

ChatGPT's Impact on Data Science Students' Learning Performance: A Systematic Review and Prospects

Tianbinhuo Feng

No. 2025, Chengluo Avenue, Chengdu City, Sichuan Province, 610106, China

2046017234@qq.com

Abstract. This paper systematically studies the effect of ChatGPT on the academic performance of data science students. Through the many articles reviewed, ChatGPT has a significant role to play in the field of data science education. Since ChatGPT can provide students with immediate feedback and personalized support, especially in helping with coding, problem-solving, and conceptual understanding, this feature shows great help and potential in improving the learning outcomes of data science education. However, over-reliance on AI tools may negatively affect students' ability to think critically and solve problems independently. We highlight the importance of using ChatGPT in a balanced way, using ChatGPT as an adjunct resource rather than a primary resource, and highlight opportunities for future research into how AI tools affect long-term learning outcomes.

Keywords: ChatGPT, data science education, learning performance, artificial intelligence

1. Introduction

The rapid development of artificial intelligence (AI) has a huge impact on all walks of life, and education is no exception. ChatGPT has been the most popular tool in artificial intelligence in recent years, which helps students gain the technical skills and conceptual understanding needed for learning through advanced language processing capabilities [1]. Especially in student programming, statistical analysis, and machine learning courses, artificial intelligence tools such as ChatGPT can make it easier for students to understand complex course knowledge, enrich the learning experience, and master more knowledge outside of data science courses [2]. ChatGPT can help students simplify difficult concepts and solve problems efficiently. This kind of autonomous, real-time, and more accurate feedback is especially beneficial for data science students, because the problems in the major are more complex than those in the general major, requiring systematic analysis and long-term solutions, so in addition to the teacher, the students' immediate response is essential for them to better prepare for the exam [3]. The potential over-reliance on artificial intelligence is one of the most worrying problems, which may lead to a decline in students' problem-solving skills, such as a lack of critical thinking and independent writing [4]. Data science requires students to have the computer skills to analyze data and derive unique insights and optimize methods from it [5]. Overuse of ChatGPT may result in students not having their thinking and only once relying on AI for help [1]. The long-term consequences of this dependence are terrible, and these missing thinking and abilities are the most important skills they need to have in the real world [2]. To address these issues and challenges, educators need to adapt strategies to address changing learning patterns, enabling blended learning models that use ChatGPT as a complement to after-school learning rather than a direct substitute for students completing assignments [3]. Educators should educate students about the appropriate use of AI, encourage students to make efforts to explore independently before using AI assistance when there is no solution, and mention the importance of critical thinking and independent problem solving [3]. Improved teaching methods allow students to experience the benefits of AI learning while remaining open to analytical thinking [4].

This study will analyze the advantages and limitations of ChatGPT, discuss the relevant research methods of other scholars in this field, and explore the future direction of integrating artificial intelligence into data science learning environments [4].

2. Theoretical Background and Technological Impact

ChatGPT is a tool for applications arising from the modeling of high-level languages and has been widely discussed since its release. Researchers have used several approaches to explore the practical benefits of ChatGPT in real academic scenarios [5]. ChatGPT is widely used in education, especially for data science majors. As a tool developed based on sophisticated machine

learning and natural language processing algorithms [4], ChatGPT has had a far greater impact on the discipline of computer science than on other disciplines, particularly the specialized field of data science. As the key discipline of data science continues to develop in academia and industry, advanced AI tools such as ChatGPT offer great potential for industrial development [4]. In education, data science students are often required to perform complex data analysis, machine learning, and predictive modeling, where ChatGPT can undoubtedly serve as a powerful assistant in their learning process [3]. The tool helps students quickly grasp complex concepts, provides immediate feedback on coding and data processing problems, and explains complex statistical methods. In addition, ChatGPT serves as a resource for the soft skills of development, communication, and collaboration that are essential for data scientists. Whether in academic papers, presentations, or industry reports, it can help students make technical explanations more understandable [3].

2.1. The Significance of ChatGPT in Data Science Education

ChatGPT is important for data science students because the major is characterized by the complex real-time processing and analysis of data. Unlike many other humanities disciplines such as history or literature, they deal most of the time with static data such as text files and archives [5]. In contrast, data science requires the ability to quickly interpret and react to changing information. This need for instant, complex analytics sets data science apart, creating challenges and opportunities that are highly matched to ChatGPT's capabilities. In scientific fields such as biology or chemistry, processing large amounts of data is as common as data science majors [5]. However, batch mode is usually used at the end of an experiment, whereas data science requires continuous analysis and feedback throughout the process [4]. This requirement highlights the need for tools that can process large amounts of data quickly and accurately. Because of this, ChatGPT is a unique tool to help data science students [20].

ChatGPT not only has a strong correlation and matching degree in data majors but also provides a learning tool with strong experience for learners of this major [6]. Because data science is designed with a large number of theoretical and technical problems, it is difficult for beginners or relatively backward students to overcome the difficulties [7]. ChatGPT can assist learning in this process [7].

ChatGPT's powerful algorithms help data science students comprehensively analyze searches and extract valuable information from them [1]. For students who need to be exposed to a lot of data information for a long time, such help is more suitable for their learning style, saving time and energy.

Data science itself integrates mathematics, statistics, computer science and other fields. ChatGPT, as a language model for natural language processing and machine learning, has many common principles and technologies with data science [8]. Therefore, ChatGPT can better adjust the feedback according to the problems of students in this major, and provide more comprehensive and personalized resources [8].

3. Materials and Methods

3.1. Search Strategy and Inclusion Criteria

This systematic review's information sources included ERIC (Education Resources Information Center), Web of Science, and Google Scholar. To retrieve relevant studies, a structured search strategy was implemented using keywords and logical operators. The search keywords are (1) Chat GPT; AND (2) Data/Computer Science (In the Chinese education system, data science and computer science are closely related and often belong to the same school, which reflects the intersection and complementarity of the two in academic research and practical applications. Data science relies on techniques and methods provided by computer science, such as programming, algorithms, database management, etc. Therefore, it is important to use computer science as a keyword when studying data science. This interdisciplinary perspective more fully explains the trajectory and applications of data science.); AND (3) " learning performance" OR "academic achievement" OR "grades". The search was restricted to studies from 2023 to 2024 to ensure relevance and timeliness. Advanced search settings were employed to filter results according to these criteria, enhancing the focus and relevance of the search results.

3.2. Summary of the Studies Screening Approach

A total of 173 initial studies were identified through searches in ERIC (n=52), Web of Science (n=41), and Google Scholar (n=80).

The following is a detailed selection criteria and process. Firstly, each study's abstract was initially scanned to ensure it specifically addressed the study's focus on ChatGPT's impact on Data Science education. Studies were excluded if the abstracts did not substantively engage with the core topics (n=30). Secondly, duplicate studies were excluded (n=8), this included two sets of studies duplicated across three databases and two sets duplicated across two databases. Duplicate studies with an earlier publication date in the two articles were excluded to ensure the inclusion of up-to-date and comprehensive data. Thirdly, exclude studies published in languages other than English (n =18). Fourthly, exclude studies on subjects not under review (Data Science), for example, data science curriculum in Business education or Engineer education, rather than study data science in this specific major. Studies unrelated to the review's focus, such as those not specifically addressing the impact of ChatGPT on learning

performance within the Data Science curriculum, were excluded (n=40). Fifthly, non-academic sources should be excluded from Google Scholar (such as newspapers, magazines, Internet sites, blogs, etc.) (n = 24). Sixthly, excluding the full text does not meet the following screening criteria: 1. studies that merely mention relevant keywords (Data Science) without providing an in-depth exploration of the research question, for example, there were only one or two paragraphs in the article that talked about ChatGPT's impact on Data Science students, not the entire article; 2. exhibit non-empirical research and methodological deficiencies such as inadequate sample sizes, lack of control groups, insufficient statistical analysis; 3. only peripherally address the topic in limited sections of the text. The above are excluded from this review (n=23). Finally, inaccessible studies were excluded due to subscription or other access issues (n=11).

After repeated reading and confirmation, 21 studies were finally left as the main objects of this review.

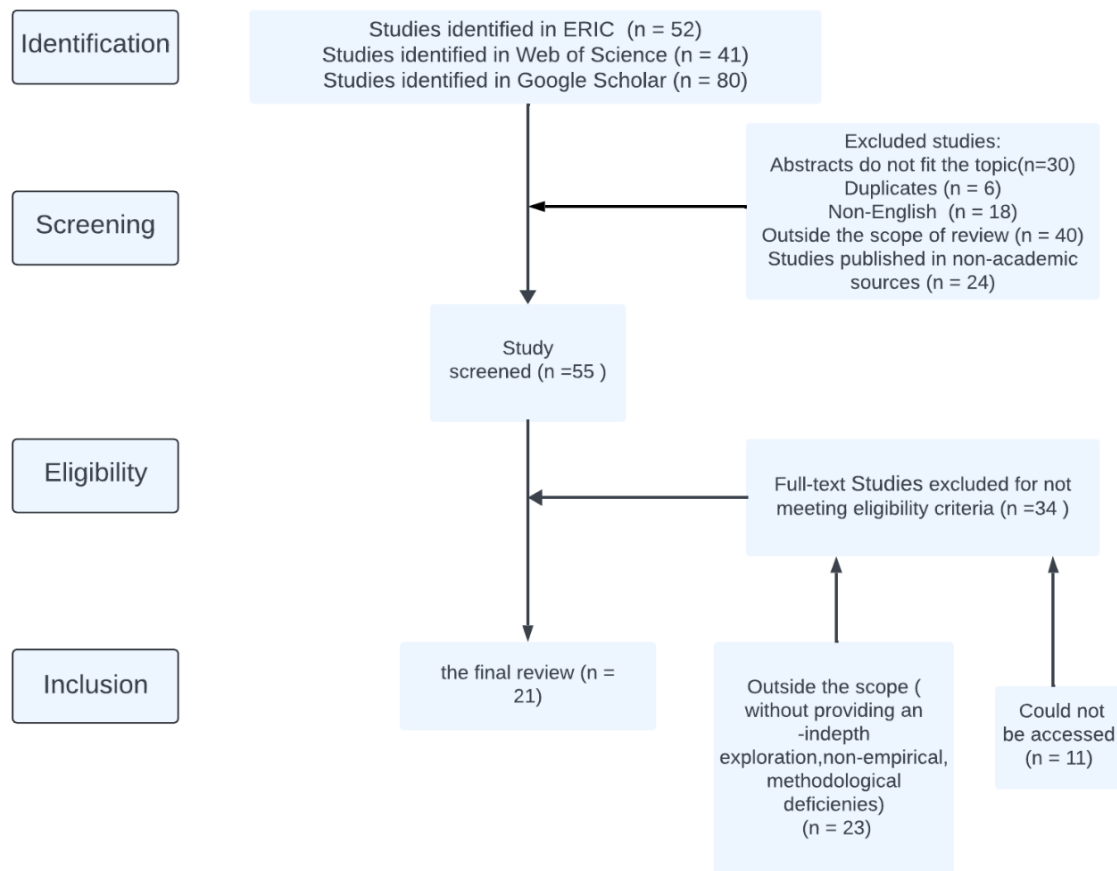


Figure 1. Flowchart of the Studies Selection Process

4. Results

4.1. Summary of the ChatGPT Benefits in Data Science Education

ChatGPT has demonstrated multiple benefits in data science education, becoming a valuable tool for supporting teachers and students [5]. It has also shown its potential to revolutionize the data science learning experience by handling versatile educational tasks [5].

One of ChatGPT's contributions to data science is supporting coding tasks in student assignments such as helping with complex code, correction, and modifying completed code by providing code generation, error checking, and debugging support [9]. This support can help students avoid mistakes and improve their learning. The popularity of this function in students' daily learning activities is very high [9]. Students often worry about code syntax, logic errors, and debugging problems, which seriously hinder the completion and progress of the entire task. By generating code along with comments, ChatGPT helps students better understand programming languages and paradigms. As such, it is an indispensable resource for data science students [11].

ChatGPT can also enhance students' understanding and interest in the course by providing an additional interactive learning experience. This ability is of great benefit to data science students, as complex concepts can be very challenging for beginners

[10]. ChatGPT can transform complex problems into more understandable ones and provide real-world examples, making it easier and more effective for students to grasp these concepts and improve their motivation to learn. In addition, ChatGPT encourages students to explore at their own pace, asking questions and receiving tailored answers to enrich their overall learning experience [10].

ChatGPT also fosters critical thinking in students, as it not only encourages users to receive information but also encourages them to actively participate in assessing the quality and credibility of the conversation [11]. When students ask questions using ChatGPT, reminders at the bottom of each conversation ask the user to evaluate the accuracy of the generated information, and as students begin to evaluate and cross-check with other sources, they reflect on potential assumptions or errors. In the process, it helps to improve their critical thinking. This also facilitates the machine learning assessment process for data science students. These processes strengthen students' core critical thinking skills, such as questioning the validity of arguments, assessing the relevance and reliability of data, and synthesizing multiple perspectives to arrive at fuller, richer answers [11]. These critical evaluation steps are more difficult than mere analysis, requiring students to think deeply about how conclusions are reached and to question them if necessary.

Providing personalized learning is another of ChatGPT's strengths. ChatGPT's high adaptability allows it to meet the unique learning needs of each student [12]. ChatGPT can be based on student's existing knowledge, communication, and specific questions. Provide different answers when different students ask the same question so that students can experience one-on-one education. This personalization is especially useful in data science, where students tend to have a wide variety of problems that may require different levels and ways of support. ChatGPT provides instant, personalized feedback and helps to bridge the gap between different learning progress, ensuring that all students complete the course effectively [12].

In addition, ChatGPT has been shown to improve academic performance, and a subset of students' academic performance has improved as a result of its support. By answering questions quickly and accurately and guiding students through complex problems, ChatGPT can help students achieve high scores in a short time [10]. In data science education, assignments often involve complex multi-step processes such as data analysis and algorithm development, in which case using tools such as ChatGPT allows students to achieve better results faster [7].

Supporting distance learning is another important benefit of ChatGPT. As educational institutions increasingly adopt online and blended learning models, tools that support distance learning become critical [13]. ChatGPT provides distance learners with an accessible, ready-to-use resource that helps bridge the gap between traditional face-to-face teaching and distance learning [14]. By providing immediate feedback, personalized support, and interactive learning opportunities, ChatGPT ensures student engagement.

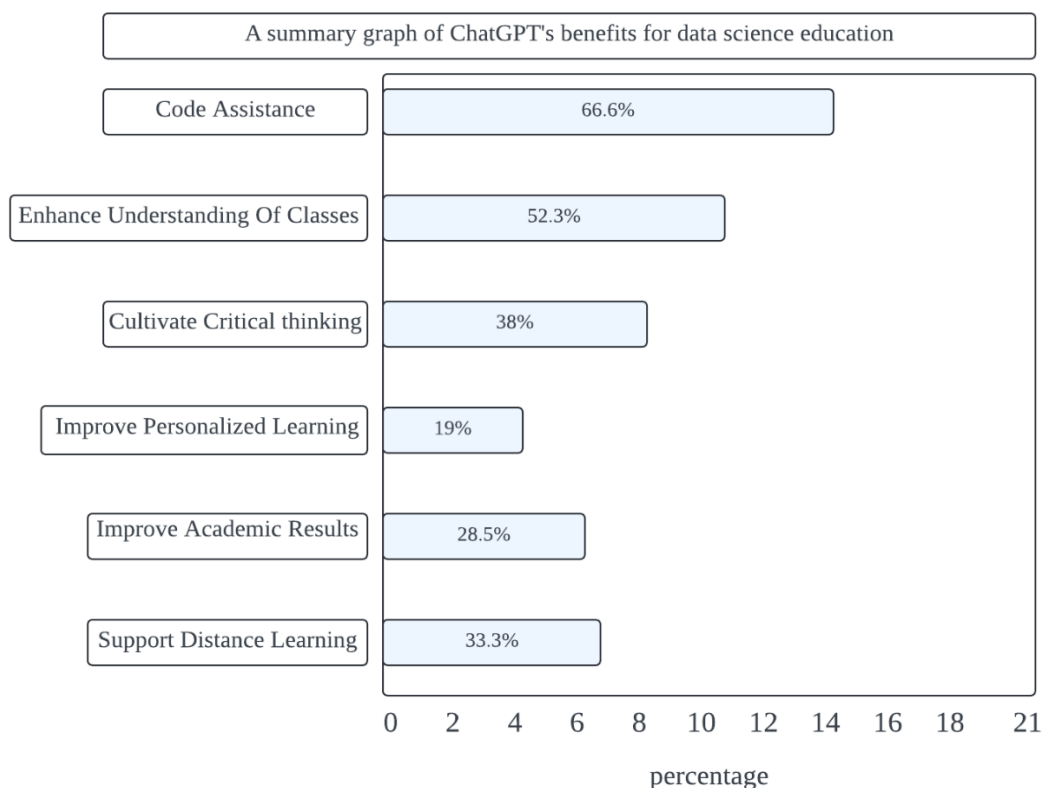


Figure 2. Percentage Table Showing the Benefits ChatGPT Brings to Data Science Education

4.2. Summary of the ChatGPT Limitations and Concerns in Data Science Education

While a groundbreaking tool in many educational settings, ChatGPT has shown specific limitations and raised concerns when applied to data science education. These issues are not only technical but also affect student learning, the effectiveness of teaching, and the integrity of the academic environment.

Most concerning was the inaccuracy of answers, mentioning ChatGPT's inaccuracy in handling complex statistical processes, especially those such as p-values and confidence intervals. These inaccuracies can lead students to a fundamental misunderstanding of statistical principles essential to data science. For example, consider the common statistical problem of how to interpret p-values. The correct interpretation would be that the p-value represents the probability of the observed statistic or more extreme case given that the null hypothesis is true. However, ChatGPT may incorrectly simplify this interpretation with statements such as "The p-value is the probability that the model is correct," which is completely false [9].

Some of the articles highlight that a significant risk of frequent use of AI like ChatGPT is over-reliance on technology to solve problems. When students turn to ChatGPT to answer all their questions, they may neglect to critically engage in learning or bypass the logical reasoning process that is essential in data science [13]. This dependency may prevent students from developing the skills to explore and obtain solutions independently, which is a key component of any data science course designed to produce innovative and self-sufficient analysts [14].

As ChatGPT can be analyzed to generate comprehensive answers, ethical issues such as academic plagiarism and misappropriation may arise [1]. Students may use AI-generated content directly as final answers for homework or even exams, which will challenge the fairness and integrity of educational assessment. This practice will undermine the value of academic evaluation, lead to the normalization of cheating, and further aggravate the academic ethics problems within the institutions [8].

While ChatGPT has advantages in developing critical thinking by encouraging students to evaluate, question, and synthesize information, there are concerns that over-reliance on AI may have disadvantages [15]. Several studies have highlighted that if students rely too much on ChatGPT, they may miss opportunities to think deeply and engage with complex problems. This can lead to a superficial approach to problem-solving. Students become proficient at using AI to find solutions, but they are unable to engage in the critical thinking processes required to truly understand, analyze, and synthesize information. This weak critical thinking poses a risk, particularly in fields such as data science, where the ability to critically evaluate data-driven insights and make informed decisions is crucial. Thus, while ChatGPT is an important tool for developing critical thinking, it must be used thoughtfully and with certain restrictions to avoid undermining the skills it is designed to improve.

In addition, ChatGPT may have adverse effects on the way students learn [11]. When students have become accustomed to the thrill of an answer via ChatGPT, this can contribute to a mindset that favors instant gratification over lasting problem-solving. [10] This shift can have implications for long-term learning outcomes, as students may prioritize speed over depth of understanding, which can impact their readiness to tackle career challenges that require thorough, detailed data analysis.

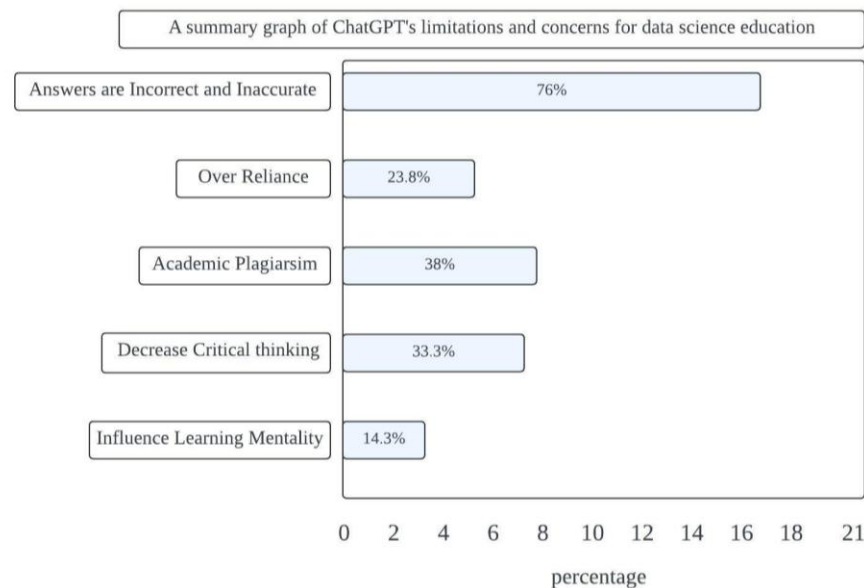


Figure 3. A Percentage Table Showing the Limitations and Concerns of ChatGPT for Data Science Education

4.3. ChatGPT's Potential Impact and Support for Data Science in the Future

ChatGPT has great potential to revolutionize and advance data science because advanced natural language processing capabilities can greatly simplify and improve all aspects of the field. ChatGPT has a potentially transformative impact on data science

education, but it also raises some complex issues to ensure the integrity and effectiveness of education. Integrating ChatGPT into the curriculum can provide students with an interactive, and dynamic learning environment, and ChatGPT's ability to handle complex problems and analytical tasks can significantly reduce students' learning disabilities [15]. By providing personalized learning, ChatGPT can tailor content to students' different needs, simplifying their understanding of difficult concepts and enabling them to grasp them quickly [16]. ChatGPT simplifies the coding and analysis process, enabling students to devote more energy to strategic aspects of data science, such as interpreting data results and applying insights to the real world. This shift has the potential to improve students' critical thinking and problem-solving skills. ChatGPT's ability to generate contextual examples and simulated data scenarios also provides students with experiential learning opportunities to engage in interactive modules [16]. This reliance on AI has caused great concern among educators. The convenience provided by AI has the potential to foster an atmosphere of passive learning, making students over-reliant on automated solutions [13]. This dependence affects the development of deep analytical skills and reduces the opportunity for students to think critically. Integrating AI tools such as ChatGPT into education will require careful monitoring by schools or teachers to ensure that these technologies support rather than replace traditional teaching methods. The effectiveness and feasibility of these AI tools must be constantly evaluated to maintain a balance that enhances learning without compromising the development of independent problem-solving [17]. The ethical aspects of AI in educational Settings also deserve serious consideration, especially how these tools might affect academic integrity and the authenticity of student assignments. In conclusion, while ChatGPT holds promise for enhancing data science education by increasing accessibility and engagement in learning, its integration must be approached with caution. Educators need to emphasize and ensure that AI tools are used to complement, not replace, undergraduate education. This balanced approach will go a long way toward preparing students not only to excel academically but also to succeed professionally in the increasingly competitive data-driven environment of the future.

ChatGPT can also be used in the daily workflow of data scientists to automate routine data pre-processing and analysis tasks, thereby increasing their productivity [18] and contributing to more complex areas such as predictive modeling and decision making. ChatGPT can also quickly interpret large data sets to gain unique insights and provide preliminary analysis for further research by human experts. It also equips more professionals and students with data science skills by assisting with real-time coding and statistical analysis. This accessibility creates a more inclusive environment for the industry, where people from all backgrounds can contribute to and benefit from data-driven innovation [9]. While the impact of ChatGPT has facilitated significant advances in the way data is analyzed and used, high standards are still needed to assess the accuracy and bias of validation results and the integrity of data science practices. In the future, the role of ChatGPT in data science may be to become a powerful adjunct to human expertise, enabling smarter, faster decision-making across a wide range of domains [6].

4.4. Evaluation Methods to Assess the Impact of ChatGPT on Data Science Learning Outcomes

To fully assess the impact of ChatGPT on the academic performance of data science students, we used a variety of assessment methods, each providing a unique perspective on the learning process and subsequent research. In examining the impact of ChatGPT on the academic performance of data science students, different articles have adopted different assessment methods. These multidimensional approaches allow educators and researchers to fully understand how AI tools like ChatGPT affect the learning process for data science students. This section provides an overview of the main research methods and discusses the advantages and potential limitations of these methods. Analyzing students' final grades provides a straightforward and quantifiable way to assess the impact of ChatGPT on the academic performance of data science students [11]. Researchers can compare the impact of ChatGPT on academic performance by examining the final grades of students who used ChatGPT in their courses and students who did not use ChatGPT. This approach provides a quantifiable result to determine whether ChatGPT use improves academic performance. However, relying solely on final grades to assess educational impact faces some problems [19]. Final grades may not adequately reflect the nuances of student understanding or the complexity of the learning process. Final grades may also be influenced by several external variables such as prior knowledge base, motivation to learn, quality of teaching, and individual study habits. These factors may mask the true effects of ChatGPT, suggesting that a more detailed methodology that considers both quantitative and qualitative data is needed to fully understand how ChatGPT affects learning outcomes in data science.

Surveys and questionnaires can directly obtain information on students' perceptions of the usefulness of ChatGPT, changes in their study habits, and attitudes toward AI-assisted learning [20]. Researchers can collect qualitative and quantitative data directly from learners to measure students' satisfaction, perceived benefits, and potential drawbacks of ChatGPT. The results show that ChatGPT can affect students' academic performance. While such surveys can be good for collecting a variety of data, they rely heavily on self-reporting, which can be biased, such as students' tendency to provide socially acceptable answers, or some respondents' lack of self-awareness [21].

In-depth interviews allow for a more qualitative assessment of the impact of the ChatGPT, allowing respondents to give detailed answers and explore complex perspectives and experiences that may not have been captured by the questionnaire. These interviews can reveal how they have integrated the tool into their learning process and how often they have adopted it. While this approach can provide rich detail, it is time-consuming and may not be representative of the wider student population as sample sizes are typically small.

A more comprehensive approach to assessing the impact of ChatGPT on student achievement might be to conduct a controlled classroom experiment [17]. In this case, the researchers divided the class into a control group and an experimental group, with only the experimental group using ChatGPT. This approach allows for the observation of the direct causal effects of the ChatGPT on learning outcomes under controlled conditions [5]. Researchers can not only monitor outcomes but also study learning behaviors during the monitoring process. The challenge here is to ensure that the two groups of students are identical except for the introduction of ChatGPT, which is difficult to achieve in a typical educational setting [8].

Each assessment method has its unique strengths as well as unique challenges. Using a combination of these methods allows researchers to comprehensively evaluate the impact of ChatGPT on the academic performance of students in data science education, ensuring that the results are comprehensive and inform educational policy and practice.

Combining these methods not only improves the reliability of the findings, but also ensures that the different dimensions of learning are valid.

5. Discussion

The impact of ChatGPT on the academic performance of data science students can be viewed in terms of both weaknesses and strengths. On the one hand, this kind of online learning can give data science students an edge in understanding and mastering challenging course content through timely feedback and personalized help. On the other hand, it will also have some negative effects on students' learning attitudes and psychology. Rapidly improving understanding of these concepts can greatly help students struggling at advanced stages of data science to develop skills and gain confidence. Having said that, simple solutions may encourage a passive approach to learning, where students may focus on completing the work rather than delving into the learning process itself [8]. This may lead to feedback literacy problems, and long-term dependence may lead to a decrease in the ability and skill of students to receive, understand, analyze, evaluate, and feedback on the actions or achievements of others. The results reveal this trend: those who rely primarily on ChatGPT for homework will primarily use it to count on quick answers to questions. Long-term effects can be harmful because to some extent they will not be used to solve the problem independently and, therefore, may be out of context. These students may complete the goal of the final assignment, but they may find it difficult to connect the dots as they answer more complex or unfamiliar questions that may come up later in the exam. Interestingly, while ChatGPT can be a useful tool at times, there are situations where over-reliance on the tool causes problems, leading to gaps in conceptual understanding, which is a fundamental factor in mastering data science. Building a foundation by mastering core theory is undoubtedly decisive for long-term achievement in the field, which requires the cultivation of inquiry, independent learning, problem-solving skills, and flexible thinking methods. These problems require a hybrid learning model and the ability to take full advantage of ChatGPT. Therefore, blended learning is the best way forward. In this model, ChatGPT would be a complement to the coursework problem rather than a primary solution. According to this study, the ChatGPT application will be used to clarify difficult concepts, propose alternative solutions, and practice technical skills. However, it clearly outlines that it should be used as a complementary tool for performance rather than as a tool for self-learning and problem-solving. By suggesting that students use AI tools such as ChatGPT in their academic pursuits to support a more balanced leadership habitat, educators can foster a more balanced collection of learning. The strategy requires students to apply critical reasoning skills and develop analytical tools. It helps them avoid over-reliance on technology.

In addition, future research should focus on investigating how ChatGPT and similar AI tools affect long-term learning outcomes. There is also a need to investigate the extent to which students can retain and apply knowledge gained through AI-assisted learning, as opposed to traditional, more active learning methods. Equally important, there is a need to explore how AI tools can be integrated into the student learning process to accommodate students' self-directed learning progress, provide personalized feedback, and encourage students to engage more deeply in learning [7]. These improved systems could provide some automated assessment methods that ensure ChatGPT is used in a balanced way, addressing key issues such as students' over-reliance on AI.

Going forward, educators need to understand how AI technologies such as ChatGPT will shape higher education [11]. As AI is increasingly applied to education platforms, there is a need to explore innovative ways to incorporate it into integrated learning environments [20]. This environment not only helps students develop technical skills, but also promotes critical thinking, ethical awareness, and problem-solving resilience. Educators and educational institutions should constantly evaluate the impact of AI tools on student learning to ensure that these technologies benefit students in the long term.

In addition, another future study may explore how ChatGPT reduces anxiety in students. In the complex subject of data science, many students are under pressure to meet deadlines and excel, which can leave students in the process feeling stressed and burned out. ChatGPT's ability to provide immediate feedback and support, allowing students to solve challenging problems at their own pace, reduces stress for students [7] and reduces anxiety that students may not understand the material quickly enough or understand the meaning of the task tools can act as a safety net, giving students the confidence to tackle difficult concepts [1].

6. Conclusions

In conclusion, this paper has systematically examined the impact of ChatGPT on the academic performance of data science students. The reviewed articles collectively suggest that ChatGPT holds significant potential to enhance the educational experience within

the field of data science. By providing immediate feedback and personalized support, particularly in areas such as coding, problem-solving, and conceptual understanding, ChatGPT can greatly improve learning outcomes.

However, it is crucial to recognize the potential downsides of over-reliance on AI tools. Excessive dependence on ChatGPT may impair students' ability to think critically and solve problems independently. Therefore, we emphasize the need to use ChatGPT in a balanced manner—viewing it as a supplementary resource rather than the primary one. This balanced approach ensures that students benefit from the strengths of AI while still developing essential skills through independent effort.

Future research should explore the long-term effects of AI tools like ChatGPT on student learning outcomes. Investigating how these tools influence critical thinking, problem-solving abilities, and overall academic performance over time will be vital in understanding their full impact. By doing so, educators can better integrate AI into their teaching strategies, fostering an environment where technological advancements and traditional learning methods complement each other to enhance educational experiences.

References

- [1] Shen, Y., Syamkumar, (2024). Implications of ChatGPT for Data Science Education. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024)*, Portland, OR, USA, March 2024, pp. 1–7. <https://doi.org/10.1145/3626252.3630874>
- [2] Lo, C. K., Hew, K. F., & Jong, M. S. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers & Education*, 219, Article 105100. <https://doi.org/10.1016/j.compedu.2024.105100>
- [3] Joshi, I., et al. (2024). ChatGPT in the Classroom: An Analysis of Its Strengths and Weaknesses for Solving Undergraduate Computer Science Questions. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education*, Vol. 1, Portland, OR, USA, March 2024, pp. 1–7. <https://doi.org/10.1145/3626252.3630803>
- [4] 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. <https://doi.org/10.1080/26939169.2023.2223609>
- [5] Zheng, Y. (2023). ChatGPT for Teaching and Learning: An Experience from Data Science Education. In *Proceedings of the 24th Annual Conference on Information Technology Education (SIGITE '23)*, Marietta, GA, USA, October 2023, pp. 1–7. <https://doi.org/10.1145/3585059.3611431>
- [6] Hassani, H., & Silva, E. S. (2023). The Role of ChatGPT in Data Science: How AI-Assisted Conversational Interfaces Are Revolutionizing the Field. *Big Data and Cognitive Computing*, 7(2), 62. <https://doi.org/10.3390/bdcc7020062>
- [7] Gelvanovsky, G., & Saduov, R. (2024). The Impact of ChatGPT on Academic Writing Instruction for Computer Science Students. *Preprint*. <https://doi.org/10.13140/RG.2.2.30899.40481>
- [8] López-Fernández, D., & Vergaz, R. (2024). Adoption and Impact of ChatGPT in Computer Science Education: A Case Study on a Database Administration Course. *arXiv Preprint arXiv:2407.12145*. <https://doi.org/10.48550/arXiv.2407.12145>
- [9] Chinsook, K., et al. (2022). Big Data in Higher Education for Student Behavior Analytics. *Higher Education Studies*, 12(1), 105–112. <https://doi.org/10.5539/hes.v12n1p105>
- [10] Richards, M., et al. (2024). Bob or Bot: Exploring ChatGPT's Answers to University Computer Science Assessment. *ACM Transactions on Computing Education*, 24(1), Article 5. <https://doi.org/10.1145/3633287>
- [11] Wang, T., Vargas Diaz, D., Brown, C., & Chen, Y. (2023). Exploring the Role of AI Assistants in Computer Science Education: Methods, Implications, and Instructor Perspectives. In *2023 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. <https://doi.org/10.1109/VL-HCC57772.2023.00018>
- [12] Xing, Y. (2024). Exploring the Use of ChatGPT in Learning and Instructing Statistics and Data Analytics. *Teaching Statistics*, 46, 95–104. <https://doi.org/10.1111/test.12367>
- [13] Ahn, S. (2024). Data Science through Natural Language with ChatGPT's Code Interpreter. *Translational and Clinical Pharmacology*, 32(2), 73–82. <https://doi.org/10.12793/tcp.2024.32.e8>
- [14] Humble, N., Boustedt, J., Holmgren, H., Milutinovic, G., Seipel, S., & Östberg, A.-S. (2023). Cheaters or AI-Enhanced Learners: Consequences of ChatGPT for Programming Education. *Electronic Journal of e-Learning*, 22(2), 16–29. <https://doi.org/10.34190/ejel.21.5.3154>
- [15] Singh, H., Tayarani-Najaran, M.-H., & Yaqoob, M. (2023). Exploring Computer Science Students' Perception of ChatGPT in Higher Education: A Descriptive and Correlation Study. *Education Sciences*, 13, 924. <https://doi.org/10.3390/educsci13090924>
- [16] Gelvanovsky, G., & Saduov, R. (2024). The Impact of ChatGPT on Academic Writing Instruction for Computer Science Students. *Preprint*. <https://doi.org/10.13140/RG.2.2.30899.40481>
- [17] Kosar, T., Ostojic, D., Liu, Y. D., & Mernik, M. (2024). Computer Science Education in ChatGPT Era: Experiences from an Experiment in a Programming Course for Novice Programmers. *Mathematics*, 12, 629. <https://doi.org/10.3390/math12050629>
- [18] Narayanan Valli, L., Sujatha, N., Mech, M., & L. V. S. (2024). Exploring the roles of AI-Assisted ChatGPT in the field of data science. *E3S Web of Conferences*, 491, Article 01026. <https://doi.org/10.1051/e3sconf/202449101026>
- [19] Budhiraja, R., Joshi, I., Akolekar, H., Challa, J. S., & Kumar, D. (2024). It's not like Jarvis, but it is pretty close: Examining ChatGPT's Usage among Undergraduate Students in Computer Science. In *Proceedings of the Australian Computing Education Conference (ACE 2024)*, Sydney, NSW, Australia, 2024, pp. 1–10. <https://doi.org/10.1145/3636243.3636257>
- [20] Mendonça, N. C. (2024). Evaluating ChatGPT-4 Vision on Brazil's National Undergraduate Computer Science Exam. *ACM Transactions on Computing Education*, 9(6), Article 3674149. <https://doi.org/10.1145/3674149>
- [21] Hassan, M. M., Knipper, A., & Santu, S. K. K. (2023). ChatGPT as your Personal Data Scientist. *ACM*, 1(1), 1–35.