Impact of AI Language Models on the Academic Performance of Afghan Students: A Case Study

Zabihullah Burhani[1*]; Abdul Ameen Akbari[2]; Buhrudin Azizi[2]

[1]*Department of Computer Science, Faculty of Education, Takhar University, Afghanistan*
[2]*Department of Geography, Faculty of Education, Takhar University, Afghanistan*
Corresponding Email: zabihullahburhani@gmail.com, Phone Number: +93705002913

Abstract

This study investigates the impact of AI language models on the academic performance of students at Takhar University. Using a sample of 164 students, the research evaluates students' familiarity with and use of these models and their effect on academic outcomes. Results indicate that students have limited familiarity with these models and that their use did not significantly impact their academic performance. Statistical analyses, including the use of SPSS and a one-sample t-test, revealed that the introduction and use of AI language tools did not lead to notable changes in grades or skills. The results suggests that targeted educational and awareness programs are essential for harnessing the potential benefits of AI language models. Enhancing knowledge and practical skills in this area is expected to improve academic performance and productivity in scientific and practical projects.

Keywords: AI Language Models, Student Performance, Academic Achievement, Educational Technology.

Introduction

Artificial intelligence has advanced significantly since its inception and has evolved into a widespread technology used in various fields, including education. The emergence of language models like ChatGPT has captured the attention of both everyday users and researchers. These models have been able to address the challenges of modern educational environments as educational assistants and require the definition of appropriate performance indicators (Ivanova et al., 2024; Sage & Matteucci, 2024). In recent years, the use of AI language models in education and learning has surged significantly. These models possess the capability to generate and process natural language (Weng & Chiu, 2023), and they can profoundly impact students' academic performances. Models such as ChatGPT play a crucial role in increasing access to information(Lund & Wang, 2023), providing educational support (Wibowo et al., 2023), empowering self-directed learning(Fazil et al., 2024), delivering personalized learning experiences (Božić & Poola, 2023; Fuchs, 2023), and developing cognitive skills (Batool & Herrera, 2024; Tabib & Alrabeei, 2024), thereby enhancing the quality of education (Rane, 2023). Afghanistan, as a country facing a shortage of educational resources, stands to

benefit greatly from these technologies by enhancing its educational landscape. This research examines both the positive and negative impacts of these technologies on Afghan students' academic performance and proposes strategies for the optimal use of these tools in Afghanistan's educational system. The primary objective of this study is to understand the level of familiarity, usage, and effects of this novel technology on the learning process and academic achievements of Afghan students. This research delves into the students' experiences of students, analyzes the challenges and opportunities associated with the use of AI language models in Afghanistan's educational system, and offers solutions for improving the effectiveness and efficiency of these technologies.

Many studies have been conducted on this title both inside and outside Afghanistan; however, our study will focus specifically on the students at Takhar University. In 2024, Fazil and colleagues at Kabul University conducted research on the impact of AI technologies on the engagement and academic performance of Afghan students. This study employed regression analysis and ANOVA, involving 200 students from various faculties to comprehensively examine the dimensions of AI integration. The results indicate that students possess a good understanding of AI, but there is a need to improve its integration into education and to pay attention to ethical considerations (Fazil et al., 2024).

A study conducted in 2023 focused on enhancing students' critical thinking skills and AI literacy by examining the use of ChatGPT in higher education. This experimental research included both experimental and control groups. The results indicated that ChatGPT enhances students' analytical abilities and independent learning; however, challenges such as reduced critical thinking and dependency on technology were also noted. It was recommended that ChatGPT be used as a complementary tool in conjunction with appropriate educational policies (Huang et al., 2023).

The 2024 study by Zahara and colleagues examined the positive and negative effects of ChatGPT on student learning at Gadjah Mada University. The results showed that ChatGPT aids in improving student interaction, engagement, and critical thinking skills; however, it also highlighted issues like over-reliance on technology and reduced motivation for independent learning. This article suggests using ChatGPT as a supplementary tool alongside appropriate educational policies and strategies, emphasizing a balanced and evidence-based approach (Zahara et al., 2024).

The 2023 research "The Influence of ChatGPT and AI Tools on Academic Performance" by Nurtayeva and colleagues investigated the impact of AI tools on students' academic performance. This study was conducted at universities in Kazakhstan using online questionnaires and semi-structured interviews. The findings revealed that AI tools help improve students' writing and research skills, however, there are concerns about dependency on technology and the authenticity of students' work. The researchers emphasized the importance of appropriate educational policies and strategies (Nurtayeva et al., 2023).

The 2023 paper titled "The Influence of ChatGPT on Student Learning and Academic Performance" by Ayman and colleagues explored the impact of ChatGPT on student learning at universities in Egypt. The findings indicate that ChatGPT enhances students' analytical and research skills and improves student interactions though it also raises concerns about technology dependency and ethical issues. The researchers emphasized the need for appropriate policies (Ayman et al., 2023).

The study "The Impact of ChatGPT Use on the Quality of Academic Support for Students" by Wibowo and colleagues (2023) examined the effect of ChatGPT on academic support. Involving 150 students from Indonesian universities, the study revealed that ChatGPT improved the quality of academic support and student satisfaction. However, the need for proper training and attention to feedback was also identified (Wibowo et al., 2023).

The article "Effects of ChatGPT on Students' Academic Performance: Mediating Role of Prompt Engineering" by Shehri and colleagues (2023) explored the impact of ChatGPT and prompt engineering. The study, which involved 200 students, demonstrated that ChatGPT enhances academic performance and that prompt engineering plays a crucial role. Challenges such as the need for effective question design were identified (Shehri et al., 2023).

Material and Method

This study employed a descriptive-analytical research method. The primary data collection tool was a questionnaire designed to assess the topics under investigation. The statistical population consisted of all currently enrolled students at Takhar University. A simple random sampling method was used to select the sample, resulting in 164 students participating in the study. The reliability of the questionnaire was assessed using Cronbach's alpha, which yielded a value of 0.843 across 6 items. This indicates a high level of internal consistency and suggests that the questionnaire is a reliable tool for measuring the intended variables. For data analysis, SPSS software and a one-sample t-test were used. Population: population for this study comprised all students at Takhar University, totaling 5,774 individuals. This information was obtained from the Students Affairs Office of Takhar University and served as the basis for the target population.

Sample
From the total population, a sample of 164 students was selected using simple random sampling. Due to time constraints, no specific formula was applied to determine the sample size. The sample was chosen to represent a diverse range of students at different academic levels and with varying degrees of familiarity with AI language models.

Data Collection Instrument
The research tool was a researcher-designed questionnaire aimed at assessing students' familiarity with and use of AI language models. After the initial distribution, several questions were removed due to a lack of internal consistency with other items, resulting in a final questionnaire consisting of 6 items. The questions were designed using a five-point Likert scale, ranging from "Very Low" to "Very High."

Validity and Reliability
The content validity of the questionnaire was assessed by consulting experts in the fields of AI and higher education. To ensure the reliability of the instrument, Cronbach's Alpha was calculated, yielding a value of 0.851, indicating an acceptable level of reliability.

Data Collection Process
Data were collected in two phases: in-person and online. Students who could attend the university were given the questionnaire in person, while others received a link to the online version of the survey.

Normality of Data Distribution
To assess the normality of data distribution, the Shapiro-Wilk test was conducted. The results indicated that the data distribution was not significantly different from normal (p = 0.082). This suggests that the data are approximately normally distributed and suitable for parametric statistical tests.
Data Analysis :Data analysis was performed using SPSS, employing a one-sample t-test to examine the impact of AI language models on students' academic performance.

Results

A total of 164 respondents participated in the study. The average age of respondents was 21.73 years, with ages ranging from 18 to 30 years. The respondents categorized into three major fields of study. Such as, 76 respondents (46.3%) were studying in Natural Sciences, 66 respondents (40.2%) we're studying Humanities, and 22 respondents (13.4%) were in Islamic Sciences. The results indicate that the majority of respondents were in the fields of Natural Sciences and Humanities, while the Islamic Sciences field had the smallest representation. The majority of respondents were in their fourth semester (62 students, 37.8%) and second semester (51 students, 31.1%). Fewer students were in the sixth semester (27 students, 16.5%) and the eighth semester (23 students, 14.0%). This distribution indicates that the highest number of respondents were in the second and fourth semesters, with fewer representatives from in the other semesters. Among the 164 participants, 43.3% rated their internet access as average. 24.4% reported having very limited internet access, and 20.7% indicated low internet access. Additionally, 9.1% of students had high internet access, while only 2.4% had very high Internet access. These results suggest that most students have moderate internet access, with very high access being rare. 19.5% of participants rated their familiarity with AI language models like ChatGPT as very low. 21.3% reported low familiarity, 37.2% reported average familiarity, 15.9% reported high familiarity, and 6.1% reported very high familiarity. These results suggest that most students had an average level of familiarity with these models, while a small percentage have a high level of familiarity with the technology. Among the 164 participants, 82.3% use ChatGPT the most frequently. 13.4% used Copilot, and 4.3% used Bing Chat. These results indicate that ChatGPT was the most popular AI language model among students, while the use of other models was less common. 37.8% of participants reported an average use of AI language models in their projects and household tasks. 26.8% reported low use, and 21.3% reported very low use. Additionally, 14.0% reported high used of these models. These results indicate that most students have an average level of use of AI language models in their activities, whereas very low and high usage is less

common. The descriptive statistics for the examined indicators are provided in the table below.

According to table 1 that the mean impact of AI language models on academic grades is 2.55, indicating a moderate effect. AI models had a mean score of 3.06 in assisting with writing papers and projects, showing a relatively positive impact. The effect on writing skills was 2.66, and the assistance in learning foreign languages was 2.68, both indicating moderate effects. For solving technical issues and coding, the mean score was 2.58, while the reduction in time spent on academic projects and household tasks was 3.00, reflecting a moderate impact of the models in these areas. For inference from the sample and generalization to the population, a one-sample t-test was used.

Table 1. Descriptive Statistics

	Mean	Std. Deviation
The impact of AI language models on academic grades.	2.5549	1.11462
The assistance provided by AI language models in writing papers and projects.	3.0610	1.06647
The influence of AI language models on writing skills.	2.6341	1.07979
The contribution of AI language models to learning foreign languages.	2.6768	1.25290
The role of AI language models in solving technical problems and coding.	2.5854	1.21293
The reduction in time spent on academic projects and household tasks due to AI language models.	3.0000	1.13505
Valid N (listwise)		

In the one-sample t-test, the test value was set at 3 to evaluate whether the mean impact of using AI language models on students' academic performance significantly differed from this value. The calculated t-value is -3.703 with 163 degrees of freedom, and the significance level (p-value) is 0.000, which is less than the conventional alpha level of 0.05. This indicates that the mean impact is significantly different from 3. The mean difference is -0.24797, with a 95% confidence interval ranging from -0.3802 to -0.1157. This means that, on average, the impact of using AI language models is significantly less than the value of 3, suggesting that students perceive a lower impact on their academic performance than the test value.

Table 2. One-Sample Test

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Overall Mean	-3.703	163	.000	-.24797	-.3802	-.1157

Discussion

The lack of familiarity and limited usage have led to a minimal impact of AI language models on students' academic performances. The results indicate that these tools did not significantly affect academic grades or other aspects of students' performances. It appears that insufficient use of AI language models, due to inadequate awareness and limited engagement, has resulted in no significant changes in students' academic outcomes. These

findings underscore the need for increased education and promotion of AI tools among students to achieve notable improvements in academic performance and other areas.

Conclusion

The results of this study indicate that the familiarity with and use of AI language models among students at Takhar University students remains at a low level and has not significantly impacted their academic performance. Given that the limited use of these tools has not resulted in substantially changed in academic grades or other skills, it is recommended that educational programs and awareness campaigns about the benefits of AI language models should be implemented. Enhancing awareness and practical skills in this area could contribute to improved academic performance and increased efficiency in conducting scientific and practical projects.

Data Availability Statement:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgment

Data Availability Statement:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ayman, S. E., El-Seoud, S. A., Nagaty, K., & Karam, O. H. (2023). The Influence of ChatGPT on Student Learning and Academic Performance. 2023 International Conference on Computer and Applications (ICCA), 1–5. https://doi.org/10.1109/ICCA59364.2023.10401713

Batool, Y., & Herrera, F. (2024). Cognitive Skills Development with ChatGPT: Enhancing Pedagogics Design through AI Tools. https://doi.org/10.13140/RG.2.2.20550.89926

Božić, V., & Poola, I. (2023). Chat GPT and education. https://doi.org/10.13140/RG.2.2.18837.40168

Fazil, A. W., Hakimi, M., Shahidzay, A. K., & Hasas, A. (2024). Exploring the Broad Impact of AI Technologies on Student Engagement and Academic Performance in University Settings in Afghanistan. RIGGS: Journal of Artificial Intelligence and Digital Business, 2(2), Article 2. https://doi.org/10.31004/riggs.v2i2.268

Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is Chat GPT a blessing or a curse? Frontiers in Education, 8, 1166682.

https://doi.org/10.3389/feduc.2023.1166682

Huang, C.-W., Coleman, M., Gachago, D., & Van Belle, J.-P. (2023, November 29). Using ChatGPT to Encourage Critical AI Literacy Skills and for Assessment in Higher Education. https://doi.org/10.1007/978-3-031-48536-7_8

Ivanova, T., Staneva, A., Borissova, D., & Rasheva-Yordanova, K. (2024, February 27). Chat GPT Performance Evaluation Model for Learning. https://doi.org/10.1007/978-3-031-54327-2_15

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? Library Hi Tech News, 40(3), 26–29. https://doi.org/10.1108/LHTN-01-2023-0009

Nurtayeva, T., Salim, M., Basheer, T., & Khalilov, S. (2023). The influence of ChatGPT and AI Tools on the Academic Performance. 22, 247–258. https://doi.org/10.37896/YMER22.06/26

Rane, N. (2023). Enhancing the Quality of Teaching and Learning through ChatGPT and Similar Large Language Models: Challenges, Future Prospects, and Ethical Considerations in Education. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4599104

Sage, R., & Matteucci, R. (2024). Technology and Learning: Issues Vital to Address. BRILL.

Shehri, F. A., Maham, R., Malik, A., & Saif, O. B. (2023). Effects of ChatGPT on Students Academic Performance:Mediating Role of Prompt Engineering. The Asian Bulletin of Big Data Management, 3(2), 137–147. https://doi.org/10.62019/abbdm.v3i2.58

Tabib, F. M., & Alrabeei, M. M. (2024). Can Guided ChatGPT Use Enhance Students' Cognitive and Metacognitive Skills? https://doi.org/10.1007/978-3-031-52280-2_10

Weng, X., & Chiu, T. K. F. (2023). Instructional design and learning outcomes of intelligent computer assisted language learning: Systematic review in the field. Computers and Education: Artificial Intelligence, 4, 100117. https://doi.org/10.1016/j.caeai.2022.100117

Wibowo, G. A., Rahman, A., & Anis, M. (2023). The Impact of ChatGPT Use on the Quality of Academic Support for Students. Technology and Society Perspectives (TACIT), 1(3), 132–138. https://doi.org/10.61100/tacit.v1i3.69

Zahara, A., Hasanah, F., Habib, H., & Sabarrudin. (2024). The Influence of Advancements in AI Chat GPT Technology in Students' Learning Process. BICC Proceedings, 2, 132–137. https://doi.org/10.30983/bicc.v1i1.117