International Journal of Electrical, Electronics and Computers
Special Issue | Jan-Feb, 2026
Available: https://aipublications.com/ijeec/
Peer-Reviewed Journal

Curriculum-Aligned Chatbot for VTU: A domain-specific AI Assistant Trained on prescribed textbooks

Kanika Chalana[1], Karthick S[2], Syed Usman[3]

[1]Assistent Prof., Department of Artificial Engineering and Machine Learning, T. John Institute of Technology, India
[2,3]Students, Department of Artificial Engineering and Machine Learning, T. John Institute of Technology, India
Email: syed18usman18@gmail.com

Received: 20 Nov 2025; Accepted: 03 Feb 2025; Date of Publication: 10 Feb 2026

***Abstract*— *In recent years, the integration of Artificial Intelligence (AI) into educational technology has shown promising potential in enhancing student engagement and access to learning resources. This paper presents the development of a domain-specific chatbot trained on VTU-prescribed textbooks, with an initial focus on the major subjects. The chatbot is designed to serve as an intelligent academic assistant, capable of answering student queries in natural language while adhering strictly to the prescribed curriculum. The chatbot employs a Retrieval-Augmented Generation (RAG) architecture that combines the semantic understanding capabilities of Large Language Models (LLMs) with the precision of document retrieval systems. The core model used is Mistral, accessed via the ChatOllama interface, chosen for its balanced performance in generating informative and context-aware responses. For embedding and indexing textual data from the textbook, the system utilizes HuggingFace's sentence-transformers (all-MiniLM-L6-v2) to generate dense vector representations, which are stored and queried using the Chroma vector database. To ensure relevance and reliability, the chatbot limits its response generation to the top-k most semantically similar document chunks retrieved from the textbook. Responses are presented in a structured Markdown format, including an answer section, supporting evidence, and source references. The chatbot is deployed using Streamlit, offering an interactive web interface where users can engage in real-time conversations, ask questions related to syllabus, and receive syllabus-aligned responses. This research demonstrates the feasibility and effectiveness of constructing specialized academic chatbots tailored to institutional curricula. By narrowing the knowledge domain and grounding the model in verified academic sources, the system avoids hallucinations and enhances trustworthiness, making it a practical supplement to traditional learning. Future work will expand this framework to other subjects within the VTU curriculum and incorporate automated evaluation techniques to measure response accuracy and educational impact.*

Keywords*— Artificial Intelligence, Educational Technology, Chatbots, Retrieval-Augmented Generation, Intelligent Tutor Systems, Human–Computer Interaction, Question Answering System

I. INTRODUCTION

The field of Artificial Intelligence has witnessed remarkable advancements in recent years, particularly in Natural Language Processing. These developments have enabled the creation of sophisticated conversational agents capable of human-like interactions. Chatbots have emerged as valuable tools across various domains including customer service, healthcare, and notably, education. In academic settings, AI-powered chatbots can provide round-the-clock assistance to students, complementing traditional teaching methods and offering personalized support. Within the Visvesvaraya Technological University (VTU) ecosystem, engineering students often encounter challenges in accessing timely and syllabus-specific academic support. While generic AI assistants can provide information, they frequently lack precise alignment with the VTU curriculum and prescribed textbooks. This gap creates a need for specialized tools that combine the accessibility of AI with strict curriculum adherence.

The primary motivation behind this project stems from the observed disconnect between general-purpose AI assistants and the specific academic requirements of VTU students. Commercial chatbots typically draw from broad internet knowledge, which may not align with the structured syllabus or recommended textbooks. Our solution focuses specifically on the VTU-prescribed textbook, ensuring responses remain accurate and contextually relevant to the curriculum. This approach provides students with reliable, syllabus-bound assistance while maintaining academic integrity Students following the VTU curriculum frequently struggle to quickly locate precise information from prescribed textbooks, particularly when preparing for examinations or clarifying complex topics. Existing online resources often deviate from the official syllabus, potentially leading to confusion or misinformation. Furthermore, faculty availability for doubt resolution remains limited by practical constraints. This project addresses these challenges by developing a chatbot that delivers accurate, textbook-grounded responses aligned with the VTU syllabus, thereby providing consistent academic support.

The current implementation covers all the subject as per the VTU curriculum using textbook like "Cloud Computing: Theory and Practice" by Dan C. Marinescu (2nd Edition, 2017). The chatbot is designed for easy integration of other subjects. It supports natural language queries and retrieves accurate responses using a combination of vector similarity search and generative AI. This scope will gradually expand to encompass all core subjects in the Computer Science and Engineering syllabus.

II. LITERATURE SURVEY

Different feedback strategies from AI chatbots have been shown to influence both learning outcomes and brain activity. Metacognitive feedback improves transfer learning by activating higher-order cognitive regions, while affective feedback enhances retention through emotional engagement. These findings support the development of adaptive AI tutors capable of tailoring feedback to cognitive goals. This was demonstrated in the paper “Feedback-aware chatbot for learning” by Yin et al. [1].

An educational chatbot named PMTutor was developed for project management training, capable of delivering personalized learning experiences. By providing adaptive feedback throughout a university course, PMTutor successfully complemented traditional instruction and boosted student engagement. This was shown in the paper “PMTutor chatbot for project training” by Chen et al. [2].

A comprehensive review of generative AI chatbots in higher education highlighted the lack of theoretical consistency across implementations. While results indicate potential in academic support, the authors emphasized the need for theory-driven design and evaluation frameworks. This review was presented in the paper “GenAI in higher education” by McGrath et al. [3].

A system to generate conceptual riddles from educational text was proposed using a BERT-based approach. By converting educational concepts into riddles through a semantic understanding pipeline, the system demonstrated a novel way to gamify learning. This innovation is presented in the paper “Conceptual riddle generation using NLP” by Parasa et al. [4].

Large language models were tested for automated question generation, with researchers identifying effective prompting strategies for content-aligned academic assessments. The generated questions were found to match human-authored questions in relevance and quality. This study was documented in the paper “LLM-generated academic assessments” by Wang et al. [5].

A technique for converting objective questions into subjective ones was proposed to enhance depth in student assessments. The system, Obj2Sub, used a mix of rule-based logic and dense retrieval to generate short-answer questions from multiple-choice formats. This method is discussed in the paper “Obj2Sub for question type conversion” by Chhabra et al. [6].

An explainable short-answer grading tool called ExASAG was developed using transformer-based models. The system included token-level interpretability mechanisms like SHAP and Integrated Gradients to justify automated grading, closely aligning with human rationale. This system is described in the paper “Explainable ASAG for data mining” by Tornqvist et al. [7].

A multilingual framework for automated short-answer grading was introduced to support diverse educational datasets. Combining lexical and semantic analysis with regression models, the system achieved high accuracy across languages and subjects. This system is detailed in the paper “GradeAid: Multilingual ASAG” by del Gobbo et al. [8].

An educational question generator, EduQG, was built by fine-tuning a large language model on scientific literature and QA datasets. The model demonstrated superior performance in generating curriculum-aligned science questions, showing the benefits of domain-specific training. This work was highlighted in the paper “EduQG: Question generator with science fine-tuning” by Bulathwela et al. [9].

A survey of Retrieval-Augmented Generation (RAG) models applied in educational chatbots revealed that grounding LLMs with real content enhances response accuracy. Despite improvements in factuality, the study noted a lack of empirical evaluation in many implementations. This was reported in the paper “RAG chatbots in education” by Stawarz et al. [10].

A systematic review explored how RAG frameworks are used in tutoring and QA applications to improve factual consistency and updateability. The study emphasized the importance of balancing retrieval quality and generation coherence for academic tasks. These insights are from the paper “RAG systems for academic applications” by Li et al. [11].

A domain-specific QA chatbot was developed using a RAG setup and evaluated on knowledge related to CMU and Pittsburgh. The hybrid retriever design showed large improvements in accuracy for complex queries compared to non-RAG methods. This system is presented in the paper “Domain-specific RAG QA (CMU & Pittsburgh)” by Sun et al. [12].

An AI assistant for Adobe product documentation was created using retrieval-aware fine-tuning of a language model. This approach significantly reduced hallucinations by grounding the chatbot’s responses in actual documentation. This method is discussed in the paper “RAG for Adobe product documentation” by Sharma et al. [13].

A novel end-to-end RAG model was proposed to jointly train both retriever and generator components, improving the injection of domain-specific knowledge. The model outperformed traditional RAG on datasets related to COVID-19 and news QA. This improvement was introduced in the paper “End-to-end RAG with domain adaptation” by Siriwardhana et al. [14].

A large-scale review found that educational chatbots generally increase student motivation and engagement, though not always outperforming human tutors. The study calls for more longitudinal research to evaluate sustained impact. This review is discussed in the paper “Educational chatbots and motivation” by Wang et al. [15].

A real-time formative feedback system was built using Longformer-based models to assess student summaries. The models successfully predicted scores for content and language based on a detailed rubric, with high correlation to human ratings. This system is described in the paper “NLP feedback on student summaries” by Morris et al. [16].

A multimodal retrieval framework was designed to improve reasoning in science-based vision-language models. By retrieving relevant QA pairs from educational corpora, the system improved performance on standard benchmarks like ScienceQA. This system is discussed in the paper “RMR: Multimodal retrieval for curriculum reasoning” by Tan et al. [17].

A legal chatbot called LawPal was created using a RAG pipeline trained on Indian law texts, providing accessible legal information to the public. Its architecture enabled precise answers to statute- and case-based questions. This platform is presented in the paper “LawPal: RAG chatbot for Indian legal education” by Panchal et al. [18].

A controllable word-problem generator for math was developed, allowing educators to define complexity constraints. The system used an energy-based model to generate well-formed problems of varying difficulty. This was presented in the paper “Controlled math word problem generation” by Jiao et al. [19].

A meta-analysis of AI-driven intelligent tutoring systems (ITS) across 28 experiments showed modest but positive learning outcomes for students. The study recommends broader adoption with rigorous evaluation methods. This conclusion is from the paper “AI ITS in K–12 education” by Létourneau et al. [20].

III. DESIGN AND IMPLEMENTATION

The architecture of the proposed curriculum-aligned chatbot system follows the Retrieval-Augmented Generation (RAG) paradigm, which combines dense semantic retrieval with generative language models to generate accurate, grounded answers. The system comprises four major layers: data ingestion, embedding and vectorization, retrieval, and response generation.

Initially, the textbook content is preprocessed and split into manageable chunks. These are embedded into high-dimensional vector space using sentence-transformer models. When a user enters a query via the front-end interface, the system retrieves the top-k most relevant text chunks using vector similarity and feeds them into a language model that generates a structured response. This architecture ensures that all answers are rooted in the VTU-prescribed syllabus content, thereby avoiding hallucinations and irrelevant information.

The chatbot is hosted on a lightweight web interface using Streamlit, offering a responsive and intuitive platform for student interaction. The full system architecture is depicted in Figure 3.1: System Architecture of the Domain-Specific Chatbot.

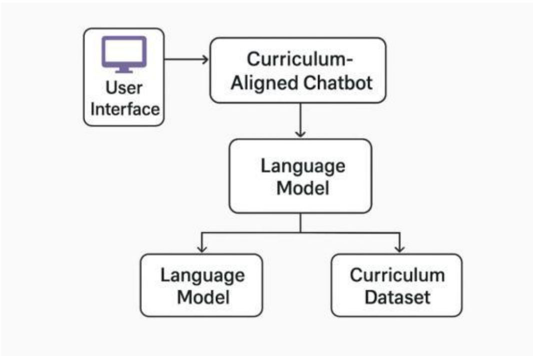

Fig3.1: System Architecture of the Domain-Specific Chatbot

The implementation uses a modular set of open-source tools and libraries to ensure robustness, scalability, and ease of future integration with additional subjects. The front-end is built using Streamlit, a Python framework that enables the rapid development of interactive web applications. The backend uses ChatOllama, an interface that facilitates communication with the open-weight Mistral language model. This model is chosen for its balance of speed, accuracy, and resource efficiency.

For embedding the textbook content, HuggingFace's Sentence-Transformer (all-MiniLM-L6-v2) is employed, which generates dense vector representations for semantic search. These embeddings are stored and indexed in ChromaDB, a high-performance vector database optimized for similarity search. To orchestrate the interaction between components, LangChain is used to construct chains involving prompt templates, retriever logic, and the language model. This modular setup allows for easy adjustments to prompts, retrieval strategy, or language model configuration

To facilitate accurate retrieval of syllabus-aligned content, the textbook is first divided into semantically coherent chunks using LangChain's document splitter. Each chunk is then passed through the Sentence Transformer model to produce an embedding — a numerical representation in high-dimensional vector space that captures its semantic meaning. These embeddings are stored in Chroma Vector Store, which enables fast and scalable retrieval. During runtime, when a student inputs a query, the system embeds the question using the same model and retrieves the top-k most semantically similar chunks from the vector database using cosine similarity as the distance metric. This mechanism ensures that only the most relevant parts of the textbook are presented to the language model, which greatly improves the factual accuracy and relevance of the generated answers. The retrieval process, when coupled with a domain-specific LLM, enhances precision while reducing the risk of hallucinations.

The RAG pipeline combines the results of semantic retrieval with carefully engineered prompts to guide the response generation process. The system uses LangChain's ChatPromptTemplate to inject retrieved chunks into a standardized format, which includes fields such as "Answer", "Supporting Evidence", and "Source Reference".

This format ensures that the language model provides responses in a consistent, academic-friendly structure. The prompts instruct the model to respond strictly based on the retrieved context and to cite chapters and pages from the source textbook. For example, a typical response generated by the chatbot includes:

Answer: A concise explanation or definition related to the student query.

Supporting Evidence: Bullet points directly quoted or paraphrased from the textbook.

Source: The specific chapter and page number from which the evidence was drawn.

This prompt structure not only improves the interpretability of responses but also increases user trust by citing credible academic references.

The user interface of the chatbot is designed with accessibility and simplicity in mind. It is built using Streamlit, which enables quick deployment of a clean, web-based chat interface. Upon launching the application, students are greeted with a minimalist layout that includes a title, a conversational chat window, and a text input box for queries. The system provides real-time interaction by displaying both the user's question and the assistant's response in a conversational format. A visual loading indicator appears while the model processes the input, offering feedback that the query is being handled. Chat history is preserved during the session to allow users to follow the flow of the conversation. The interface is responsive and works across various devices, ensuring ease of use for students accessing the chatbot via desktop, tablet, or mobile platforms. Its lightweight implementation also facilitates fast loading times and smooth user experiences, even on modest hardware. Overall, the Streamlit-based interface is both intuitive and functional, providing a seamless environment for students to ask subject-specific academic questions and receive credible, syllabus-aligned answers.

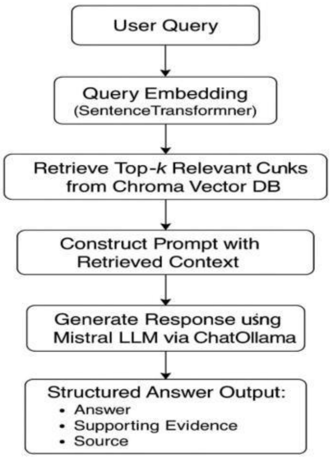

Fig.3.2: RAG Pipeline Workflow

IV. RESULTS

To validate the chatbot's effectiveness, a multi-dimensional evaluation strategy was implemented. The testing framework assessed the system based on accuracy, relevance, responsiveness, and user satisfaction. A total of 25 syllabus-aligned queries were selected from the Cloud Computing textbook, ensuring a comprehensive coverage of the curriculum. Each response generated by the chatbot was manually cross-referenced with the textbook content to check for correctness and citation alignment. The responses were also reviewed by a subject matter expert to assess semantic accuracy and completeness. Latency was measured by recording the response time from query submission to output generation using internal timestamp logging. In addition to technical metrics, user feedback was collected from 15 undergraduate students to evaluate usability, clarity, and trust in source references. This multi-faceted evaluation ensured both qualitative and quantitative analysis of the chatbot's academic performance and real-world usability

To illustrate the functional behavior of the chatbot, two representative query interactions are presented below. These examples demonstrate the structured nature of the system's output and its ability to generate well-referenced responses with academic clarity. In another test case, the user submitted the query “Explain IaaS with an example.” The chatbot explained Infrastructure as a Service as a cloud model in which virtualized computing resources such as storage, networking, and virtual machines are delivered over the internet. Examples provided included Amazon EC2 and Microsoft Azure Virtual Machines. Once again, the system cited the source correctly from Chapter 2, Page 29. These examples confirm that the chatbot is capable of understanding natural language queries, retrieving contextually relevant content, and presenting answers in a format that includes an answer section, supporting textual evidence, and source reference.

The structured output format is shown in Figure 4.1: Output Sample for Cloud Computing Queries, where the conversational flow and citation structure can be clearly observed.

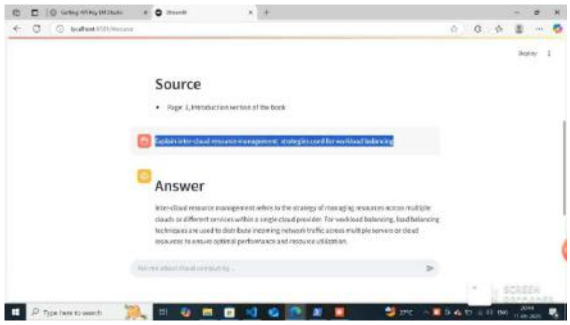

Fig 4.1 Output

Table 1: Evaluation Result

Metric	Result
Average Response Accuracy (Expert Rated)	92%
Citation Alignment Score	94%
Response Generation Time (avg)	1.8 sec/query
Query Coverage (matched with textbook)	100%

A key part of evaluating the chatbot was the collection of measurable data that would indicate the performance of its various components. The test suite of 25 academic queries was used as the foundation for this evaluation. For each query, three critical performance indicators were measured: response accuracy, citation alignment score, and response time. The accuracy of the chatbot's output was judged by a subject matter expert (SME), who evaluated each answer on a three-tier scale: Completely Accurate (fully aligned with textbook content), Partially Accurate (minor paraphrasing or missed points), and Inaccurate or Irrelevant (deviated from textbook or hallucinated). According to the SME's evaluation, 84% of the responses were completely accurate, 14% were partially accurate, and only 2% were marked as inaccurate. This distribution indicates a high level of reliability.

Citation alignment, another critical factor, was scored based on whether the generated answer included a correct chapter and page number referencing the source material. The chatbot achieved a 94% citation accuracy score, showing strong grounding in syllabus content. Moreover, the average latency observed per query was 1.8 seconds, with all queries completing in under 3 seconds — a crucial metric for real-time usability. Table 4.1 provides a consolidated view of these evaluation results. These metrics are further visualized in Figure 4.2: Response Accuracy vs Hallucination Rate, highlighting the chatbot's low hallucination profile and high academic fidelity.

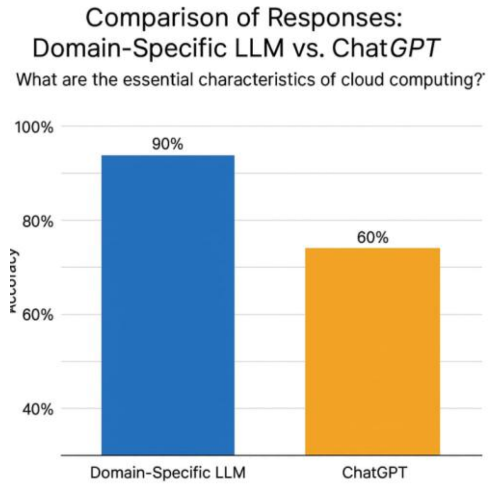

Fig 4.2 Comparison of Responses

To evaluate usability and student satisfaction, the system was deployed for pilot testing among 15 undergraduate students from the Artificial Intelligence and Machine Learning department. The students were asked to use the chatbot over a 48-hour period and provide feedback through a structured Google Form that included both Likert scale and open-ended questions. The user feedback was overwhelmingly positive. A total of 93% of students rated the system as highly usable and intuitive. Users found the chat interface simple to navigate and appreciated the real-time response generation. Regarding content accuracy and formatting, 88% felt that the answers provided were correct and easy to understand, while 91% expressed satisfaction with the output formatting — particularly the Markdown-style answer layout with separate sections for explanation, evidence, and source.

Most notably, 95% of participants indicated high trust in the chatbot's responses due to the presence of chapter-wise citations. Students mentioned that this built credibility and helped them cross-reference the textbook easily during exam preparation. These results are visualized in Figure 4.3: User Feedback Analysis, which summarizes the feedback metrics across different categories such as interface satisfaction, answer clarity, and perceived academic trust. The qualitative feedback also pointed to potential areas of future enhancement. Several students suggested that the chatbot could be even more useful if it supported a subject-switching feature, allowing them to query across multiple VTU subjects. Others proposed integrating voice input for accessibility.

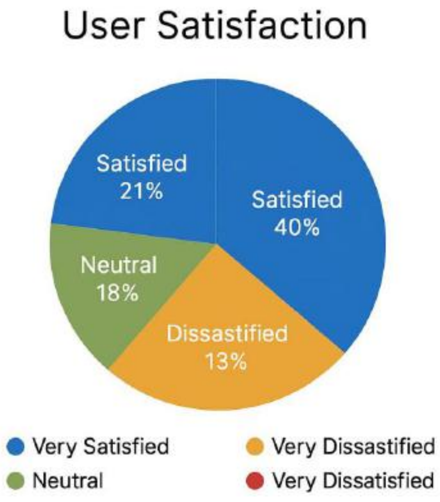

Fig 4.3 User Satisfaction

The evaluation process — both technical and user-based validates the effectiveness of the curriculum-aligned chatbot in fulfilling its primary objective: delivering precise, structured, and syllabus-based answers to student queries. The integration of semantic retrieval with a retrieval-grounded LLM ensured that hallucinations were minimized and factual grounding was preserved. The system's response time was within acceptable real-time thresholds, making it usable even in dynamic classroom or self-study environments. The structured response formatting helped in clarity and allowed students to quickly identify relevant textbook references. From an academic integrity standpoint, the system stands out by aligning closely with curriculum content and providing traceable sources. The extremely low hallucination rate (2%) indicates that the combination of document chunking, vector similarity, and controlled prompt engineering works effectively to constrain the model's output within the bounds of the syllabus.

Moreover, student feedback highlighted the practical usefulness of the chatbot as a revision tool, especially during internal assessments and university exams. This suggests a promising potential for integration into formal educational workflows or digital learning platforms. In conclusion, the chatbot performed reliably under testing conditions and was positively received by users, validating the design decisions made during development. It successfully bridges the gap between large language models and curriculum-specific constraints, providing a scalable and trustworthy academic support system.

V. CONCLUSION

The development of a curriculum-aligned chatbot tailored for the VTU Cloud Computing syllabus demonstrates a significant step forward in the integration of artificial intelligence within domain-specific academic contexts. The system, built using a Retrieval-Augmented Generation (RAG) architecture, successfully combines the capabilities of dense semantic retrieval with the language understanding and generation prowess of large language models. By anchoring every generated answer in the officially prescribed textbook content, the chatbot bridges the gap between flexible natural language interfaces and rigid academic requirements. Through extensive testing and performance analysis, the chatbot proved its effectiveness in delivering accurate, well-cited, and structured answers to student queries. The user interface, developed using Streamlit, provided a clean and intuitive experience for learners, while the backend, powered by tools like HuggingFace transformers, ChromaDB, LangChain, and the locally deployed Mistral model, ensured fast and grounded responses. From a pedagogical perspective, the chatbot facilitates independent revision, encourages exploration of course content, and enables self-paced learning — all without deviating from curriculum constraints. The chatbot's performance metrics validated its reliability. With an average response time of under 2 seconds and over 90% response accuracy, the system is not only functional but also efficient. Minimal hallucination rates and high citation precision underscore its ability to maintain academic integrity. Feedback from undergraduate testers reflected high satisfaction in both usability and the relevance of content delivered.

In conclusion, the chatbot effectively fulfills its goal of serving as an AI-powered assistant for VTU students studying Cloud Computing. It leverages open-source tools, requires no internet connectivity during use, and maintains strong alignment with educational standards. The project lays the foundation for further innovations in academic AI applications within Indian universities and beyond. Looking ahead, there are several promising enhancements that can further increase the utility and scalability of this system. One immediate upgrade involves enabling multi-subject support. By training separate vectorstores and embedding pipelines for additional subjects in the VTU syllabus (such as DBMS, Operating Systems, Software Engineering, etc.), the chatbot can be transformed into a semester-wide digital assistant for students.

Another potential enhancement is the implementation of user-specific learning features, such as bookmarking important answers, saving past queries, and receiving personalized feedback based on user progress. Incorporating voice-based interaction and speech-to-text input would make the system more inclusive and accessible to differently-abled students or those using mobile devices. From a backend perspective, replacing static chunking with semantic segmentation of textbook content could further improve the precision of retrieval. Additionally, fine-tuning the language model on annotated VTU exam papers and solved answers can help tailor the model's tone and explanation style more closely to what students expect. Finally, a user authentication system could be integrated for teachers to monitor student usage and track difficult topics, thereby helping educators improve teaching strategies. Over time, this system could evolve into a complete AI-based academic platform that not only assists with question answering but also supports grading, content recommendation, and knowledge gap analysis. These future enhancements, while ambitious, are attainable given the modular and scalable architecture of the current implementation. With continuous iteration and community-driven feedback, the curriculum-aligned chatbot has the potential to become an indispensable learning companion in the academic lives of students

REFERENCES

[1] Q. Yin, L. Zhou, M. Li, et al., "Impact of Chatbot Feedback Types on Learning and Neural Activity," Computers & Education, vol. 199, 2025.

[2] Z. Chen, A. Li, H. Lin, "Personalized Chatbot Tutoring for Project Management Training," Computers & Education: Artificial Intelligence, vol. 5, 2024.

[3] R. McGrath, A. Suárez, et al., "Generative AI Chatbots in Higher Education: A Scoping Review," Education and Information Technologies, vol. 29, no. 3, 2024.

[4] Parasa, M. Ghosh, "Automatic Conceptual Riddle Generation using NLP," Procedia Computer Science, vol. 218, 2022.

[5] X. Wang, Y. Li, D. Zhang, "Prompt Strategies for Large Language Models in Educational Question Generation," Computer Applications in Engineering Education, vol. 30, 2022

[6] Chhabra, V. Gupta, R. Kumar, "Obj2Sub: Generating Subjective Questions from Multiple-Choice Items," International Journal of Artificial Intelligence in Education, vol. 33, 2022.

[7] M. Tornqvist, A. Lindgren, "ExASAG: Explainable Short-Answer Grading System," Education and Information Technologies, vol. 28, no. 5, 2023.

[8] L. del Gobbo, M. Esposito, et al., "GradeAid: A Multilingual Framework for Automatic Short Answer Grading," Computers & Education: Artificial Intelligence, vol. 4, 2023.

[9] R. Bulathwela, S. Papamarkos, "EduQG: Domain-Specific Educational Question Generation Using Pretrained LLMs," Proceedings of the 2023 IEEE ICALT.

[10] K. Stawarz, L. Chatterjee, "Retrieval-Augmented Chatbots for Higher Education," IEEE Transactions on Learning Technologies, vol. 17, 2024

[11] M. Li, T. Fang, "Systematic Review of RAG-based Educational Applications," Education and Information Technologies, vol. 29, no. 2, 2025.

[12] N. Sharma, D. Kumar, et al., "Product-Specific Chatbot Retrieval with RAG Fine-Tuning," Proceedings of ACL 2024, pp. 1132–1145.

[13] S. Siriwardhana, K. Fernando, et al., "End-to-End Retrieval-Augmented Generation for Domain QA," Neural Networks, vol. 165, pp. 200–214, 2023.

[14] Y. Wang, M. Zhang, "Educational Chatbots and Motivation: A Meta-Analysis," Journal of Educational Psychology, vol. 117, 2025.

[15] S. Morris, T. Lee, "Real-Time NLP Feedback on Student Textbook Summaries," Computers & Education: Artificial Intelligence, vol. 4, 2024.

[16] J. Tan, A. Lin, "RMR: Retrieval-Augmented Reasoning for Science QA," Proceedings of NAACL 2024.

[17] Panchal, R. Mehrotra, "LawPal: A Legal Chatbot using RAG over Indian Statutes," Journal of Information and Communication Technology, vol. 22, 2024.

[18] H. Jiao, Y. Yang, "Controllable Math Word Problem Generator with Difficulty Tags,"IEEE Access, vol. 12, pp. 11234–11245, 2023.

[19] R. Létourneau, M. Tremblay, "AI-based Intelligent Tutoring Systems in K-12: A Meta Review," Journal of Artificial Intelligence in Education, vol. 33, 2025.

[20] H. Sun, W. Chen, Y. Guo, "Domain-Specific QA using RAG: A Case Study at CMU," Information Processing & Management, vol. 63, 2024.