TransInformação

THEMATIC SECTION

Digital information, data management and governance, and research information systems: an educational approach

Editor
Carlos Luis Gonzáles-Valiente

Support
Henan Province Teacher Education Curriculum Reform Research Project "Reform and Implementation Strategies of Education and Training Courses for Normal University Students Focusing on Core Literacy" (number: 2023-JSJYZD-017).

Conflict of interest
The author declare have no conflict of interests.

Received
Jan 29, 2024

Final versión
Apr 10, 2024

Approved
Apr 11, 2024

Research on information construction of knowledge graph based on literature retrieval in english learning

Pesquisa sobre construção de informação do gráfico do conhecimento baseado na recuperação de literatura na aprendizagem de inglês

Guolong Zhao[1]

[1] Xinyang Normal University, College of Teacher Education. Henan, China. E-mail: <zgl5127@xynu.edu.cn>.

How to cite this article/*Como citar este artigo*: Zhao, G. Research on information construction of knowledge graph based on literature retrieval in english learning. *Transinformação*, v. 36, e2410862, 2024. https://doi.org/10.1590/2318-0889202436e2410862

Abstract

This study aimed to explore the construction of an English language knowledge graph based on literature retrieval to support intelligent education. A questionnaire was administered to collect data on students' experiences with traditional and technology-enhanced learning approaches. Literature was also retrieved and analyzed to populate the knowledge graph domains. The results showed that implementing a knowledge graph significantly improved learning personalization and fostered greater student engagement compared to conventional teaching methods. Real-time analytics and continuous feedback further optimized the learning process. Post-implementation assessments found notable gains in students' academic performance and inclination toward English learning. The personalized, adaptive learning environment facilitated by the knowledge graph more effectively sustained interest and promoted achievement. In conclusion, knowledge graphs constructed through literature analysis hold promising potential for advancing English education when incorporated into intelligent tutoring systems. By mapping interconnections within the subject domain visually and computationally, they can power highly customized instruction tailored to individual needs.

Keywords: Controlled vocabulary search. Document visibility. Information technologies.

Resumo

O objetivo deste estudo foi explorar a construção de um gráfico de conhecimento da língua inglesa baseado na recuperação de literatura para apoiar a educação inteligente. Um questionário foi administrado para coletar dados sobre as experiências dos alunos com abordagens de aprendizagem tradicionais e aprimoradas pela tecnologia. A literatura também foi recuperada e analisada para preencher os domínios do gráfico de conhecimento. Os resultados mostraram que a implementação de um gráfico de conhecimento melhorou significativamente a personalização da aprendizagem e promoveu um maior envolvimento dos alunos em comparação com métodos de ensino convencionais. A análise em tempo real e o feedback contínuo otimizaram ainda mais o processo de aprendizagem. As avaliações pós-implementação encontraram ganhos notáveis no *desempenho acadêmico dos alunos e na inclinação para o aprendizado do inglês. O ambiente de aprendizagem personalizado e adaptativo facilitado pelo gráfico de conhecimento sustentou de forma mais eficaz o interesse e promoveu o desempenho. Em conclusão, os gráficos de conhecimento construídos através da análise da literatura têm um potencial promissor para o avanço da educação de inglês quando incorporados em sistemas de tutoria inteligentes. Ao mapear as interconexões dentro do domínio da disciplina visual e computacionalmente, eles podem fornecer instruções altamente personalizadas, adaptadas às necessidades individuais.*

Palavras-chave: *Pesquisa em linguagem controlada. Visibilidade dos documento. Tecnologias da informação.*

Introduction

The advent of information technology and the ongoing advancements in artificial intelligence have led to the education sector's progressive digitalization and intelligent transformation. In light of the current era, educators, researchers, and educational institutions have initiated a proactive exploration of leveraging advanced technological resources to enhance the standard of instruction, optimize learning efficiency, and effectively address the evolving demands of contemporary society in talent development (Xin, 2021). Teaching English garners significant attention as a vital instrument for facilitating worldwide communication. Enhancing the effectiveness and appeal of English language instruction for pupils has been a persistent concern among educators. Integrating modern information technology and artificial intelligence technology has become a viable avenue for innovation in the digital era (Menekse, 2023). The integration of this fusion presents novel opportunities and potential for the field of English language instruction, hence stimulating comprehensive investigations and scholarly inquiries into inventive pedagogical methodologies. This research will utilize English learning based on literature retrieval as a focal point, integrating information technology and artificial intelligence technology to investigate the advancement of English education within the digitalization and intelligence framework.

This study aims to develop intelligent English Learning based on a literature retrieval system that utilizes sophisticated technology to enhance teaching effectiveness, optimize learning, and foster students' language proficiency and overall competence (Ruiz; Rebuschat; Meurers, 2021). Using extensive research and practical application, a state-of-the-art, highly appealing, and efficient instructional approach will be developed for English education. This objective aims to create a knowledge graph based on literature retrieval utilizing machine learning and deep learning methodologies to enhance the provision of intelligent and individualized assistance for English language instruction. Using a Knowledge Graph based on literature retrieval can improve the integration and organization of knowledge by functioning as a semantic network, which, in turn, can offer educators and learners more streamlined access to learning materials and personalized learning trajectories. This study aims to investigate the construction of English Learning based on a literature retrieval knowledge graph based on machine learning and deep learning approaches and the subsequent implementation of this knowledge graph based on literature retrieval in English education (Chiu; Tseng, 2021). The evaluation of the effects of this intelligent teaching approach on students' academic performance and learning interest will involve an examination of students' academic achievements, levels of engagement, and instructional techniques. This objective aims to generate novel insights and methodologies for enhancing the teaching of English. This study aims to develop intelligent English Learning based on a literature retrieval system by leveraging advanced technology teaching practices and integrating machine learning and profound learning principles. The objective is to enhance traditional English Learning based on literature retrieval mode and facilitate the sustainable development of English education.

Through questionnaires and data analysis of students adopting machine learning and deep learning, this study aims to compare the differences between these two modes of instruction regarding students' interest in learning and academic performance. An in-depth understanding of the substantial impact of these emerging technologies in English language teaching and how they shape the learning process and academic performance is pursued (Cui, 2023). The benefits of machine learning and deep learning in English language teaching can be clearly and unambiguously demonstrated through constructing a knowledge map. This visual presentation contributes a more comprehensive understanding of how these advanced technologies affect students' learning process, interest, and performance (Shiranirad; Burnham; English, 2021). A questionnaire survey was conducted covering students using machine learning, deep learning, and traditional teaching modes to compare the impact of different teaching modes on English Learning based on literature retrieval interest and academic achievement. The questionnaire survey results were supplemented by interviews with students adopting machine learning, deep learning, and traditional teaching modes (Liu; Hwang; Liu, 2021). English teachers should embrace more machine learning and deep learning to teach students, and applying machine learning and deep learning to English Learning based on literature retrieval can effectively improve students' English proficiency.

Theoretical Research

Machine learning

Machine learning is an artificial intelligence technique that rearranges the structure of knowledge to improve the performance of a system by analyzing data and using the results. Through algorithms and models, systems can learn patterns of laws from large amounts of data and use this knowledge to improve their performance. This learning ability allows the system to gradually optimize its behavior to respond more intelligently to different situations. Scholars point out that among the many areas of artificial intelligence, machine learning is considered crucial because it enables computers to not only perform pre-programmed tasks but also to learn, understand, and adapt from data. Systems that cannot learn tend only to be able to perform specific predefined tasks, making it difficult to adjust to new scenarios and requirements and, therefore, demanding to be called truly intelligent. Of particular interest is the far-reaching impact of machine learning on education. Through intelligent education systems, machine learning can be personalized to adapt to student's learning styles and needs, providing customized learning content and advice. This personalized education helps to improve learning efficiency and stimulate students' interest in learning, which promotes the innovation and development of education. Machine learning can also help educators analyze student learning, develop more scientific teaching strategies, and achieve more efficient education management, thus promoting the upgrading and improvement of the entire education system. The core tasks of machine learning include data analysis and data mining, which allow in-depth study of various fields, such as education. For example, machine learning can extract meaningful information from massive amounts of data by analyzing students' subject learning over a year, which includes analysis of learning behaviors, subject preferences, and learning progress, providing students and educators with a deeper understanding. The results of utilizing machine learning can provide precise insights into the root causes of student learning problems, not just the surface. For example, it is possible to identify bottlenecks in subject learning, clarify learning difficulties, or reveal correlations between subjects. This insight lays the groundwork for developing personalized learning recommendations, enabling educators to target teaching strategies to stimulate students' interest and motivation.

Machine learning relies on historical instances and running records to quickly generate the results of the current execution through pattern recognition and profound data analysis. During this process, the system automatically adjusts algorithms and models based on previously accumulated experience to flexibly adapt to the task requirements in different contexts and achieve more accurate and efficient execution and decision-making. This ability to continuously learn and adapt allows machine learning to maintain accuracy and efficiency in an ever-changing environment, bringing great value and innovation to various fields. In this system, the environment is the starting point for education, and the source of information and stimuli constantly received during the learning process. The climate impacts the learning function, which shapes the characteristics and difficulty of the learning task. Machine learning, as an essential part of the learning system, corrects, updates, and optimizes the existing knowledge base by analyzing and processing the data and information provided by the environment. This updated knowledge base is the basis of the execution phase, guiding the machine to the following decisions and operations. The execution phase is the practical application phase of the learning system, which relies on the updated knowledge base. The machine performs specific tasks based on the guidance information in this knowledge base, constantly interacting with the environment and generating new experiences. After each execution, the system incorporates the expertise from this run into the knowledge base, creating a learning cycle. This cycle-guided learning process gradually accumulates rich experience, continuously strengthens the knowledge base, and improves the system's self-learning ability and judgment ability. The learning system can adapt to different environments and tasks through this feedback mechanism, optimize itself continuously, and achieve more efficient and intelligent operation and decision-making.

To explain the nature of machine learning more clearly, scholars resort to an example. Suppose a machine needs to determine whether a particular mathematical operation's result is correct based on the user's evaluation. In the process of machine learning, the device makes use of loop structures and judgment statements in computer language programs for analysis. This learning process must be done under proper guidance to ensure accurate results and avoid misleading the learner. Suppose machine learning is performed under artificial misguidance. In that case, the system may give incorrect answers, such as incorrectly determining that a mathematical operation equals a specific value, highlighting that the machine-learning process takes place through the influence of the environment. The machine receives user evaluations and correct answers in this environment, and by analyzing and learning from this information, it gradually adjusts its judgmental abilities to provide accurate results for mathematical operations. Machine learning is achieving correct judgment and efficient decision-making based on interaction and feedback with the environment and constantly adjusting and optimizing its models and algorithms. The information and feedback in the background are the basis of machine learning, which is crucial for building intelligent systems and improving their performance.

Deep learning models

Artificial neural networks are the model underlying deep learning, consisting of many simple neurons interconnected to form a complex network structure. This model is designed to simulate the operation of the biological nervous system to realize the simulation and judgment of natural things in the computer. Through signal transmission and processing between neurons, the artificial neural network can gradually adjust the connection weights and optimize itself to adapt to different patterns and data for more accurate judgment and learning. This method of simulating biological neural systems provides a powerful tool for deep learning, is widely used in image recognition, speech processing, natural language understanding, etc., and promotes the rapid development of artificial intelligence. Each neuron can receive multiple input signals, process them through weighted sums and activation functions, and then produce an output signal, which can be used as an input for other neurons. This information transfer and processing method allows artificial neural networks to learn and adapt. By repeatedly adjusting the connection weights between neurons, the network can gradually optimize its model, thus improving the accuracy of its judgment of complex things r. Biological nervous systems inspire the design and operation of artificial neural networks but are also designed to be reasonably simplified and abstracted. This model has already achieved remarkable results in various fields, such as image recognition, natural language processing, and medical diagnosis. It is one of the essential foundations for breakthroughs in deep learning technology and provides a powerful tool for computers to simulate human intelligence.

Artificial neural networks include input, hidden, and output layers, with inter-layer connection weights to transfer information. Neurons in the same layer are not directly connected. Gradient vanishing and gradient explosion can occur in deep networks, and parameters must be adjusted to avoid these problems. Recurrent neural networks are an upgraded version of artificial neural networks. Unlike traditional neural networks, there are recurrent connections between neurons in recurrent neural networks, allowing the output of the current neuron to be influenced by previous works, capturing sequential information about the input data, and making model predictions more accurate. However, deep recurrent neural networks also face the problem of gradient vanishing and explosion, which needs to be solved by adjusting the time step. Long-short-term memory is a commonly used variant of recurrent neural networks, which calculates connection weights between neurons through the leaky unit and threshold mechanism, establishes long-term connections, and combines short-term memory to produce accurate outputs.

Both long-term and short-term memory systems comprise input, forgetting, and output gates. These gates are crucial in preventing gradient disappearing and gradient explosion issues using weight adjustment. A Convolutional Neural Network (CNN) is a feed-forward neural network that employs convolutional propagation. Its purpose is to address the challenges of conventional artificial neural networks, including several parameters, limited efficiency, training complexities, gradient issues, and susceptibility to overfitting. Convolutional Neural Networks enable the preservation of planar structural information inherent in the input data using convolution kernels. This facilitates the more efficient processing of visual and spatial data through the three-dimensional organization of neurons. The architecture comprises several key components: an input layer, a convolutional layer, a pooling layer, a fully connected layer, and a Softmax layer. The convolutional layer is the central component that carries out convolution operations and acquires knowledge from the input data using filters. This process effectively reduces the number of parameters and mitigates the occurrence of overfitting. The pooling layer is a technique used in neural networks to decrease the size of the input matrix and the number of parameters, achieved by computing the maximum or average value within a pooling region. Two pooling methods that are often employed are complete pooling and average pooling. These techniques are beneficial as they contribute to reducing the complexity of the model and enhancing computational efficiency. The fully linked layer is responsible for the categorization and processing of information. On the other hand, the Softmax layer provides a probability distribution that represents the final forecast. The architecture and methodology of Convolutional Neural Networks have yielded notable achievements in image recognition, computer vision, and natural language processing, making them extensively employed in practical applications.

The Transformer model is a deep learning model based on a self-attention mechanism that allows the model to focus on different parts of the input sequence at various locations, thus capturing long-distance dependencies within the series. Through the self-attention agency, the model can learn and understand the associations between inputs more efficiently, a property that significantly improves the training speed and performance of the model. Regarding parallel computation, the Transformer model has an advantage because it can process multiple positions of the input sequence simultaneously without relying on sequential processing like the RNN model. This parallel computing capability makes Transformer more efficient and faster at processing large-scale data without degrading the accuracy and performance of the model. These advantages have made the Transformer an essential model in natural language processing, image processing, etc., and have led to excellent results in many tasks. The Transformer model consists of four parts: input, encoder, decoder, and output. The input is sequence data with additional positional coding to provide positional information. The encoder consists of multiple coding modules that convert the sequence data into a matrix and generate a new matrix Z by Multi-Head Attention. Residual block X prevents network degradation, and normalization improves training speed and stability. Then, it is trained by a two-layer fully connected neural network and enters the decoder. Decoder inputs are available for training and prediction, containing Mask codes to mask specific values and Padding Mask to implement sequence padding to ensure attention is focused on important information. The decoder consists of multi-layer decoding units that are linearly transformed, and Softmax is used to obtain the probability distribution and final prediction.

Intelligent tutoring

Scholars believe immediate feedback is an essential part of the learning process. Traditional manual feedback, where the teacher assumes the role, cannot provide feedback anytime, anywhere. However, anytime, anywhere learning support is possible with the help of machines, which adds great flexibility and convenience to the learning process. Smart devices can instantly guide the learning process while teaching knowledge. As learners answer questions, smart devices can provide immediate guidance until the learner gets the correct answer. Compared with giving answers directly, this teaching method can better motivate learners to participate in thinking, understand knowledge, and accumulate experience in solving problems. Therefore, using personalized instant feedback provided by smart devices can effectively improve the effectiveness and efficiency of learning. Scholars developed the landmark first truly adaptive learning machine, SAKI, in the 1980s. The SAKI was innovative because it was designed to train keyboard operators to teach how to process data on punching equipment. Unlike previous machines, SAKI can learn adaptively, adjusting its learning model in real time based on the immediate performance of the learner. This breakthrough enables SAKI to build dynamic probabilistic learning models that optimize learning strategies and provide real-time personalized learning advice on the learner's performance at every step.

The introduction of SAKI lays the foundation for personalized education and paves the way for developing subsequent intelligent education systems. This adaptive learning model provides learners with a more effective learning experience and pioneered adaptive education technology. In the 1980s, two main ideas emerged in the development of artificial intelligence: adaptive learning machines, such as SAKI, and Computer-Assisted Instruction (CAI) systems, of which the PLATO teaching system had a notable impact. The PLATO allowed students to access standardized teaching materials through remote terminals, which realized human-computer interaction for the first time and changed the information transfer mode between learners and computers. The PLATO has changed the way information is transferred between learners and computers. With the rise of personal computers, CAI was widely used, but limited by the level of artificial intelligence technology at that time, CAI could not fully understand the characteristics of unique learners and could not realize brilliant tutoring. Computer-assisted instruction has discovered accurate intelligence with the significant improvement of computer technology.

Intelligent tutoring originated from CAI, a typical application of artificial intelligence in education, relying on Intelligent Tutoring Systems (ITS). At the end of the twentieth century, Intelligent Tutoring Tools were planned to be developed to generate Intelligent Tutoring Assistants (ITA) that targeted different subject areas. Teachers could create ITA and knowledge databases on demand, which learners accessed via the Web. However, the system did not enable personalized and precise tutoring. In the 21st century, ITS aims to design Intelligent Pedagogical Agents that can model, predict, and monitor learners' learning behaviors, abilities, and mental states and provide personalized help and feedback. The ITS achieves this goal by leveraging various technologies, moving away from search-based tutoring tools and toward assistive devices that include human-computer interactions. The ITS attempts to replicate the role of the teacher by automating the question-generating, selecting, and giving feedback. The ITS tries to replicate the part of the teacher and automate teaching functions such as question generation, selection, and feedback. However, AI cannot completely replace the teacher, so it is shifting to a blended learning model, realizing a complementary teaching model in which the teacher is the primary and secondary teacher.

Specifically, the research progress in the above literature has laid a necessary foundation for the established use of knowledge graphs in intelligent tutoring systems. They outline the best practices for visualizing and computing structured domain-specific knowledge to enhance the adaptive learning experience. Due to current research attempting to establish an English knowledge graph, these references provide appropriate information for research design. In addition, relevant literature extensively explores knowledge mapping techniques applied to second language acquisition. Discussed how semantic networks effectively represent the interrelated concepts and patterns learned over time. When this study examines the English learning environment, the above literature provides an important background for the study.

Methodology

Research questions and questionnaire analysis

Aiming to apply the constructed knowledge graph based on literature search to English learning through literature search practice, personalized teaching, intelligent assisted learning, and other aspects (Zhang *et al.*, 2021). This dissertation aims to enhance college students' recognition of English Learning based on literature retrieval by applying the Intelligent English Learning based on literature retrieval System to the college English classroom (Yang *et al.*, 2021). This innovative teaching method aims to explore the impact of AI technology in English education on students' academic performance and subject interest, optimize the effectiveness of university English Learning based on literature retrieval, stimulate students' interest in English Learning based on literature retrieval, and improve academic performance (Ye *et al.*, 2023). By constructing a knowledge map of English Learning based on literature retrieval from machine learning and deep learning perspectives, this paper aims to provide more comprehensive guidance for English Learning based on literature retrieval. The following key questions must be addressed: 1) (Charifi *et al.*, 2021; Eng; Wolock; Wieczorek, 2023). Can machine learning and deep learning stimulate students' interest in English learning based on literature retrieval? Can intelligent systems tailor learning content and teaching methods to students' preferences and characteristics to make learning more attractive and challenging? 2) (Aouichaoui *et al.*, 2022). Can machine and deep learning increase students' interest in learning? Can personalized learning paths and real-time feedback mechanisms enable students to learn more efficiently?

Using a questionnaire as the central methodology, this study aims to gain insights into students' perceptions of the application of machine learning and deep learning in English language teaching, as well as the potential impact of these new technologies on students' interest in English language learning and their interest in learning (Moore; Schneider; Strickland, 2021). The questionnaire is well-designed and divided into five parts, including the title, preface, guidelines, questions and optional answers, and conclusion, ensuring the questionnaire is logical and easy to read (Jiang *et al.*, 2022). To ensure the quality of the questionnaire, a pilot evaluation was first conducted to test the feasibility and accuracy of the questionnaire so that it could be adjusted and optimized as necessary (Geisler *et al.*, 2021). The questionnaire design focused on clarity of questions and avoided vague, leading, sensitive, and complex queries to ensure students could understand clearly and respond accurately (Huang *et al.*, 2022). For the experimental design, two parallel university classes were carefully selected to provide the proximity of students' English proficiency and to maintain consistency in content and credit hours (Roberts, 2023). The experimental class utilized an advanced machine learning and deep learning teaching model (Xu *et al.*, 2024). In contrast, the control class followed traditional teaching methods to compare the effectiveness of the two teaching models (Matthews; Proctor, 2021). Through this design, the author aims to comprehensively assess the impact of machine learning and deep learning in English Learning based on literature retrieval and learning and provide a reliable empirical basis for evaluating the effects.

Before commencing the experiment, an initial round of questionnaire surveys was administered to comprehensively understand students' English Learning based on literature retrieval, encompassing their educational interests, learning preferences, and other relevant aspects. This preliminary phase served as a basis for subsequent experimental procedures and offered a vital point of reference for the design and implementation of the experiment. Following the inquiry, the next set of questionnaires was administered after the study to ascertain any potential alterations in the students' inclination towards English Learning based on literature retrieval and their overall interest in learning after their exposure to machine learning and deep learning techniques in the context of English instruction. The researchers ensured the maintenance of questionnaire design consistency while also making careful adjustments to the order of the questions to enhance the objectivity and accuracy of the collected data. The questionnaire designs utilized in this study were designed by the research topic and informed by the findings of other studies. The questionnaires underwent reliability testing to ascertain their legitimacy and validity. By examining students' sentiments and assessments regarding the implementation of machine learning and deep learning in English instruction, a more profound comprehension of the significant influence of these technologies on English Learning based on literature retrieval can be attained. This, in turn, will offer robust backing for subsequent enhancements and refinements in teaching methodologies, ultimately enhancing the efficacy of English instruction and enriching students' learning encounters.

The research data showed that the questionnaire of this experiment demonstrated high reliability and internal consistency on the whole, with a Cronbach's alpha value of 0.933, which indicates that the questionnaire design has good reliability. The 15-question questionnaire design covered students and their interest in learning. Each question provided three options for students to choose from, namely "not at all," "mostly," and "very much". The scale design allowed for a quantitative assessment of students' perceptions and attitudes in these areas. At the same time, to reflect the students' overall level more intuitively, the sum of the scores of each of the five questions was used as the level score of an item, with higher scores representing better performance in that area. Such a design helps to gain insights into students' perceptions and feelings about machine learning and deep learning in English language teaching, providing strong support for the study's accuracy.

Data collection and processing

Before the questionnaire survey was officially conducted, the research team pilot-tested the questionnaire. After all questionnaires were collected, the questionnaire samples were tested for reliability and validity using IBM®SPSS® 26.0 statistical software. Based on the test results, inappropriate questions were eliminated and reordered to ensure that the final questionnaire could be formally used in the survey. The entire formal survey study was conducted at a university, and the research team distributed the questionnaire to the subjects. Before testing, subjects were informed that the questionnaire was designed for academic research purposes and would not be used to assess other aspects of performance. They were encouraged to answer the questions as accurately and honestly as possible. If there was any doubt about a question or needed an explanation, they could seek the experimenter's help to ensure that the most valid and reliable statistics were obtained – this questionnaire aimed to investigate innovative applications of machine learning and deep learning in English language teaching. One hundred questionnaires were distributed to the experimental and control classes, of which 100 valid questionnaires were received. This design ensured the sample's adequacy and the credibility of the findings.

After collecting the questionnaires, the obtained data were entered into the computer and analyzed in depth using the appropriate statistical analysis software r (Aouichaoui *et al.*, 2022). Firstly, the English Learning based on the literature retrieval situation of the students in the experimental and control classes was analyzed to explore the general status of English Learning based on the literature retrieval level of both, including academic performance, interests, and attitudes. Second, critical factors related to interest and interest in learning are identified by analyzing the main factors affecting English proficiency, such as study time and extracurricular activities. In the third step, the actual extent of the impact of machine learning and deep learning in English language teaching was assessed to determine whether this new teaching model has had a significant effect on academic performance and interest in learning by comparing and analyzing the English language learning levels of students in the experimental class with those in the control class. Subsequently, an interview outline was designed based on the results of the questionnaire. Five students from the machine learning and deep learning teaching students were randomly selected to participate in the interviews. The interviews included questions about English Learning based on literature retrieval interest after the experiment, changes in learning interest, the impact of machine learning and deep learning on English proficiency, and attitudes and expectations about future teaching methods. These questions aimed to gain insights into the students' feelings about the new teaching mode and their views on future teaching methods. The researcher usually negotiated with the interviewees about the time, place, and possible adjustments before the formal interviews to ensure that the discussions were convenient and smooth for both parties, reflecting the interview's seriousness and discretion. With the interviewees' consent, the researcher chose an appropriate recording device, usually a cell phone, to record the entire interview for subsequent analysis and organization.

Empirical Process

Impact of machine learning and deep learning on students' interest in learning

Before experimenting, the students' English Learning based on literature retrieval interest in both classes was equal, satisfying the experimental design's requirements. However, upon conducting data analysis after the experiment, it was determined that the *p*-value associated with the practical class was below the threshold of 0.05. This finding suggests a noteworthy disparity exists between English Learning based on literature retrieval interest observed in the experimental and control classes after implementing machine learning and deep learning techniques. The practical course exhibits a notably higher level of learning interest than the control class. This may be attributed to the introduction of machine learning and deep learning techniques, which bring about novel learning approaches and methodologies. Subsequently, upon retesting following the experiment, it was observed that the *p*-value of the experimental group remains below the threshold of 0.05. This finding indicates a statistically significant increase in students' learning interest in the practical course, thus providing further evidence to support the beneficial influence of machine learning and deep learning on students' learning interests. The instructional approach of machine learning and deep learning, facilitated by sophisticated algorithms and tailored learning trajectories, can more effectively address students' diverse learning requirements and deliver a more precise and efficient educational encounter. Implementing a personalized learning method facilitates enhanced concentration and a more profound comprehension of the subject matter among students. Consequently, this technique contributes to superior academic achievement in examinations and practical applications (Table 1).

Table 1 – Questionnaire on students' interest in learning English.

Classes or grades in school	Pre-experimental learning interest	Post-experimental learning interest	*t*-value	*p*
Control subjects	5	5	0	∞
Experimental class	5	9.66	-10.51	0
t-value	0	-10.51		
p	∞	0		

Note: ∞: represents an infinite *p*-value (significance level)
Source: Own elaboration (2024).

Applying machine learning and deep learning stimulated the students' interest in the experimental class, which may stem from the novel learning approach and more participatory course design, making learning less tedious and more engaging and interactive. This positive learning experience positively impacted students' academic achievement and motivation. At the same time, by constructing a knowledge graph based on literature retrieval, it is clear that there is a strong connection between students' interest in learning English and machine learning and deep learning. As seen in Figure 1, in traditional English Learning based on literature retrieval, students often need help with single-course content and relatively fixed teaching methods, which can easily lead to a decline in learning interest. However, innovative teaching methods introducing machine learning and deep learning provide students with a more diverse and personalized learning experience. As shown in Figure 2, machine learning and deep learning can personalize the course content and teaching methods according to students' learning habits and interest characteristics. This personalized matching makes students feel more engaged and enthusiastic about learning. The customized learning experience improves learning efficiency and stimulates students' interest in learning English. Students can acquire better knowledge and enhance their academic achievement through these novel learning approaches, increasing their confidence and enthusiasm for learning English. The knowledge graph based on literature retrieval demonstrates a positive association between machine learning and deep learning and students' interest in learning English. This association enables educators to better design instructional strategies to meet students' learning needs and provides valuable insights for improving students' interest in learning, as shown in Figures 1 and 2.

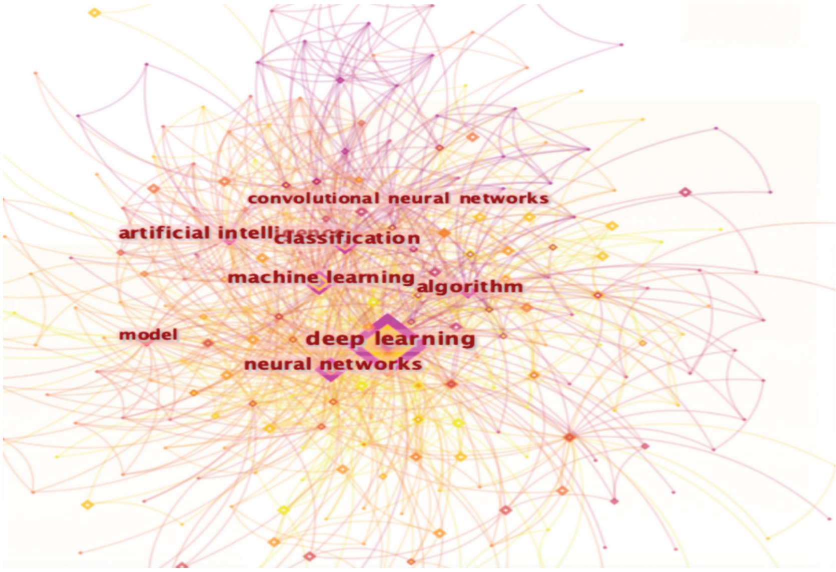

Figure 1 - Knowledge graph based on literature retrieval between machine learning and deep learning.
Source: Own elaboration (2024).

Figure 2 – Knowledge graph based on literature retrieval between machine learning vs. deep learning.
Source: Own elaboration (2024).

Impact of Machine Learning and Deep Learning on Student Academic Achievement

The significant effect of machine learning and deep learning in practical teaching can be observed by comparing the grades of the experimental class with those of the control class. For example, the average midterm rise of the control class is 36.6. The final step is 41.4, with a difference of 4.79 between the midterm and final average scores, while the middle midterm grade of the experimental class is 40.05. The intermediate final phase is 46.91, with a difference of 6.86 between the midterm and final average scores. These data show that the experimental class consistently outperforms the control class. The gap between the practical class's midterm and final average scores widened, and the final scores were significantly improved compared to the control class. For example, in the midterm test, the gap between the mean scores of the experimental and control types widened from 11.95 to 12.86, i.e., the practical class performed slightly lower than the control class. However, it is encouraging to note that in the following July post-test scores, the gap between the experimental and control classes narrowed from 12.86 to 9.59. This indicates that the students in the practical class improved their scores in the follow-up study, and the gap between the two types gradually narrowed, which is a remarkable demonstration of the effectiveness of machine learning and deep learning (Table 2). These comparative results highlight the positive role of machine learning and deep learning in English Learning based on literature retrieval, whose personalized and targeted learning approach helps improve students' performance and widen the gap in academic performance, providing strong support for technological innovation and improvement in education.

Table 2 – Academic Achievement of Students.

	Average score difference in the previous period	Medium-term average score difference	Average score difference at the end of the period
Experimental and control classes	11.95	12.86	9.59

Source: Own elaboration (2024).

Meanwhile, by analyzing Figure 3, it is evident that machine learning and deep learning significantly help English Learning based on literature retrieval. These phenomena are closely related to machine learning and deep learning, forming an interconnected and complex network. First, machine learning and deep learning technologies can help teachers better understand students' learning characteristics, preferences, and academic needs through extensive data analysis and pattern recognition. Second, these technologies can design personalized learning paths and content for English language teaching, tailoring the most appropriate teaching program to each student's learning level and interests. In addition, machine learning and deep learning can provide real-time feedback so students can keep abreast of their academic performance, identify deficiencies, and make improvements, as shown in Figure 3.

The findings of this study also have important implications for educational innovation and the practical application of technologies in teaching. Specifically, the results provide clear guidance on translating these research insights into actionable instruction strategies.

The study demonstrates the meaningful impacts of machine learning and deep learning approaches, and highlights their tremendous potential to enhance learning when integrated thoughtfully into pedagogy. Educators can leverage these techniques to develop personalized, engaging lessons that maximize student outcomes.

Figure 3 – English Learning based on literature retrieval and Artificial Intelligence.
Source: Own elaboration (2024).

For example, intelligent tutoring systems could be designed according to the principles of individualization, real-time support, and progressive learning outlined.

Moreover, the findings indicate a clear need to move beyond conventional classroom models that fail to sustain interest or achievement at comparable levels. As such, the results present a compelling case for educators and administrators.

Conclusion

First, implementing intelligent education systems using machine and deep learning significantly improved students' motivation to learn. By tailoring teaching materials and methods to individual needs and characteristics, these technologies made lessons more engaging and relevant for each learner. This personalized approach boosted learners' interest in their studies.

Second, real-time assessment and feedback allowed teachers to closely monitor learning progress and address weaknesses. Students received timely guidance to overcome obstacles, keeping them focused and driven to succeed. This continuous support from teachers optimizes their instruction and maximizes learning outcomes.

Finally, comparing pre- and post-implementation data revealed growth in students' academic performance after adopting machine learning-driven education. In contrast, the conventional teaching model appeared unable to sustain learners' involvement or promote the same levels of achievement.

Moving forward, research should focus on quantifying exact gains in test scores and comparing different applications of these advanced technologies. Understanding which capabilities most contribute to interest and performance will help refine intelligent systems for maximal educational impact. As the use of machine learning in classrooms increases, continued evaluation is essential to realize its full potential benefits.

References

Aouichaoui, A.R.N. *et al.* G.Uncertainty estimation in deep learning-based property models: Graph-based on literature retrieval neural networks applied to the critical properties. *AIChE Journal*, v. 68, n. 6, e17696, 2022. Doi : https://doi.org/10.1002/aic.17696

Charifi, R. *et al.* A. Sedimentary phosphate classification based on spectral analysis and machine learning. *Computers & Geosciences*, v. 150, 2021.

Chiu, C.K.; Tseng, J.C.R. A Bayesian Classification Network-Based Learning Status Management System in an Intelligent Classroom. *Educational Technology & Society*, v. 24, n. 24, p. 67-78, 2021.

Cui, Q. Multimedia Teaching for Applied Linguistic Smart Education System. *International Journal of Human-Computer Interaction*, v. 56, n. 56, p. 89-101, 2023. Doi: https://doi.org/10.1080/10447318.2022.2122111

Eng, K.; Wolock, D.M.; Wieczorek, M. Predicting baseflow recession characteristics at ungauged stream locations using a physical and machine learning approach. *Advances in Water Resources*, v. 11, n. 15, p. 45-58, 2023. Doi: https://doi.org/10.1016/j.advwatres.2023.104440

Geisler, E.L. *et al.* A Role for Artificial Intelligence in the Classification of Craniofacial Anomalies. *Journal of Craniofacial Surgery*, v. 32, n. 3, p. 967-969, 2021. Doi: https://doi.org/10.1097/SCS.0000000000007369

Huang, W. *et al.* A generic intelligent routing method using deep reinforcement learning with graphs based on literature retrieval neural networks. *IET Communications*, v. 69, p. 69-77, 2022. Doi https://doi.org/10.1049/cmu2.12487

Jiang, D. *et al.* Multiview feature augmented neural network for knowledge graph based on literature retrieval embedding. *Knowledge-Based Systems*, v. 56, n. 56, p. 65-77, 2022. Doi: https://doi.org/10.1016/j.knosys.2022.109721

Liu, Y. F.; Hwang, W.Y.; Liu, Z.Y. Effects of Mobile Drama With Authentic Contexts on English Learning based on Literature Retrieval: *Journal of Educational Computing Research*, v. 59, n. 7, p. 1294-1318, 2021. Doi: https://doi.org/10.1177/0735633121994289

Matthews, S.; Proctor, M.D. Can Public Health Workforce Competency and Capacity be Built through an Agent-based Online, Personalized Intelligent Tutoring System? *Educational Technology & Society*, v. 24, n. 1, p. 77-86, 2021.

Menekse, M. Envisioning the future of learning and teaching engineering in the artificial intelligence era: Opportunities and challenges. *Journal of Engineering Education*, v. 233, n. 233, p. 45-59, 2023. Doi: https://doi.org/10.1002/jee.20539

Moore, S.K.; Schneider, D.; Strickland, E. How Deep Learning Works: Inside the Neural Networks that Power Today's AI. *IEEE Spectrum*, v. 10, n. 58, 2021. Doi: https://doi.org/10.1109/MSPEC.2021.9563965

Roberts, G.J. Tall trees; weak roots? A model of barriers to English language proficiency confronting displaced medical healthcare professionals: *Language Teaching Research*, v. 27, n. 4, p. 820-836, 2023. Doi: https://doi.org/10.1177/1362168820968366

Ruiz, S.; Rebuschat, P.; Meurers, D. The effects of working and declarative memory on instructed second language vocabulary learning: Insights from intelligent CALL: *Language Teaching Research*, v. 25, n. 4, p. 510-539, 2021. Doi: https://doi.org/10.1177/1362168819872859

Shiranirad, M.; Burnham, C.J.; English, N.J. Machine-learning-based many-body energy analysis of argon clusters: Fit for size? *Chemical Physics*, v. 552, e111347, 2021.

Xin, D. Application Value of Multimedia Artificial Intelligence Technology in English learning Based on literature retrieval Practice. *Hindawi Limited*, v. 45, n. 45, p. 23-40, 2021.

Xu, Z. *et al.* The effects of web-based text structure strategy instruction on adult Chinese ELLs' reading comprehension and reading strategy use. *Language Teaching Research*, v. 28, n. 4, p. 1288-1310, 2024. Doi: https://doi.org/10.1177/13621688211022308

Yang, C. *et al.* Machine Learning-Based Student Modeling Methodology for Intelligent Tutoring Systems: *Journal of Educational Computing Research*, v. 59, n. 6, p. 1015-1035, 2021. Doi: https://doi.org/10.1177/0735633120986256

Ye, Y. *et al.* Improving machine learning-based phase and hardness prediction of high-entropy alloys using Gaussian noise augmented data. *Computational Materials Science*, v. 67, n. 67, p. 78-99, 2023. Doi: https://doi.org/10.1016/j.commatsci.2023.112140

Zhang, B. *et al.* Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers. *Journal of Artificial Intelligence Research*, v. 71, n. 71, p. 45-60, 2021. Doi: https://doi.org/10.1613/jair.1.12895