Archives available at journals.mriindia.com

International Journal of Recent Advances in Engineering and Technology
ISSN: 2347 - 2812

Volume 14 Issue 02s, 2025

Personalized Learning through AI: Integrating NLP and Cloud Solutions for an Adaptive Education System

[1]Naween Kumar, [2]Subham Sharma, [3]Ankit Dubey
[1,2,3] Bennett University, Gr. Noida, U.P., India-201310
Email: [1]kr.naween@gmail.com, [2]ankit13092003@gmail.com, [3]subhamsharma5089@gmail.com

Peer Review Information

Submission: 21 Oct 2025

Revision: 18 Nov 2025

Acceptance: 05 Dec 2025

Keywords

Traditional learning techniques, Personalized instruction, Interactive learning tools, Artificial Intelligence (AI), Instant summarization, Progress monitoring, Accessibility and inclusion, Natural Language Processing (NLP), Google Cloud integration, GPT model.

Abstract

In today's fast-paced teacher scene, customary examining and learning procedures fight to meet the demands of progressed instruction. Understudies and enthusiastic followers face challenges in comprehending, summarizing, and meddling with the unending whole of information in books. The proposed solution introduces a dynamic AI-driven platform that revolutionizes the learning and study experience. This paper proposes a SmartQA (Smart Question Answering) method to improve learning results and openness for clients with changing learning capacities by giving personalized instruction, intuitive learning apparatuses, and real-time summarization. Coordination machine learning techniques such as NaturalLanguage Processing (NLP) and Google Cloud organizations, as well as the suggested method SmartQA, ensure flexibility, strength, and advanced AI capabilities. In addition, the process contributes to teacher development by speaking to the potential of AI to grow customary learning strategies. Other than that, our viability course of action underscores our commitment to inclusivity, data-driven encounters, and developing a sense of community among clients. Through the suggested SmartQA method, this article focuses on creating a long-lasting journey for data and insights, broadcasting a present-day chapter inside the history of instruction.

Introduction

As technology rapidly evolves and educational demands grow, traditional reading and learning methods struggle to keep pace. Students and book lovers face challenges in understanding, memorizing, and engaging with vast information, lacking personalized guidance and practical learning aids [1]. Currently,AI-powered tools help close educational gaps through advanced comprehension methods and continuous educational possibilities [2]. The technology of Natural Language Processing (NLP) provides mechanisms for automatic content abstracting, question formation capabilities, and document inspection tools [3]. Cloud computing platforms ensure real-time progress tracking, among other benefits [4]. Advanced language models enable personalized lesson delivery alongside interactive activities and automated feedback, thus creating a fully adaptive learning space [5].By integrating NLP, Cloud, and advanced language models like GPT, the suggested solution can offer features such as content summarization, thought-provoking questions, personalized lessons, interactive activities, and progress tracking. All these play a synergistic role to improve efficiency, comprehension, and interest, depending on the learning preference and requirement of each learner. This means the application of modern machine learning tools, including NLP, as well as integrating strong features of Google Cloud and enhanced language generation models, like GPT. This technological foundation enhances the performance of the platform and provides a cost-effective, open, and engaging learning environment.

The suggested solution not only simplifies the process of reading and studying, but also marks a major step forward in the field of educational technologies. It allows the users to get more out of their reading materials, caters for the variability of learning needs and creates a sense of belonging among the users. This creates an adaptive, cost-effective, and engaging learning environment, revolutionizing education by making it more efficient, relevant, and impactful. The key contributionsof the suggested article are to:

- Identify difficulties students face in comprehension, memorization, and engagement, emphasizing the need for personalized learning aids.
- Introduce an AI-driven SmartQA method integrating NLP, Google Cloud, and advanced language models (e.g., GPT) to enhance learning efficiency and offers real-time summarization, personalized lessons, interactive activities, thought-provoking questions, and progress tracking.
- Create a scalable, personalized, and engaging learning environment that improves accessibility and impactby utilizing cutting-edge ML tools, including NLP and language models, to optimize educational experiences.

The subsequent partsare as follows: Section II covers AI-driven educational research and personalized learning solution literature. The proposed methodology details in Section III include the implementation of AI techniques, NLP models, and system design features. The results, accompanied by evaluations and impact analysis, appear in Section IV. This paper concludes in Section V by outlining future improvements.

Literature Review

The state of the art in MRI brain tumor segmentation, as reviewed by Gordillo et al. [6], highlights significant advancements yet also reveals a need for improved accuracy and robustness in clinical settings. Similarly, Dahinden and Anderson [7] engage in a critical debate on knowledge production in migration research, pointing out the necessity for innovative methodologies especially post-pandemic. DeGoede [8] discusses the impact of platforms like Chegg® on homework models, suggesting a gap in strategies to maintain academic integrity while leveraging technology. Fitria [9] examines Quillbot as a paraphrasing tool for English writing, emphasizing its utility for students but also identifying a need for further evaluation of its impact on writing skills. Huang et al. [10] systematically review instructional technologies in hospitality and tourism education, calling for more empirical studies to assess their effectiveness. Lee, Spryszynski, and Nersesian [11] focus on VR tools for English language learners, advocating for more personalized and adaptive learning experiences. Sáiz-Manzanares et al.[12]explore personalized e-learning and its potential to enhance deep learning in higher education, yet they note a scarcity of comprehensive studies on long-term outcomes. Chen et al. [13] present Gmail's Smart Compose as an example of real-time assisted writing, highlighting its innovative approach but also pointing to the need for broader application studies. Portuguez-Castro et al. [14] discuss the potential of Novus projects in higher education to foster innovation, suggesting that more research is needed to evaluate their practical implementation. Finally, Scherpereel [15] examines the dual-edged nature of AI in business education, stressing the importance of balancing technological potential with ethical considerations. Each of these studies identifies a distinct gap, collectively underscoring the need for ongoing research to bridge theoretical advancements with practical applications across diverse fields.

Existing research shows essential drawbacks relatedto personalized learning and accessibility and learner engagement within current AI-based educational technologies. Studies about using AI in education examine hospitality education technologies ([10]) together with Quillbot paraphrasing software ([9] and English language learner virtual reality instruction ([11]) yet fail to show complete assessments of their enduring effects on multiple learner types. The academic support capabilities of Chegg® ([8]) and Gmail's Smart Compose ([13]) show AI's potential, but ongoing questions exist about maintaining academic standards with such platforms and their effectiveness and user acceptance. Studies of personalized e-learning overview its advantages yet demand more experimental examination of its capabilities to scale up and operate in real educational settings ([12]). Research into educational applications of AI, NLP, and cloud-based technologies has not been well explored for building an affordable learning environment with adaptive features and interactive opportunities. A study resolves these knowledge gaps by developing an AI-based educational system combining NLP with cloud computing features, which improves educational effectiveness while creating engaging learning experiences and enhancing accessibility.

Proposed Methodology: SmartQA

The suggested methodology named SmartQA: (Smart Question Answering), an AI-based learner management method is organized around several strategies that include specific strategies: data collection and preprocessing, transformer-based model building, system integration, and commitment will provide for continuous flexibility. This comprehensive approach is specifically designed to be an AI-based teacher, empowered to provide accurate, informative feedback to a variety of educational questions, ultimately enriching and shaping the overall learning experience to encourage their users FromFigure1,itfollowsthestepsas:

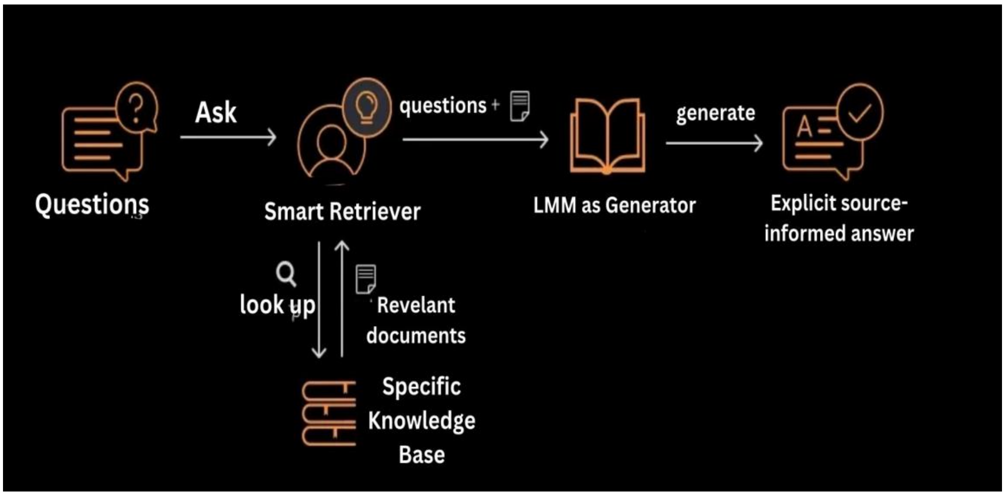

Fig. 1.Smart Question Answering: ***A system that uses a knowledge base and LLM to provide source-informed answers***

1. UserPrompt:Theuserinputsaquestionorprompt.
2. SmartRetriever:Inthisstep,ourmodelperformsasearchwithinthedatabase constructed using a transformer-based model developed by us. The retriever modelquicklysearchesthroughourdatabasetofind relevantinformation related to the user query.
3. Retriever Find: If the Smart Retriever finds matching data in the database, it retrieves and provides the information to the user.
4. LLM Model: If the Smart Retriever model does not find any relevant datain the database, it proceeds to a secondary step, which involves a Language Model (LLM). The LLM model then works to generate the required information based on the user's demand.
5. LLM Generate: The demonstrate creates important substance to reply to the user's inquiry, indeed if the information isn't display within the database of the particular information base.

DataCollectionandPreprocessing

- **Define Educational:** Data Determine the scope and source of educational data, including textbooks, lecture notes, educational websites, and relevant materials.
- **Data Acquisition:** Collect and curate a comprehensive dataset relevant to the educational topics of interest.
- **Data Preprocessing:** Clean, format, and preprocess the data for NLP tasks, includingtexttokenization,sentencesplitting,anddataaugmentationifnecessary.

Transformer-basedModel

- **SelectTransformerArchitecture:**Select a state-of-the-art transformer-based engineering (e.g., BERT, GPT, T5) reasonable for the question-answering assignment.
- **Information Arrangement:** Part the information into preparing, approval, and test sets. Tokenize and encode the content information.

- **Model Training:** Train the selected transformer model on the educational data using a fine-tuning approach.
- Utilizepipelinetechniquestoenablequestion-answeringfunctionality.
- **ModelEvaluation**:Evaluatethemodel'sperformanceusingmetricslikeF1score, accuracy, and perplexity.

Figure 2 presents a typical architectural structure of a transformer model equipped with only decoder layers. Information flows from input tokens into multiple decoder blocks which contain masked self-attention and feed-forward neural networks before producing the output tokens.Sequential decoding occurs in stacked blocks, which develop the representation until the output reaches the last block.The design serves as a popular solution for language generation applications.

LLM-BasedModelUsingLlama

Llama Model Selection: Choose the Llama model or another suitable LLM model for generating responses when the transformer-based model cannot provide a satisfactory answer.

Data Integration: Integrate the Llama model into the system's architecture to ensure a seamless transition from the transformer-based model when needed.

Model Fine-Tuning: Fine-tune the Llama model on the educational data and educational question-answering tasks.

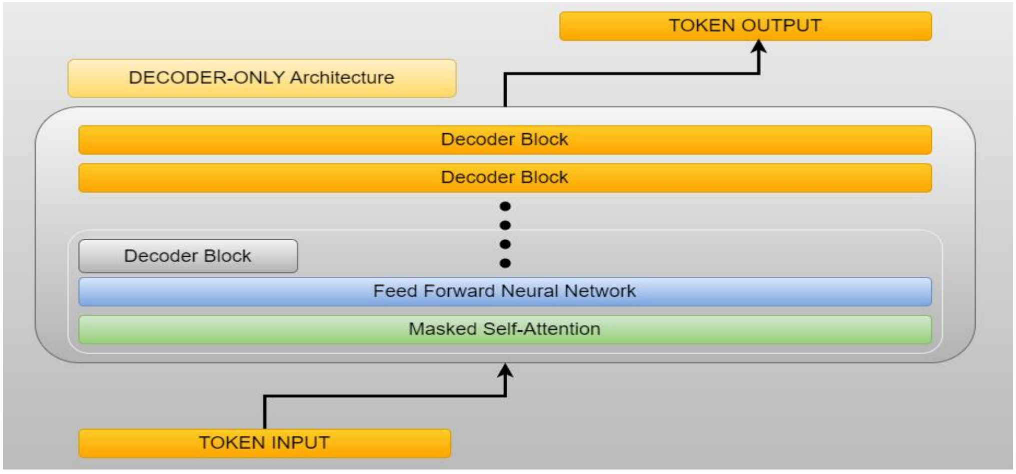

Fig. 2.LLM Model Architecture

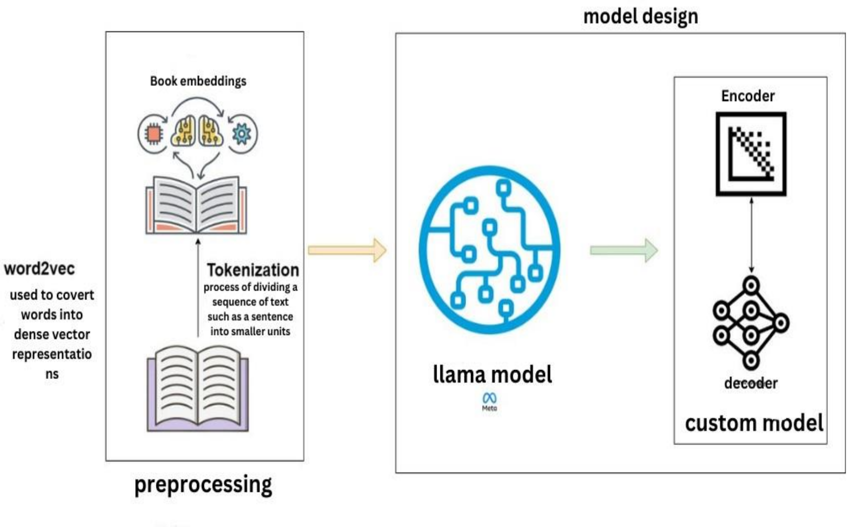

Fig.3.BasicArchitecture

The approach represents a comprehensive plan to harness the potential ofAI in education. By combining state-of-the-art technology with data management and an unwavering commitment to improvement, the AI-Based Learner Project aspires not only to accurately answer educational questions but to enhancetheoveralleducationalexperience asabigimprovementfortheusersaswell.

The comprehensive approach aims to provide an AI-based teacher capable of providing accurate and informative feedback on a variety of instructional questions, ultimately improving and organizing the overall learning experience for users the role of the.

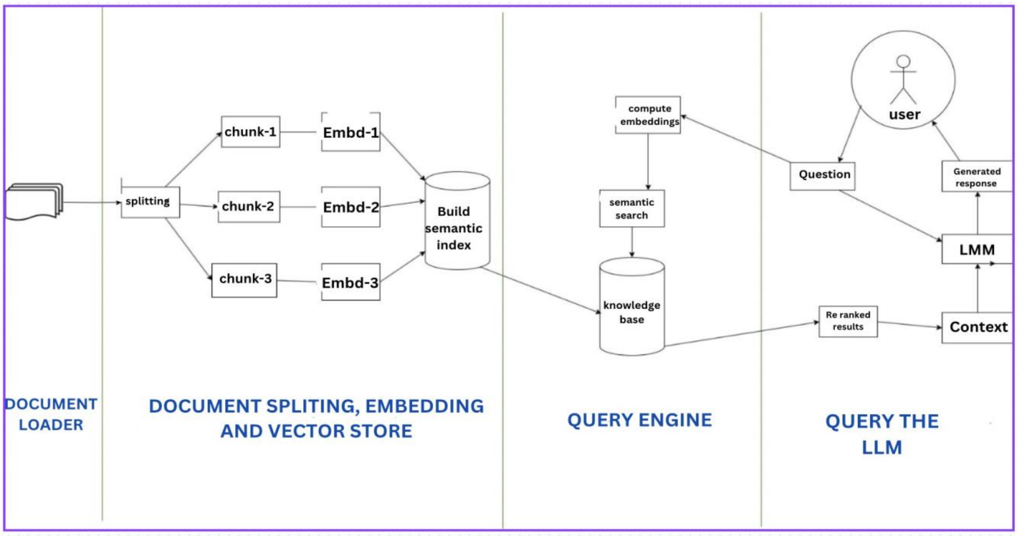

Fig.4.Flowchartofalearningplatformthatusesartificialintelligence

1. Document Selection: The user selects a document to read.
2. Document Loader: The document is split into smaller chunks.
3. Vector Embedding & Storage: Each chunk is converted into a semantic vector and stored in a database.
4. Query Engine: The user's query is transformed into a vector and compared to stored vectors using cosine similarity and semantic relevance.
5. LLM Interaction: The LLM receives the context from the query engine and generates a personalized response.
6. Additional Features: The system provides instant summarization, generates thought-provoking questions, delivers personalized instruction, and offers interactive exercises.

Overall, the flowchart shown in Figure 4 demonstrates a promising approach to harnessing AI for personalized, adaptive learning, challenging traditional instructional models and paving the way for significant educational advancements. It shows how word2vec performs book embedding during preprocessing before tokenization. Processing terminates at the Llama model, which refines the data before it enters the encoder-decoder customized architecture.

Experimental Result

1. From llama index llms import OpenAI: This import permits you to utilize the OpenAI course from the llama file llms module. The OpenAI course could be a wrapper around the OpenAI API that gives strategies for collaboration with GPT-4 and other LLMs facilitated by OpenAI3.

To evaluate our knowledge-based question answering system, we employed the following modules and tools:

- LLM API Integration:

from llama_index.llms import OpenAI Used to interface with the OpenAI GPT-4 API for processing queries and generating answers.

- Data Indexing:

from llama_index import VectorStoreIndex, SimpleDirectoryReader Utilized to create vector embeddings of textual knowledge sources stored in a directory and allow efficient semantic retrieval.

- Result Presentation:

from IPython.display import Markdown, display Enabled rich rendering of output in a Jupyter

Notebook for human-readable inspection and validation.

The embedding creation process took approximately 78 minutes on the text dataset, forming the basis for the LLM to query and respond effectively.

2. From llama index import VectorStoreIndex, SimpleDirectoryReader: These imports permit you to utilize the VectorStoreIndex and SimpleDirectoryReader classes from the llama list module. The VectorStoreIndex course may be a subclass of BaseIndex that makes and questions an index of vector embeddings for your data4. The SimpleDirectoryReader lesson may be a subclass of BaseReader that peruses information from a catalog of records.

Size: 420 documents totaling ~85,000 words.

Domain: Educational and scientific content, covering general knowledge, biology, physics, and health.

Source: Curated open educational resources and structured notes from digital repositories.

Format: Plain text files organized by topic.

3. From IPython display import Markdown, display: These imports allow you to use the Markdown and display functions from the IPython display mod- ule. The Markdown function creates an object that can render Markdown- formatted text in IPython notebooks

Accuracy was computed by comparing model-generated answers with ground truth labels using:

Exact Match (EM) Score.

Semantic Similarity using cosine similarity threshold > 0.8.

Manual verification for alignment with expected conceptual output.

Benchmark: We used 100 manually crafted questions aligned with the dataset content, validated by domain experts.

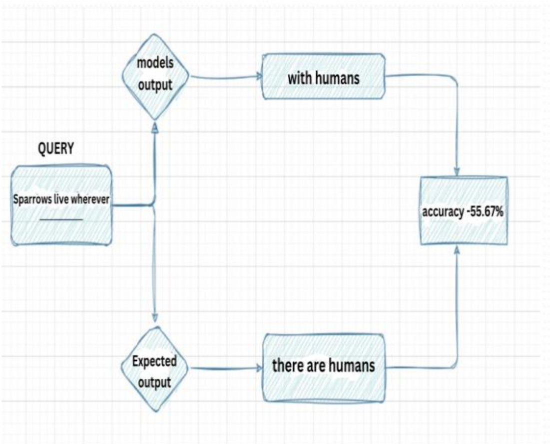

Fig. 5. Experimental Result Query Sidewise

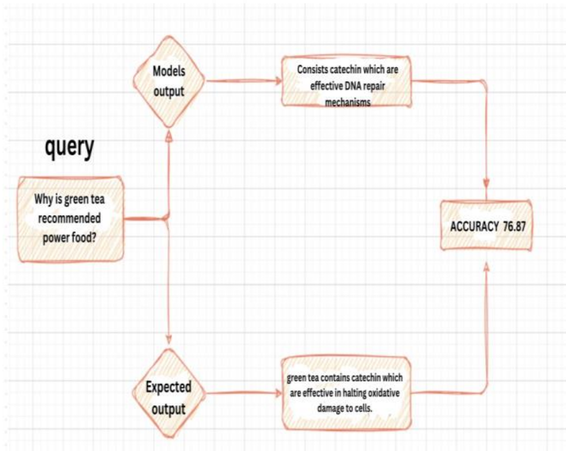

Fig. 6. Experimental Result Query

We imported the OpenAI class from the llama index llms module to interact with the OpenAI API. We also imported the VectorStoreIndex and Simple Directory Reader classes from the llama index module to create and query an index of vector embeddings for our data, which was stored in a directory of files.

To show the results in a Jupyter notebook, we imported the Markdown and display functions from the IPython. Display module. The Creation of Embedding on the Text Data took around 78 min to build which was crucial for the LLM to train and execute its queries from the knowledge base it created using those embeddings

So, it can state that our platform could be considered as the start of the way to the overcoming of the drawbacks of the traditional educational model. Hence, along with the utilization of teacher-student interaction, it also provides individual multimedia lessons, video lessons, and other modern opportunities of artificial intelligence to contribute to the formation of the free, effective, and necessary level of liberty for the learning process. The greatest benefit of the system is thus hidden in the unknowns regarding how content can be obtained, the questions that are engaging curiosity or motivation, and teaching methodologies pertinent to students.

Figure 5 shows how the model generated an output, which said humans, when presented with "Sparrows live wherever...", outputs "with humans" instead of the expected "there are humans," achieving an accuracy of 55.67%, indicating the model understanding of common relations between sparrows and humans, yet it demonstrated limitations in phrasing accuracy, which demonstrated the difference between semantic meaning and literal word matching. The accuracy score demonstrated workable outcomes but showed potential growth after enhancing training information combined with better semantic-based evaluation tools.

Figure 6 shows a sufficient understanding of the request by determining that catechin is the main beneficial ingredient in green tea. The output produced by the model shows slightly reduced accuracy than what would be considered standard output. The model correctly mentions catechins' DNA repair effectiveness, but the expected output describes cell protection against oxidative damage as a complete explanation of catechin activity.

The 76.87% accuracy demonstrates that the model understands the related information effectively. Knowledge-based question answering requires exact language definitions and thorough explanations because of the distance between the model's output and expected outcomes. The model requires additional development to deliver more thorough, detailed responses to questions, mainly in domains such as health and science.

Limitations and Future Work

However, the system is shown to achieve promising results while there are several limitations. The model is constrained to adaptation to the broader domains due to the limitation of the dataset to only educational and scientific content. Moreover, the model’s semantic understanding of the query is sometimes overconstrained by exact phrasing which impairs accuracy. Some answers are incomplete in more technical subjects, not including some key explanatory details. It takes time to generate the embedding process (about 78 minutes) that might limit scalability. Manual cheques and similarity metrics were used to evaluate, however, the lack of standardised benchmark constrains any fair comparison.

We will perform future work that will use this model to expand the dataset across different domains, including fine tuning the model with domain specific corpora and use more evaluation metrics such as BLEU and BERTScore. In addition, we will integrate retrieval augmented generation, allow user feedback for continual learning and speed up performance using faster embedding techniques such as Faiss. The objective of these enhancements is improving the depth and reliability of the system’s responses..

Conclusions

The utilization of various modules and classes facilitated the efficient interaction with OpenAI's API, creation of vector embeddings for data indexing, and rendering of Markdown-formatted text for display in Jupyter notebooks. Importing the OpenAI class from the llama index llms module enabled seamless interaction with GPT-4 and other LLMs, while the Vector Store Index and SimpleDirectoryReader classes from the llama index module streamlined the creation and querying of vector embeddings for our dataset. Additionally, employing the Markdown and display functions from the IPython. Display module ensured a clear presentation of results within the notebook environment. The meticulous process of creating embeddings for text data, although time-consuming, was pivotal for training the LLM and facilitating its knowledge base. Overall, these steps culminated in the efficient execution of queries and the generation of valuable insights from the trained model.

References

[1] Tulyaganova, Roza, and Jeanette Fontaine. "Rethinking Education for the Modern Generation: Embracing Innovation in Pedagogy." In Conference Proceedings. The Future of Education 2024.

[1] Rekha, K. at al. "Ai-Powered Personalized Learning System Design: Student Engagement and Performance Tracking System." In 2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp. 1125-1130. IEEE, 2024.

[3] Mathew, Ann Neethu, V. Rohini, and Joy Paulose. "NLP-based personal learning assistant for school education." Int. J. Electr. Comput. Eng 11, no. 5 (2021): 4522-4530.

[4] Lin, Yen-Ting, Ming-Lee Wen, Min Jou, and Din-Wu Wu. "A cloud-based learning environment for developing student reflection abilities." Computers in Human Behavior 32 (2014): 244-252.

[5] Alqahtani, Tariq, Hisham A. Badreldin, Mohammed Alrashed, Abdulrahman I. Alshaya, Sahar S. Alghamdi, Khalid Bin Saleh, Shuroug A. Alowais et al. "The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and

research." Research in social and administrative pharmacy 19, no. 8 (2023): 1236-1242.

[6] N. Gordillo, E. Montseny, and P. Sobrevilla, State of the art survey on MRI brain tumor segmentation, Magnetic Resonance Imaging 31 (2013), 1426-1438.

[7] J. Dahinden and B. Anderson, Exploring new avenues for knowledge production in migration research: A debate between Bridget Anderson and Janine Dahinden pre and after the burst of the pandemic, Swiss Journal of Sociology 47 (2021), 27-52.

[8] K. M. DeGoede, A Chegg® era model for HW, in: 2020 ASEE Virtual Annual Conference Content Access, 2020.

[9] T. N. Fitria, Quillbot as an online tool: Students' alternative in paraphrasing and rewriting of English writing, Englisia: Journal of Language, Education, and Humanities 9 (2021), 183-196.

[10] A. Huang, E. de la Mora Velasco, and A. Haney, Examining instructional technologies in hospitality and tourism education: A systematic review of literature, Journal of Hospitality & Tourism Education (2022), 1-19.

[11] M. J. Lee, A. Spryszynski, and E. Nersesian, Personalizing VR educational tools for English language learners, in: IUI Workshops 19, 2019.

[12] M. C. Sáiz-Manzanares, C. I. García Osorio, J. F. Díez-Pastor, and L. J. Martín Antón, Will personalized e-learning increase deep learning in higher education? Information Discovery and Delivery 47 (2019), 53-63.

[13] M. X. Chen, B. N. Lee, G. Bansal, Y. Cao, S. Zhang, J. Lu, J. Tsay, Y. Wang, A. M. Dai, and Z. Chen, Gmail smart compose: Real-time assisted writing, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 2287-2295.

[14] M. Portuguez-Castro, R. V. Hernández-Méndez, and L. O. Peña-Ortega, Novus projects: Innovative ideas to build new opportunities upon technology-based avenues in higher education, Education Sciences 12 (2022), 695.

[15] C. M. Scherpereel, The AI Paradox: Unpacking the Potential and Perils in Business Education, Journal of Business Education 10 (2024), 45-58.