THE FUTURE OF EDUCATION: PERSONALIZED LEARNING THROUGH ADAPTIVE INTELLIGENT TUTORING SYSTEMS WITH NATURAL LANGUAGE AND DEEP LEARNING

IL FUTURO DELL'EDUCAZIONE: APPRENDIMENTO PERSONALIZZATO ATTRAVERSO ADAPTIVE INTELLIGENT TUTORING SYSTEMS MEDIANTE NATURAL LANGUAGE E DEEP LEARNING

Umberto Barbieri
Pegaso Telematic University
umberto.barbieri@unipegaso.it

Emanuele Marsico
Pegaso Telematic University
University of Foggia
emanuele.marsico@unipegaso.it

Luigi Piceci
l.piceci@gmail.com

Raffaele Di Fuccio
Pegaso Telematic University
raffaele.difuccio@unipegaso.it

Francesco Peluso Cassese
Pegaso Telematic University
francesco.pelusocassese@unipegaso.it

Double Blind Peer Review

Citazione
Barbieri, U., Marsico, E., Piceci, L., Di Fuccio, R., & Peluso Cassese, F. (2024). The future of education: personalized learning through adaptive intelligent tutoring systems with natural language and deep learning. *Giornale Italiano di Educazione alla Salute, Sport e Didattica Inclusiva*, 8(2), Edizioni Universitarie Romane.

Doi:
https://doi.org/10.32043/gsd.v8i3.1179

gsdjournal.it
ISSN: 2532-3296
ISBN 978-88-7730-494-0

ABSTRACT
The present study fits into the context of Adaptive Learning (AL) and Intelligent Tutor Systems (ITS), focusing on the effectiveness of an Artificial Intelligence (AI) conversational agent capable of emulating a virtual tutor, in optimizing students' learning processes. In a controlled context, we performed a pilot study on a sample of 28 students that was divided into an experimental group, which interacted with the AI system called Albert, and a control group, which followed a conventional learning approach with a human tutor support in presence. The chatbot was developed using Deep Learning models with multi-layered neural networks, allowing for personalized interaction with students. Statistical analysis shows that the experimental group achieved higher scores in all tests to evaluate the knowledge learned, with significantly reduced task completion time. These results indicate the promising role of the virtual tutor as an innovative tool to foster student learning, underlining the importance of incorporating advanced technologies into educational contexts.

Lo studio presentato si inserisce nel contesto delle ricerche in ambito Adaptive Learning (AL) e Intelligent Tutor Systems (ITS), concentrandosi sull'efficacia di un agente conversazionale di Intelligenza Artificiale (IA) in grado di emulare un tutor virtuale nell'ottimizzazione dei processi di apprendimento degli studenti. In un contesto controllato, abbiamo eseguito uno studio pilota su un campione di 28 studenti che è stato diviso in un gruppo sperimentale, che ha interagito con il sistema IA chiamato Albert, e un gruppo di controllo, che ha seguito un approccio di apprendimento convenzionale con un tutor umano in presenza. Il chatbot è stato sviluppato utilizzando modelli di Deep Learning con reti neurali multistrato, consentendo un'interazione personalizzata con gli studenti. L'analisi statistica mostra che il gruppo sperimentale ha ottenuto punteggi più alti in tutti i test per valutare le conoscenze apprese, con tempi di completamento delle attività significativamente ridotti. Questi risultati indicano il promettente ruolo del tutor virtuale come strumento innovativo per favorire l'apprendimento degli studenti, sottolineando l'importanza di integrare le tecnologie avanzate nei contesti educativi.

KEYWORDS
Adaptive Learning, Personalized Learning, Intelligent Tutor Systems, Deep Learning, Natural Language
Apprendimento adattivo, Apprendimento Personalizzato, Intelligent Tutor Systems, Deep Learning, Linguaggio Naturale

Received 11/05/2024
Accepted 18/06/2024
Published 24/06/2024

Introduction

In the context of educational neuroscience, the evolution of innovative technologies is profoundly restructuring traditional paradigms of education and learning. Modern learning platforms and Artificial Intelligence (AI) tools are opening up new perspectives for personalization and optimization of the educational experience (Yu et al., 2020). However, despite significant advancements in the field of Adaptive Learning (AL) systems and Intelligent Tutoring Systems (ITS), research has often focused on task difficulty modulation and post-execution feedback, sidelining the exploration of approaches capable of vertically adapting to individual learner characteristics in line with learning objectives (Mousavinasab, 2021; Jing, 2023). In response to this gap, this study focuses on the effectiveness of an AI conversational agent capable of emulating a virtual tutor, integrated into a Chatbot. The core of the investigation is to analyze how interaction with such a Chatbot can influence the process of knowledge acquisition, reinforcement, and application by students. The central objective of this pilot study is to test the effectiveness of a tool that embodies both the flexibility needed to adapt to their individual needs and the ability to constantly guide them throughout the entire learning path (Sen et al., 2022; Breckner, 2022). Obviously, the results obtained have exclusively exploratory purposes (given the low number of samples) but play a potentially significant role in directing future research in this field of study. The tool used in the experiment has been structured in such a way as to allow students to face specific educational challenges, also providing a deeply structured framework to articulate their learning process. This research holds both practical and theoretical significance. In this regard, it concretely introduces a new resource to enhance the learning experience, accelerating the assimilation of knowledge. Furthermore, it contributes to the theoretical understanding of the effectiveness of ITS in the academic educational context, offering valuable insights for the integration and enhancement of teaching through AL.

1.1 Theoretical context

The present study aims to explore the effectiveness of a Chatbot based on a Large Language Model (LLM) in providing adaptive learning experiences to university students. To contextualize the research, a literature review is presented on two main areas: adaptive learning and educational Chatbots.

Adaptive learning (AL) is a pedagogical approach that aims to personalize instruction based on students' characteristics and needs, optimizing learning outcomes and motivation. It relies on the use of information systems that can monitor students' progress and provide them with appropriate feedback and support (Xie, 2019; Ochukut, 2019; Zhao, 2020). Among the most commonly used information systems for adaptive learning are Intelligent Tutoring Systems (ITS), which can simulate the behavior of a human tutor and interact with students flexibly and dynamically. The literature has demonstrated the effectiveness of intelligent tutoring systems in various educational contexts, such as coding education (Troussas, 2021), online higher education (Jotsov, 2021), and exam preparation (Donnermann, 2022).

However, despite the advantages of adaptive learning, its implementation presents some challenges, such as technical complexity, teacher resistance, and a lack of standardization. To overcome these difficulties, some researchers have proposed the use of Chatbots as an alternative or complementary tool. Chatbots are conversational agents that can communicate with users in natural language, providing assistance and support in various domains. In the education sector, Chatbots can assume various roles, such as tutor, assistant, companion, or evaluator. The literature has highlighted the potential of educational Chatbots in enhancing learning and services offered to students (Pérez, 2020; Hwang, 2021). It has also investigated the factors influencing the adoption of Chatbots for educational purposes by university students, such as the perception of convenience and performance improvement (Malik, 2021). However, it has also emphasized the need for further exploration of their pedagogical role, evaluation of effectiveness, and addressing implementation challenges (Thomas, 2020).

A distinctive element of this study lies in the use of an LLM for the development of the Chatbot. Unlike previous approaches that often rely on less sophisticated models, the model used in this research allows for accurate semantic analysis and flexible interaction with students. This feature enables the Chatbot to effectively emulate the behavior of a human tutor, adapting information to the student's needs and providing relevant responses. Furthermore, the adopted approach makes learning genuinely adaptive, as it allows students to select the information they wish to delve into, making the process customizable and flexible.

From this review, it becomes evident that adaptive learning and educational Chatbots are two promising areas for improving education, but they require further exploration and evaluation.

2. Materials and Methods

To evaluate the effectiveness of interaction with a Chatbot in AL, a pilot study was conducted: the experimental group interacted with the Chatbot and the control group received support from a tutor.

Partecipants

28 university students enrolled in a Digital Skills course participated in the study, randomly divided into two groups of 14 each, balanced by gender, age and level of prior knowledge.

Experimental protocol

Phases:

- Pre-test: all participants attended a GDPR lecture, lasting 30 minutes. During the lecture, students could take notes and ask questions to the teacher. At the end of the lecture, participants received course materials from the instructor and had 90 minutes to study and review the material. During this period, participants in the experimental group could interact with Albert, asking him questions and receiving personalized responses. Participants in the control group could ask questions aloud to a human tutor, who was available behind the lectern to assist students.
- Test: after a break, all participants filled out a questionnaire to collect their demographic data. Additionally, they completed a multiple-choice quiz to assess their level of knowledge acquisition regarding GDPR and a quiz to evaluate their ability to apply the knowledge learned through case studies. The knowledge acquisition quiz consisted of 15 questions, while the application quiz included 5 case studies selected by the course instructor.
- Post-test: two days later, students took another multiple-choice quiz to assess their level of knowledge retention regarding GDPR.

2.3 Technical Features of the Chatbot:

The Chatbot used in the study was developed using a Generative Pre-trained Transformer (GPT) model, which is a machine learning model capable of generating coherent and relevant text based on a question or keyword. It underwent fine-tuning to specialize its tutoring abilities. Additionally, the Chatbot was trained on a set of documents related to GDPR, the European Data Protection Regulation, to provide accurate and personalized responses to students. The Chatbot was evaluated using a text quality metric based on coherence, relevance, and grammaticality. The Chatbot's name is Albert, and it introduced itself to students with a welcome message on the Telegram platform.

3. Results

For data analysis, R-Study software was used for inferential analysis, and Jamovi 2.3.17 was used for descriptive analysis. To compare the performance between the two groups, an independent samples t-test was conducted for each dependent variable. The level of statistical significance was set at 0.05.

The main dependent variables were as follows:

- Knowledge acquisition: measured through the score obtained in the Test.
- Knowledge retention: measured through the score obtained in the quiz after two days.
- Application of knowledge: measured through the score obtained in the application task.
- Quiz completion time.

The analyses performed explored the observable differences between the two groups of participants in terms of knowledge acquisition, retention, and application of GDPR knowledge.

Descriptives

	Sample	N	Missing	Mean	Median	SD	Minimum	Maximum
Knowledge acquisition	Control	14	0	10.57	11.00	2.409	5	14
	Experimental	14	0	10.86	11.50	3.416	5	15
Completion time - quiz acquisition	Control	14	0	6.09	6.32	1.550	3.07	9.13
	Experimental	14	0	4.63	4.32	1.873	2.22	8.83
Application of knowledge in case studies	Control	14	0	2.43	2.50	0.646	1	3
	Experimental	14	0	3.00	3.00	0.679	2	4
Reiteration of knowledge	Control	14	0	9.36	9.00	2.023	7	13
	Experimental	14	0	10.00	10.00	2.449	6	13
Completion time - reiteration quiz	Control	14	0	7.41	5.74	4.543	3.72	21.75
	Experimental	14	0	5.72	5.17	1.773	2.63	9.07

Tab. 1 Descriptive statistics

Before delving into a detailed analysis of the dependent variables directly related to the learning processes, it is necessary to report the results obtained for a variable that cuts across all others: quiz completion time. Table 1 displays the descriptive statistics of quiz completion times for both groups. The results indicated a significant difference between the two groups (t(26) = 2.241, p = 0.034). The analysis carried out by applying the t-test to independent samples shows a significant difference between the two groups (t(26) = 2.241, p = 0.034). This result, while contextualized within a pilot study, appears relevant for the purposes of the research, because it shows how the interaction between the Chatbot and the students has had a positive impact on the response times to the questions in the questionnaire. Therefore, being a transversal variable to the others, it has a certain weight because it influences consequently all the others. Instead, for the learning domains analysis, the knowledge acquired was measured by the score obtained in the test, which consisted of 15 multiple choice questions on the GDPR with a maximum possible score of 15 points. The t-test analysis showed no significant differences between the two groups (t(26) = -0.256, p = 0.800). Different is the case of the variable "Application of the knowledge", measured from the score obtained in the quiz to multiple choice on the practical cases. The quiz consisted of 5 real business case studies, with 4 answer options each. The maximum possible score was 5 points. In fact, the values associated with this domain results are significantly higher in the experimental group than in the control group (t(26) = -2.280, p = 0.031). This means that interaction with the Chatbot has had a positive effect on the application of GDPR knowledge in case studies. Finally, the last thoughtful variable, namely the repetition of knowledge, is measured by the score obtained in the quiz after two days and the completion time for the same quiz. The maximum possible score was 15 points. Both groups exhibited a slight decline compared to the test in the initial phase, both in terms of score and time but the experimental group showed a smaller decrease and this group exhibited greater variability in both scores and completion times compared to the control group. Additionally, it can be noted that the experimental group had some outliers for both scores and times, indicating that some participants showed a very high improvement or deterioration between the test and the quiz after two days. The results indicated no significant difference between the two groups for either score (t(26) = -0.757, p = 0.456) or completion times (t(26) = 1.295, p = 0.207). This means that interaction with the Chatbot did not have a significant effect on the reiteration of GDPR knowledge.

4. Discussion

The results of this pilot study provided empirical evidence to support the hypothesis that the use of a Chatbot based on a LLM can improve the learning effectiveness of college students compared to traditional teaching methods.

Specifically, the experimental group that interacted with the Chatbot named Albert completed tests faster and achieved higher average scores on all tests than the control group, which used a human tutor. These results support the hypothesis that a Chatbot-based ITS may be more effective in adapting the learning process to the individual needs of students.

Albert's ITS architecture, consisting of a learner model, a domain model and a teaching model, made it possible to constantly monitor student performance and provide personalized support throughout the learning process. The student model used text classification algorithms, such as Support Vector Machine (SVM) and Naive Bayes, to categorize student queries according to the type of support required (memorization, comprehension or neutral/out-of-context queries). The domain model used semantic analysis techniques, such as dependency analysis and semantic relationship recognition, to understand the meaning of student requests and provide relevant responses. The teaching model adapted teaching strategies according to students' individual needs, allowing them to independently select the topics they wanted to learn more about, making the process customizable and flexible. This approach made learning effectively adaptive.

The results of this study are in line with the existing literature on the effectiveness of ITS in increasing the effectiveness of student learning (Xie, 2019; Ochukut, 2019; Zhao, 2020). However, the present study stands out for using an LLM as the basis for the chatbot, enabling accurate semantic analysis and flexible interaction with students, effectively emulating the behavior of a human tutor.

Conclusions

Despite promising results, this study has some limitations. Albert is based on a LLM, which is an artificial intelligence model that can generate natural text from input. However, LLMs are not without issues. Some of these are:

- Lack of coherence and cohesion in the generated text. LLMs can produce text that is grammatically correct but semantically inconsistent or irrelevant to the context or goal of the interaction (Kasneci et al., 2023).
- The difficulty of handling user questions that require a precise or specific answer. LLMs may provide vague or general answers, or avoid the question altogether (Min et al., 2023).

- The presence of bias or prejudice in the generated text. LLMs can reproduce or amplify biases present in the data used to train them, such as gender, racial, or ideological bias (Abid et al., 2021).

These limitations may have negatively affected the learning experience of the students who interacted with the Chatbot, reducing their understanding, confidence, and interest in the GDPR content. In addition, the fine-tuning was performed by paying attention to the attitude the Chatbot should have instead of focusing on structuring the knowledge acquisition path.
Another limitation is related to the methodology of the study. The study involved a relatively small sample of college students (N = 28), divided into two groups (experimental and control) randomly assigned. The experimental group interacted with the chatbot for about 90 minutes, while the control group interacted with a human tutor. This methodology has some critical issues that may have limited the validity and generalizability of the results. Some of these are:

- Sample characteristics. The sample consisted of few participants from only one university and one faculty. This reduces the representativeness of the sample with respect to the population of college students and the possibility of extending the results to other contexts or domains.
- The short duration of the intervention. The Chatbot intervention lasted only 90 minutes, which may not be sufficient to appreciate the effects of adaptive learning. Previous studies have suggested that adaptive learning through ITS takes a longer time to manifest, as it depends on student progression and monitoring (Paladines, 2020).
- Lack of long-term follow-up. The study only measured students' level of learning up to two days after the Chatbot intervention. This may not be enough to assess the duration and transferability of the acquired knowledge. Previous studies have indicated that adaptive learning can also have positive effects in the long term, as it promotes students' metacognition and self-regulation (Alebeisat, 2022).

These critical issues may have adversely affected the sensitivity of the study in detecting differences between the two groups of participants, reducing the statistical power and significance of the results. Therefore, future prospects commit us to overcome the obstacles presented in order to structure ITSs capable of making learning more effective and satisfying.
In summary, this study contributes to the growing understanding of the role of chatbots in adaptive learning and provides insights for future developments in educational technologies. The use of an LLM-based chatbot opens new

opportunities for the design of innovative and personalized learning tools that can support students in achieving their educational goals.

References

Abid A., Farooqi M., & Zou J. (2021). *Persistent Anti-Muslim Bias in Large Language Models.* Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, 298–306. https://doi.org/10.1145/3461702.3462624

Alebeisat F., Altarawneh H., Alhalhouli Z. T., Qatawneh A., & Almahasne M. (2022). *The Impact of Human and Computer Interaction on eLearning Quality.* Int. J. Interact. Mob. Technol., 16, 58–69.

Breckner B., Săcărea C., & Zavaczki R. R. (2022). *Improving User's Experience in Exploring Knowledge Structures: A Gamifying Approach.* Mathematics, 10(5), Articolo 5. https://doi.org/10.3390/math10050709

Chen Y., Jensen S., Albert L. J., Gupta S., & Lee T. (2023). *Artificial Intelligence (AI) Student Assistants in the Classroom: Designing Chatbots to Support Student Success.* Information Systems Frontiers, 25(1), 161–182. https://doi.org/10.1007/s10796-022-10291-4

Donnermann M., Schaper P., & Lugrin B. (2022). *Social Robots in Applied Settings: A Long-Term Study on Adaptive Robotic Tutors in Higher Education.* Frontiers in Robotics and AI, 9. https://api.semanticscholar.org/CorpusID:247439829

Essel H. B., Vlachopoulos D., Tachie-Menson A., Johnson E. E., & Baah, P. K. (2022). *The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education.* International Journal of Educational Technology in Higher Education, 19(1), 57. https://doi.org/10.1186/s41239-022-00362-6

Huang W., Hew K. F., & Fryer L. K. (2022). *Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning.* Journal of Computer Assisted Learning, 38(1), 237–257. https://doi.org/10.1111/jcal.12610

Hwang G.J., & Chang, C.Y. (2021). *A review of opportunities and challenges of chatbots in education.* Interactive Learning Environments, 0(0), 1–14. https://doi.org/10.1080/10494820.2021.1952615

Jing Y., Zhao L., Zhu K., Wang H., Wang C., & Xia Q. (2023). *Research Landscape of Adaptive Learning in Education: A Bibliometric Study on Research Publications from 2000 to 2022.* Sustainability, 15(4), Articolo 4. https://doi.org/10.3390/su15043115

Jotsov V. S., Akramova A., Tkach, G. F., Kerimbayev N., Madyarova G., Beisov N., & Bolyskhanova M. (2021). *Development of a Virtual Conference Online Platform for Adaptive Learning.* 2021 International Conference Automatics and Informatics (ICAI), 106–110.

Kasneci E., Sessler K., Küchemann S., Bannert M., Dementieva D., Fischer F., Gasser U., Groh G., Günnemann S., Hüllermeier E., Krusche S., Kutyniok G., Michaeli T., Nerdel C., Pfeffer J., Poquet O., Sailer M., Schmidt A., Seidel T., … Kasneci G. (2023). *ChatGPT for good? On opportunities and challenges of large language models for education.* Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Malik R., Shrama A., Trivedi S., & Mishra R. (2021). *Adoption of Chatbots for Learning among University Students: Role of Perceived Convenience and Enhanced Performance.* Int. J. Emerg. Technol. Learn., 16. https://api.semanticscholar.org/CorpusID:240590075

Min B., Ross H., Sulem E., Veyseh A. P. B., Nguyen T. H., Sainz O., Agirre E., Heintz I., & Roth D. (2023). *Recent Advances in Natural Language Processing via Large Pre-Trained Language Models: A Survey.* ACM Computing SurveysAcceptato. https://doi.org/10.1145/3605943

Mousavinasab E., Zarifsanaiey N., R. Niakan Kalhori S., Rakhshan M., Keikha, L., & Ghazi Saeedi M. (2021). *Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods.* Interactive Learning Environments, 29(1), 142–163. https://doi.org/10.1080/10494820.2018.1558257

Ochukut S. A., & Oboko R. (2019). *A Learner Model for Adaptive e-Leaning Based on Learning Theories.* 2019 IST-Africa Week Conference (IST-Africa), 1–8. https://doi.org/10.23919/ISTAFRICA.2019.8764826

Paladines J., & Ramirez J. (2020). *A Systematic Literature Review of Intelligent Tutoring Systems With Dialogue in Natural Language.* IEEE Access, 8, 164246–164267. https://doi.org/10.1109/ACCESS.2020.3021383

Pérez J. Q., Daradoumis T., & Puig J. M. M. (2020). *Rediscovering the use of chatbots in education: A systematic literature review.* Computer Applications in Engineering Education, 28, 1549–1565.

Sandu N., & Gide E. (2019). *Adoption of AI-Chatbots to Enhance Student Learning Experience in Higher Education in India.* 2019 18th International Conference on

Information Technology Based Higher Education and Training (ITHET), 1–5. https://doi.org/10.1109/ITHET46829.2019.8937382

Sen W., Hong Z., & Xiaomei Z. (2022). *Effects of human-machine interaction on employee's learning: A contingent perspective.* Frontiers in Psychology, 13, 876933. https://doi.org/10.3389/fpsyg.2022.876933

Thomas H. (2020). *Critical Literature Review on Chatbots in Education.* https://api.semanticscholar.org/CorpusID:235739218

Troussas C., Krouska A., & Sgouropoulou C. (2021). *A Novel Teaching Strategy Through Adaptive Learning Activities for Computer Programming.* IEEE Transactions on Education, 64, 103–109.

Wu R., & Yu Z. (2023). *Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis.* British Journal of Educational Technology, n/a(n/a). https://doi.org/10.1111/bjet.13334

Xie H., Chu H.C., Hwang G.J., & Wang C.C. (2019). *Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017.* Computers & Education, 140, 103599. https://doi.org/10.1016/j.compedu.2019.103599

Yin J., Goh T.T., Yang B., & Xiaobin Y. (2021). *Conversation Technology With Micro-Learning: The Impact of Chatbot-Based Learning on Students' Learning Motivation and Performance.* Journal of Educational Computing Research, 59(1), 154–177. https://doi.org/10.1177/0735633120952067

Yu L. Q., Wilson R. C., & Nassar M. R. (2021). *Adaptive learning is structure learning in time.* Neuroscience & Biobehavioral Reviews, 128, 270–281. https://doi.org/10.1016/j.neubiorev.2021.06.024

Zhao Y., Fan Y., Yin M., & Fang C. (2020). *Research on the Construction of a Student Model of an Adaptive Learning System Based on Cognitive Diagnosis Theory.* International Journal of Digital Crime and Forensics (IJDCF), 12(4), 20–31. https://doi.org/10.4018/IJDCF.2020100102