(6), 1126-1127.
- [2] Waghmare, C. (2023) *Introduction to ChatGPT. Unleashing The Power of ChatGPT: A Real World Business Applications*. Berkeley, CA: Apress, 1-26.
- [3] Lo, C.K. (2023) *What is the impact of ChatGPT on education? A rapid review of the literature. Education Sciences*, 13(4), 410.
- [4] George, A.S., George, A.S.H. (2023) *A review of ChatGPT AI's impact on several business sectors. Partners universal international innovation journal*, 1(1), 9-23.
- [5] Ellis, A.R., Slade, E. (2023) *A new era of learning: considerations for ChatGPT as a tool to enhance statistics and data science education. Journal of Statistics and Data Science Education*, 31(2), 128-133.
- [6] Dibble, M. (2023) *Schools ban ChatGPT amid fears of artificial intelligence-assisted cheating. VOA News*.

- [7] Schmidt, L., Olorisade, B.K., McGuinness, L.A., et al. (2020) *Data extraction methods for systematic review (semi) automation: A living review protocol*. *F1000Research*, 9.
- [8] Mann, B., Ryder, N., Subbiah, M., et al. (2020) *Language models are few-shot learners*. *arXiv preprint:2005.14165*, 1.
- [9] Ouyang, L., Wu, J., Jiang, X., et al. (2022) *Training language models to follow instructions with human feedback*. *Advances in neural information processing systems*, 35, 27730-27744.
- [10] Achiam, J., Adler, S., Agarwal, S., et al. (2023) *Gpt-4 technical report*. *arXiv preprint:2303.08774*.
- [11] Tlili, A., Shehata, B., Adarkwah, M.A., et al. (2023) *What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education*. *Smart learning environments*, 10(1), 15.
- [12] Xu, Y., Shieh, C.H., van, Esch, P., et al. (2020) *AI customer service: Task complexity, problem-solving ability, and usage intention*. *Australasian marketing journal*, 28(4), 189-199.
- [13] Wang, W., Liu, H., Lin, W., et al. (2020) *Investigation on works and military applications of artificial intelligence*. *IEEE Access*, 8, 131614-131625.