Karshi Multidisciplinary International Scientific Journal

www.kmisj.reapress.com

K. Multidiscip. Int. Sci. J. Vol. 2, No. 2 (2025) 60–64.

Paper Type: Original Article

Artificial Intelligence in E-Learning: An Empirical Study of ChatGPT as a Learning Tool

Irma Guga[1], Aurora Binjaku[2,*]

[1] Department of Engineering, European University of Tirana, Tirana, Albania; Irma.guga@uet.edu.al.
[2] Department of Engineering, European University of Tirana, Tirana, Albania; aurora.binjaku@uet.edu.al.

Citation:

Received: 23 December 2024 *Revised: 12 March 2025* *Accepted: 01 May 2025*	Guga, I., & Binjaku, A. (2025). Artificial intelligence in e-learning: An empirical study of ChatGPT as a learning tool. *Karshi multidisciplinary international scientific journal,* 2(2), 60-64.

Abstract

Artificial Intelligence (AI) is transforming not only traditional learning processes but also e-learning methods. This paper aims to statistically analyze the impact of AI tools on the academic performance of higher education students. An experimental study was conducted with active students of the European University of Tirana. Students participated in a structured learning session utilizing ChatGPT to engage with specific concepts from the curriculum, followed by an assessment designed to evaluate their comprehension and knowledge acquisition. The sample yielded 26 valid test responses. The data gathered through this process is used to measure the impact of AI tools on learning performance. The students' actual average grade was also treated as a controlled variable and measured as part of the study. The study aims to determine if there is a relationship between learning performance through ChatGPT and overall academic performance. The findings yield descriptive statistical measurements, concluding that learning performance is not very high when using an AI tool. The data were transformed into categorical variables, and a contingency table was used to conduct a test of independence to determine whether students' academic achievement influences learning performance facilitated by an AI tool. Additionally, a correlation analysis is conducted to examine the potential relationship between variables. The findings indicate that the variables are independent, suggesting that other factors may affect learning performance. The study is of high importance for transforming the traditional learning process in the education field.

Keywords: E-learning, Artificial intelligence, ChatGPT, Education.

1 | Introduction

The use of Artificial Intelligence (AI) and digital education has initiated necessary transformations in teaching and learning processes. E-learning platforms, such as Learning Management Systems (LMS), have traditionally provided access to educational content, assessment tools, and progress monitoring. However, recent advancements in AI technologies have introduced new opportunities for personalization, interactivity, and

Corresponding Author: urora.binjaku@uet.edu.al
https://doi.org/10.22105/kmisj.v2i2.89

learner autonomy. This study investigates the potential of AI tools, specifically ChatGPT, as integral components of e-learning systems. It seeks to evaluate whether these tools can not only support but also potentially replace traditional instructional approaches, thereby redefining the educator-learner dynamic. The integration of AI into e-learning environments has garnered increasing scholarly interest, particularly regarding its potential to transform traditional educational practices. While existing research has primarily focused on students' perceptions of AI tools such as ChatGPT, the present study seeks to move beyond perception and assess the extent to which such technologies can become alternatives to conventional instructional methods. The purpose of this paper is to conduct an empirical analysis of the learning results achieved by using ChatGPT in higher education institutions. Furthermore, it aims to find correlations and relationships between the students' performance, measured by their average grade, and their learning results with ChatGPT; hence, we aim to contribute to understanding AI's role in the contemporary education process.

2|Literature Review

AI is a central topic in contemporary discourse. Its spread is massive in every possible area, ranging from manufacturing and tourism to health and care, as well as teaching and learning. Many see it as a much-needed achievement for humanity, but on the other hand, it is equally "dangerous" and "devouring" jobs, leading to increasing unemployment levels.

Until the early 21st century, the educational model remained largely uniform, characterized by a traditional classroom environment in which a teacher directed the learning process. However, the advancement of computer technology has significantly reshaped the educational landscape, introducing new methods of instruction and access to knowledge. Online learning, also known as e-learning, is a teaching method that has gained significant popularity recently, particularly following the COVID-19 pandemic. The need for continuity during those challenging times led to a shift in the entire international education system towards online learning. Even though it is years away from the pandemic, e-learning continues to remain a widely used and popular method, especially in universities. The pros and cons of each method lead to an important question: whether it is better to teach face-to-face/classically, online, or in a hybrid format. While some argue that the abrupt and unplanned shift to online learning—often implemented without adequate training, limited internet bandwidth, and minimal preparation—may result in a suboptimal user experience that hinders long-term educational progress, others contend that this disruption could catalyze innovation.

Another concern is that instructors may find it difficult to assess students' learning status. of their learners, raising questions about the quality of e-learning [1]. Interaction in online learning programs promotes student-centered learning, encourages broader student participation, and yields more in-depth and reasoned discussions than traditional face-to-face programs [2]. Additionally, Warschauer advocates for interaction in online environments, as there is less opportunity for intimidation between individuals and less time pressure on them compared to face-to-face settings [3].

Cavanaugh and Jacquemin [4] in their work conclude that there is little to no difference in grade-based student performance between instructional modes for face-to-face and online courses. Using a dataset from five online economics and five online finance courses, it was found that the higher the students' GPA and the more time they spent on online coursework, the better their performance in online formats [5]. Another study concludes that 88.2% of respondents use virtual assistants (e.g., ChatGPT, Siri, Google Assistant, etc.) [6].

3|Research Methodology

Data are collected through a multinomial experimental study. First, the variable of interest was identified, and then the other independent variable was measured. After gathering the data, it is displayed in a contingency table, and an independence test is conducted to determine if the variables are independent or not. Descriptive statistics are used to gather information about the distribution of the data and to obtain general indicators of the variables. Correlation analysis is used to get insight into the relationship among variables. The

experimental units consisted of higher education students, with 14 students participating in the experiment once and six students attending it twice, resulting in a total of 26 results. They are first introduced to some key concepts in the statistics course they are attending. Students were given a limited amount of time to learn about the concepts in ChatGPT in class. Afterwards, a test was administered to answer theoretical and practical questions regarding the concepts. The authors graded the tests without knowing the students' names. Tests had a minimum of 0 points and a maximum of 8 points. At the end of the tests, the average grade of each student was recorded as the controlled variable, and the ChatGPT results were recorded as the variable of interest.

4 | Results

The data are statistically analyzed using SPSS software. First, numerical data are imported and are transformed into new string variables. They are coded as follows:

Average grade–numerical data of the student's average grade until the moment the experiment was conducted. It measures the students' academic performance.

ChatGPT_result_gen represents the numerical data of the test results transformed into a 0-10 grade scale. It measures the performance of ChatGPT as an effective AI tool for use in the learning process.

Avg_categories – scale type data after transforming the Average_grade variable into two categories

0-8.599-->"Other" and 8.6-10-->"Excellent". The Excellent category is grouped with "Good" and "Excellent" results.

ChatGpt_results_categories-scale type data after transforming ChatGPT_result_gen variable into two categories 0-4.099-->"Fail" and 4.1-10-->"Pass"

For the scale variables, there is a grading system: 41/100 Pass, 86-90 Good, and 91-100 Excellent.

Descriptive indicators give insight into the distribution of the data, whereas the sample is considered heterogeneous in terms of the controlled variable. The average value of the results, 5.93850, from a maximum value of 10, doesn't indicate high learning performance.

Table 1. Descriptive indicators of variables generated from SPSS.

Descriptive Statistics					
	N	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Average_grade	26	6.130000000000	10.00000000000	8.625769230769	.9244162399132
ChatGPT_results_gen	26	1.250	10.000	5.93750	2.699826
Valid N (listwise)	26				

Testing the independence of ChatGPT results and Average Grade requires transforming the data into categorical form. ChatGPT results are categorized into two groups: students who "Passed" the test and those who "Failed" the test, thereby measuring the performance of ChatGPT as an AI tool in e-learning. Average results are categorized into two groups: students who are "Excellent and Good" in terms of grades and "Other," while students with an average of less than 8.6 are grouped.

The chi-square test cannot be applied as some of the expected observations are less than 5 under the assumption of independence among categories. In this case, Fisher's exact test is conducted despite the low number of proportions and samples. This test is widely advised in cases when the sample size is small. The exact probability test devised by Fisher, Irwin, and Yates provides a way out of the difficulty when the numbers in a fourfold table are too small for the chi-square distribution [7]. Null hypothesis: the relative proportions of ChatGPT results are independent of the relative proportions of Average_grade. Failure to reject the null hypothesis means that the fact that students pass or fail by using ChatGPT depends on their average grade, which in turn depends on their overall academic performance.

Table 2. Crosstabulation table generated from SPSS.

			ChatGpt_Results_Categories		Total
			Fail	Pass	
Avg_categories	Excellent	Count	4	10	14
		Expected Count	4.8	9.2	14.0
		% within Avg_categories	28.6%	71.4%	100.0%
		% within ChatGpt_results_categories	44.4%	58.8%	53.8%
		% of Total	15.4%	38.5%	53.8%
	Other	Count	5	7	12
		Expected Count	4.2	7.8	12.0
		% within Avg_categories	41.7%	58.3%	100.0%
		% within ChatGpt_results_categories	55.6%	41.2%	46.2%
		% of Total	19.2%	26.9%	46.2%
Total		Count	9	17	26
		Expected Count	9.0	17.0	26.0
		% within Avg_categories	34.6%	65.4%	100.0%
		% within ChatGpt_results_categories	100.0%	100.0%	100.0%
		% of Total	34.6%	65.4%	100.0%

Table 2 indicates that there is no distinctive difference in proportions among variables. Fisher's exact test results are shown in *Table 3*.

Table 3. Independence test results from SPSS

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	0.490[a]	1	0.484		
Continuity Correction[b]	0.082	1	0.775		
Likelihood Ratio	0.490	1	0.484		
Fisher's Exact Test				0.683	0.387
N of Valid Cases	26				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 4.15.
b. Computed only for a 2x2 table

Fisher's exact test p-value of 0.683 is greater than the 5% significance level, indicating that it is not statistically significant. The null hypothesis cannot be rejected. Students' academic performance doesn't affect the overall performance of the learning process through ChatGPT.

Table 4 presents the results of the correlation among variables, indicating a correlation of 0.157, which suggests that there is no dependency among the variables.

Table 4. Correlation results from SPSS

		Average_Grade	ChatGPT_Results_Gs
Average_grade	Pearson Correlation	1	0.157
	Sig. (2-tailed)		0.444
	N	26	26
ChatGPT_results_gs	Pearson Correlation	0.157	1
	Sig. (2-tailed)	0.444	
	N	26	26

5 | Conclusion

The paper aimed to give statistical results and analysis regarding the general performance of ChatGPT in the learning process of higher education students. Results show that 65.4% of the students passed the tests after using an AI chatbot. Interestingly, 28.6% of good-performing students failed the tests, indicating that certain factors affect this learning process. Furthermore, the independence test and correlation analysis indicated that

there is no relationship between students' overall academic performance and their performance on the learning process using ChatGPT. However, this study is limited by the small number of experimental units. And results, thus it is recommended to use its descriptive results. It is recommended that further studies expand the field of study by increasing the sample size and the range of subjects. Additionally, this study is limited to students' academic performance as the independent variable. Studies should focus on other variables that may affect the learning process through AI tools.

Author Contribution

Conceptualization, I.G. and A.B.; methodology, I.G. and A.B.; literature review, I.G.; experiment design, I.G. and A.B.; data maintenance, I.G. and A.B.; statistical analysis and conclusions, I.G. and A.B.; writing-creating the initial design, I.G. and A.B.; writing-reviewing and editing, I.G.. All authors have read and agreed to the published version of the manuscript.

Data Availability

Data available on request due to ethical reasons

Conflicts of Interest

The authors declare that they have no conflict of interest. "Funders played no role in the design of the study, in the collection, analysis, or interpretation of the data, in the writing of the manuscript, or in the decision to publish the results."

References

[1] Qu, S., Li, K., Wu, B., Zhang, X., & Zhu, K. (2019). Predicting student performance and deficiency in mastering knowledge points in MOOCs using multi-task learning. *Entropy*, *21*(12), 1216. https://doi.org/10.3390/e21121216

[2] Karayan, S. S., & Crowe, J. A. (1997). Student perceptions of electronic discussion groups. *The journal*, *24*(9), 69–71. https://thejournal.com/articles/1997/04/01/student-perceptions-of-electronic-discussion-groups.aspx?Page=2

[3] Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *The modern language journal*, *81*(4), 470–481. https://doi.org/10.1111/j.1540-4781.1997.tb05514.x

[4] Cavanaugh, J., & Jacquemin, S. J. (2015). A large sample comparison of grade based student learning outcomes in online vs. face-to-face courses. *Online learning*, *19*(2). https://corescholar.libraries.wright.edu/biology/826/

[5] Calafiore, P., & Damianov, D. S. (2011). The effect of time spent online on student achievement in online economics and finance courses. *The journal of economic education*, *42*(3), 209–223. https://www.tandfonline.com/doi/abs/10.1080/00220485.2011.581934

[6] Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Education sciences*, *15*(3), 343. https://doi.org/10.3390/educsci15030343

[7] Swinscow, T. D. V., Campbell, M. J., & others. (2002). *Statistics at square one*. Bmj London. http://publish.uwo.ca/~gzou2/Stats_at_Square1.pdf