Grenze International Journal of Engineering and Technology, June Issue

SIGBOT using Natural Language Processing

B Bhuvaneshwari[1], Maheshwari V[2], Priti Sinha[3], Vinotha S[4] and Saran Kumar V[5]
[1]Assistant Professor, Electronics and Communication Engineering, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India
Email: bbhuvaneswari@cit.edu.in
[2-5]Electronics and Communication Engineering, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India
Email: 71762204030@cit.edu.in, 71762204042@cit.edu.in, 71762204063@cit.edu.in, 71762304205@cit.edu.in

***Abstract*—Chatbots are rapidly becoming indispensable tools for improving student engagement and learning outcomes, especially in the context of personalized educational support. This project focuses on developing a subject-specific chatbot powered by Natural Language Processing (NLP), designed to align with academic curriculum and provide context-aware tutoring. The chatbot pipeline begins with PDF content extraction using PyMuPDF, with Tesseract OCR applied to scanned documents. Extracted text and images are structured into a knowledge base, where content is divided into overlapping chunks and linked with metadata such as file names, page numbers, and figures. For information retrieval, SIGBOT combines semantic similarity search through Sentence Transformers (all-MiniLM-L6-v2) with TF-IDF as a fallback, ensuring both meaning-based and keyword-based query handling. Answer generation is powered by a text-to-text model (FLAN-T5), which rephrases retrieved content into student-friendly explanations. Additionally, the system is integrated with the Ollama framework, enabling efficient local deployment and management of large language models (LLMs) for faster and privacy-preserving inference.**

***Index Terms*— Natural Language Processing, Educational Chatbot, Semantic Retrieval, Vector Embeddings, OCR, Transformer Models, FAISS, FLAN-T5, Ollama, Academic Assistance.**

I. INTRODUCTION

Natural Language Processing (NLP) plays a pivotal role in enabling machines to understand, interpret, and generate human language, forming the foundation for intelligent conversational systems. However, traditional learning platforms and static digital resources often lack interactivity, contextual understanding, and personalized feedback, limiting their effectiveness in academic environments. This study aims to overcome these challenges by developing SigBot, an NLP-driven academic assistant capable of delivering real-time, contextually aware, and semantically accurate responses to subject-related queries. The proposed system leverages the Ollama framework to deploy transformer-based language models locally, ensuring data privacy, reduced latency, and efficient resource utilization. To enhance contextual understanding, SigBot integrates an embedding-based semantic retrieval mechanism that preprocesses academic materials such as textbooks and research papers. These resources are segmented and encoded into vector embeddings, allowing for high-speed semantic search and retrieval of relevant information during user interactions. When a user submits a query, the system performs a semantic similarity search across the indexed embeddings to retrieve the most contextually relevant segments. These segments are then passed to the transformer-based generative model, which synthesizes a coherent and contextually accurate response.

Grenze ID: 01.GIJET.12.2.313_21

Furthermore, an advanced caching mechanism preserves conversational memory across sessions, maintaining context continuity and improving system responsiveness.

The research involves developing the retrieval and generation pipeline, integrating it with the Ollama-based model, and conducting experiments to evaluate SigBot's accuracy, contextual understanding, and performance compared to traditional rule-based systems. The results demonstrate that the combination of semantic retrieval, local transformer deployment, and context caching offers a scalable, intelligent, and privacy-preserving approach to academic assistance. By bridging the gap between static educational resources and intelligent conversational learning, this study presents SigBot as a practical and adaptive tutoring solution for modern educational environments.

II. Literature Survey

A. The Original Author's View of the Problem

Prior research in educational technology and intelligent tutoring systems has identified several persistent challenges in delivering personalized academic support. Traditional e-learning platforms often lack adaptability to individual student needs, offering static question-answer formats that fail to promote deeper conceptual understanding. Studies have shown that human tutors remain more effective due to their ability to interpret context, tailor explanations, and maintain engagement. However, the scalability of human tutoring is limited by high labor costs and availability constraints. Moreover, while general-purpose chatbots have become increasingly popular in educational settings, they frequently suffer from limited domain specificity, inadequate context retention, and difficulties in aligning responses with structured academic curricula.

B. The Approach and Technology Used

Recent advancements in Natural Language Processing (NLP) and Large Language Models (LLMs) have inspired new approaches to intelligent tutoring. Prior systems have utilized rule-based frameworks or retrieval-augmented models that rely on pre-defined responses, but these often struggle with content comprehension and contextual relevance. Hybrid approaches combining semantic embeddings and traditional keyword search have been explored to improve accuracy in information retrieval. For response generation, transformer-based architectures such as T5 and BERT have demonstrated strong performance in educational question answering. Nevertheless, existing systems often depend heavily on cloud-based models, raising concerns about latency, privacy, and deployment costs in institutional environments.

C. Critical Viewpoints

Despite these advancements, several limitations remain. Many educational chatbots lack curriculum alignment, providing generic or inaccurate explanations that fail to support structured learning objectives. The dependence on third-party APIs or cloud infrastructures increases operational costs and limits data privacy—critical factors for academic institutions. Furthermore, most models exhibit limited interpretability, offering answers without sufficient reference to source materials, which hinders student trust and academic integrity. Integration with institutional learning management systems (LMS) and support for multimodal educational content (text, images, and scanned documents) also remain underdeveloped, reducing their applicability in real-world classroom environments.

D. How Your Work Stands Apart

Our project introduces SIGBOT, a subject-specific educational chatbot designed to deliver curriculum-aligned, context-aware tutoring through a locally deployable NLP framework. The system leverages PyMuPDF for PDF content extraction and integrates Tesseract OCR to handle scanned documents, ensuring comprehensive digitization of academic materials. Extracted data are structured into a knowledge base of overlapping text chunks enriched with metadata such as file names, page numbers, and figures for precise content referencing. For information retrieval, SIGBOT employs a hybrid approach that combines semantic similarity search using Sentence Transformers (all-MiniLM-L6-v2) with TF-IDF as a fallback, ensuring both meaning-based and keyword-level query coverage.

Answer generation is powered by FLAN-T5, which rephrases retrieved content into coherent, student-friendly explanations. Unlike cloud-dependent systems, SIGBOT integrates with the Ollama framework for efficient local deployment of large language models (LLMs), ensuring faster inference, reduced costs, and enhanced privacy. The system's modular architecture supports easy customization for different subjects and institutions, promoting scalable, sustainable, and accessible AI-driven learning support.

III. MOTIVATION

Students often face challenges in understanding complex subject concepts without interactive, context-aware support, as traditional learning materials like textbooks and lecture notes provide limited assistance and fail to adapt to individual learning styles or pace. Generic AI tools, though available, often misinterpret subject-specific terms due to their general-purpose training, leading to inaccurate responses. To address these issues, the proposed system SigBot leverages Natural Language Processing (NLP) to deliver intelligent, domain-specific which is DSP that accurately interprets technical terms, understands user intent, and generates contextually relevant explanations aligned with the syllabus. Unlike static systems, SigBot dynamically adapts to user queries, providing precise and conceptually sound responses. Deployed on the Ollama framework, it operates efficiently without cloud dependency, ensuring data privacy and accessibility.

IV. PROBLEM DOMAIN

In the proposed system, the existing challenges in digital learning platforms—such as lack of subject-specific clarity, generic AI responses, and absence of interactive conceptual guidance—are addressed through a domain-focused chatbot architecture. The system leverages Natural Language Processing (NLP) techniques and document retrieval to provide accurate, context-aware, and course-aligned answers. By fine-tuning the model for a specific academic domain and integrating semantic understanding, the chatbot minimizes vague or inaccurate responses. This ensures that learners can interact with an intelligent tutoring system capable of evolving with user input, thereby enhancing conceptual understanding and engagement.

V. PROBLEM DEFINITION

The proposed system addresses key challenges in digital learning platforms, including lack of domain-specific clarity, limited contextual understanding, and absence of interactive conceptual support. Existing AI-based tools often produce generic responses that fail to meet academic needs. To overcome these limitations, the system employs advanced NLP techniques, semantic retrieval, and transformer-based generative modeling to generate accurate, context-aware, and course-aligned answers. By integrating these components within the Ollama framework for local model deployment, the solution ensures data privacy, adaptability, and precision—offering a scalable and intelligent alternative to conventional AI chatbots for personalized academic assistance.

VI. STATEMENT

Providing accurate and context-aware answers in specialized domains is often a challenging and error-prone task when handled manually, as it requires extensive expertise and constant human effort. To overcome these limitations, domain-specific chatbots are proposed as an automated solution that delivers precise, efficient, and consistent responses by leveraging intelligent natural language processing and knowledge-driven customization.

VII. INNOVATIVE CONTENT

This project introduces an innovative approach for delivering domain-specific knowledge through intelligent chatbots that combine advanced natural language processing techniques with knowledge engineering. A transformer-based language model was fine-tuned on a curated domain dataset to accurately interpret user queries and generate context-aware responses. This enables real-time understanding and retrieval of specialized information, ensuring precise and relevant answers.

Additionally, semantic similarity analysis and intent classification were implemented to measure the alignment between user queries and domain knowledge. By analyzing linguistic input without requiring extensive manual rule creation, the system dynamically adapts to varied phrasing and terminology. This synergy between fine-tuned language models and semantic analysis allows for autonomous dialogue management and accurate domain-specific assistance, aligning with expert-level standards and significantly reducing human intervention. The approach offers a scalable, efficient solution for knowledge dissemination across specialized fields.

VIII. SOLUTION METHODOLOGIES

A. Design Methodology

The proposed domain-specific chatbot system is developed to enable efficient and precise academic query handling using Natural Language Processing (NLP), semantic retrieval, and transformer-based response generation. The design follows a modular approach, where each subsystem performs a critical function to achieve end-to-end intelligent interaction. The system's layered design ensures scalability and adaptability across various subject domains and educational contexts.

- User Interface (UI) Input: The system begins with the user submitting a subject-specific query through a conversational interface. This input dictates the processing flow for subsequent modules.
- Query Pre-processing (NLP Unit): The input query undergoes tokenization, normalization, and intent recognition. This step ensures that the chatbot interprets both linguistic structure and semantic meaning, enabling accurate understanding of user intent.
- Embedding Generation: The processed query is converted into a high-dimensional vector representation using transformer-based sentence embedding models. This representation captures semantic meaning beyond simple keyword matching.
- Semantic Retrieval (FAISS Vector Search): The query embedding is compared against pre-stored document embeddings (from textbooks, lecture notes, and research papers) using cosine similarity. This enables retrieval of the most contextually relevant academic content.
- Caching Mechanism: To enhance efficiency and maintain conversational continuity, recent queries and responses are stored in cache memory. This reduces redundant computation and allows the chatbot to recall prior interactions.
- Response Generation (Ollama Integration): The retrieved text segments are passed to the Ollama-powered NLP engine, which synthesizes a coherent, context-aware, and domain-appropriate response. This stage transforms raw retrieved data into human-like explanations.
- Dialogue Management: The chatbot maintains conversational flow by linking current responses with prior context. This ensures continuity across multiple user sessions and simulates tutor-like guidance.
- Output Display: The generated response is formatted and presented to the user through the chatbot interface. This stage integrates retrieval, generation, and caching modules to deliver a seamless academic interaction.
- Post-Interaction Review (Optional): An optional validation phase allows educators or users to review chatbot responses for accuracy and alignment with syllabus requirements. This can be performed manually or through automated evaluation metrics.

B. System Architecture Overview

The architectural design of the proposed NLP-based SigBot system is visually represented through the workflow diagram, which illustrates the sequential flow and interconnection between the major functional modules. Each block in the architecture corresponds to a specific stage in the chatbot pipeline, beginning from user input and culminating in intelligent, context-aware response generation. The process initiates when the user provides a query, which is processed through the NLP pre-processing module for linguistic analysis and intent recognition. The pre-processed query is then transformed into an embedding vector to capture its semantic meaning. Using cosine similarity, the system retrieves the most relevant document chunks from the vector database that align with the contextual intent of the query. The retrieved information is subsequently passed to the Ollama model, which forms the core NLP engine responsible for contextual understanding and natural language response formulation. The generated output is displayed to the user through the chatbot interface, ensuring real-time and coherent interaction. Additionally, all user interactions are stored in cache memory to enable context retention and faster responses in subsequent sessions. The sequential flow between these modules—from query processing and semantic retrieval to contextual generation and caching—ensures a structured, efficient, and adaptive operation. By integrating advanced language modelling, semantic search, and caching mechanisms, the proposed architecture achieves high accuracy, responsiveness, and scalability, making it a robust solution for intelligent academic assistance.

C. Libraries and Frameworks used

The SIGBOT system is developed in Python and incorporates several open- source libraries, each contributing to different components of the chatbot architecture.

Ollama framework – It is the primary execution environment for large language models. It allows the chatbot to load and run locally hosted models such as Llama 3, Mistral, or Phi 3, enabling fast and secure query generation without cloud dependency. The Ollama API provides a simple interface for sending queries and receiving contextually rich text responses.

Sentence Transformers – This library is based on pre-trained transformer models such as all-MiniLM-L6-v2, is employed to convert both the query and document text into numerical embeddings that preserve semantic meaning. These embeddings are stored for fast cosine similarity matching.

FAISS (Facebook AI Similarity Search) – This library provides high- speed vector search and indexing capabilities. It efficiently retrieves the most relevant text chunks corresponding to the input query.
The implementation of the SigBot system utilizes several Python libraries stated below.for PDF processing, libraries such as PyPDF, PyPDF2, and pdfplumber are used to extract textual content from academic PDF documents, while PyMuPDF (fitz) enables the extraction of both text and embedded images from more complex or scanned files. To handle scanned or image-based documents, pytesseract is employed for OCR, converting visual text into machine-readable form. Supporting this process, Pillow facilitates image manipulation, enhancement, and preprocessing to improve OCR accuracy, and pdf2image is used to convert PDF pages into image formats suitable for OCR operations. Additionally, general utility tools like pip, setuptools, and wheel are integrated for environment setup, package installation, and dependency management.

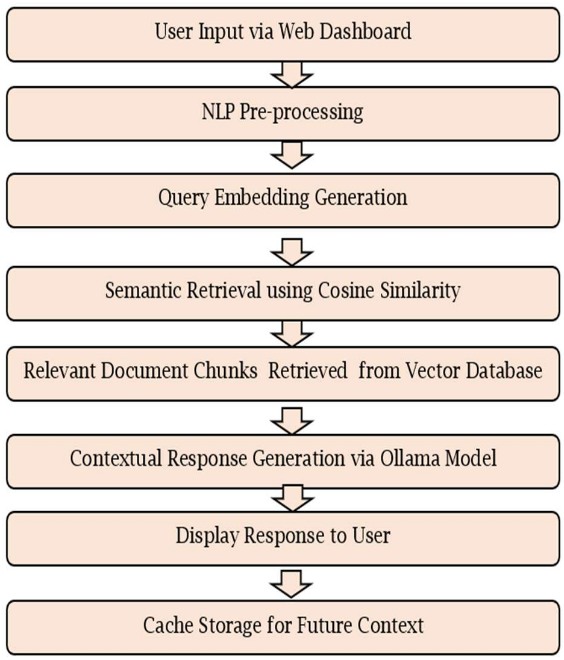

Figure. 1: Workflow diagram of SigBot Model

TABLE I: LIBRARIES FOR NLP AND MACHINE LEARNING BACKEND

NLP (Natural Language Processing)	Spacy	Tokenization, Part -of-Speech (POS) tagging, Named Entity Recognition (NER), NLP pre-processing
	Transformers	Pre- trained models (FLAN-T5, BERT, etc,) for Q&A, summarization, rephasing
	Sentence - transformers	Semantic embeddings for natural language search
Machine Learning Backend	Torch	Deep learning backend (required for transformers& embeddings)
	Scikit- learn	TF-IDF, cosine similarity, ML utilities
	FAISS	High – speed vector similarity search

TABLE II: LIBRARIES FOR PDF PROCESSING, OCR AND IMAGE PROCESSING AND UTILITIES SETUP

PDF Processing	Pypdf, PyPDF2, pdfplumber	Extract text from PDFs
	Pymupdf	Extract both text & images from PDFs
OCR & Image Processing	pytesseract	OCR for scanned PDFs images
	pillow	Image manipulation & preprocessing
	Pdf2image	Convert PDF pages to images (for OCR)
Utilities/ Setup	Pip, setuotools,wheel	Packet installation & management

D. Cosine Similarity

Following embedding, the system proceeds to the Document Retrieval Unit, which serves as the knowledge access core of SigBot. A preprocessed database of academic materials—including textbooks, lecture notes, and research papers—is stored in the form of embedding vectors. Each segment, or "chunk," of these documents has already undergone text extraction and embedding during preprocessing. When a new query vector is generated, the retrieval unit compares it with all stored document vectors using cosine similarity, a metric that measures the cosine of the angle between two vectors in multi-dimensional space. This similarity measure determines how closely related the meanings of the two texts are, regardless of their length or phrasing. A cosine similarity score close to 1 indicates a high semantic match. Based on these scores, the system retrieves the top-ranked text segments that are most contextually relevant to the user's question. This ensures that even when the query wording differs from the original text, SigBot still identifies conceptually aligned information.

$$\text{Cosine Similarity} = \frac{\boldsymbol{A} \cdot \boldsymbol{B}}{||\boldsymbol{A}|| \cdot ||\boldsymbol{B}||}$$

IX. RESULTS

It discusses the performance of various subsystems such as text preprocessing, embedding generation, semantic retrieval using cosine similarity, and intelligent response generation. The output results displayed through the command prompt demonstrate the successful integration and execution of the entire pipeline — from user query input to accurate and context-aware answer generation.

A. Command-Line Execution of SigBot as Pdf Reader

The Figure.2 shows the execution of the SigBot system, where the embedded Python script reads and processes an academic PDF file. The system operates through a command-line interface, enabling the chatbot to extract text content from uploaded PDF. In the proposed system, the existing challenges in digital learning platforms—such as lack of subject-specific clarity, generic AI responses, and absence of interactive conceptual guidance—are addressed through a domain-focused chatbot architecture. The system leverages Natural Language Processing (NLP) techniques and document retrieval to provide accurate, context-aware and aligned answers. By fine-tuning the model for a specific academic domain and integrating semantic understanding, the chatbot minimizes vague or inaccurate responses. This ensures that learners can interact with an intelligent tutoring system capable of evolving with user input, thereby enhancing conceptual understanding and engagement.

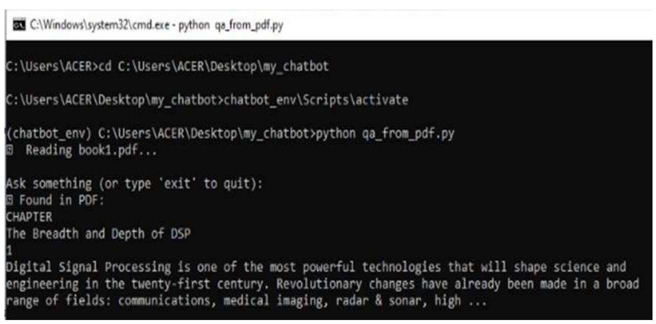

Figure. 2: Execution of the SIGBOT system

B. Keyword-Based Query Search from Pdf Document

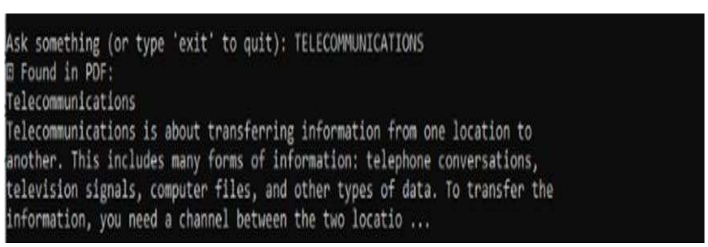

Figure. 3: Simulation of Keyword-Based Text Retrieval in SigBot

The above Figure.3 represents the working of SigBot's basic text retrieval mechanism. When the user enters a keyword such as the system scans the embedded PDF and locates the exact section containing that term. It then displays the extracted portion of text exactly as it appears in the original document. This demonstrates SigBot's ability to perform direct keyword-based searches, enabling users to quickly access relevant sections of academic materials without manually browsing through the entire file.

C. Result of Query-based Information Extraction in Sigbot

The below Figure. 4 shows the result of SigBot's query-based text retrieval process, where the system interprets a user's question and extracts the most relevant information from the embedded academic PDF. When the user asks the query, SigBot identifies the corresponding section from the source document and displays the retrieved explanation, demonstrating its ability to understand conceptual queries and return contextually accurate responses.

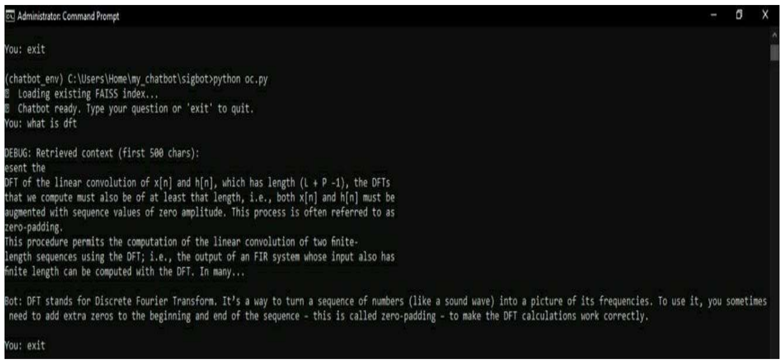

Figure. 4: Simulation of Query-based Information Extraction in SigBot

D. Result of Query-Based Image and Text Extraction in Sigbot

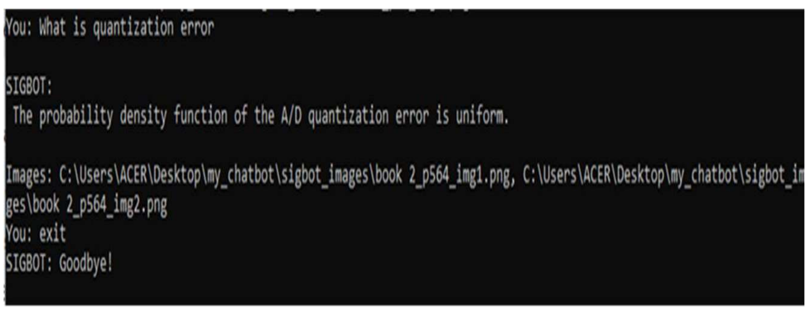

Figure. 5: Simulation of Query -Based Image and Text Extraction in SigBot

The above Figure.5 shows the simulation of the SigBot, where the system successfully responded to user queries by retrieving both textual and visual information. The figure illustrates an example interaction between the user and SigBot. When the user inputs the query the system processes it using the NLP and semantic retrieval pipeline and generates the output. The above output confirms that the chatbot not only interprets and answers the user's question accurately but also identifies the most relevant visual references extracted from the digital textbook database. The image file paths correspond to figures related to the concept of quantization error, retrieved automatically from preprocessed PDF materials.

E. Comparison of Cosine Similarity in Existing Vs. Proposed Models

In contrast, the proposed SigBot model not only retrieves text chunks with similar cosine similarity values but also ensures higher conceptual accuracy due to its enhanced semantic-processing pipeline. This improvement is enabled by advanced query embedding, refined text chunking, and contextual filtering supported by the Ollama-based RAG generation. As a result, SigBot significantly reduces false matches, even when similarity scores numerically fall within a similar range as the existing system. This comparison clearly demonstrates that cosine similarity alone cannot determine the correctness of retrieved answers. What truly matters is the combination of similarity scoring with intelligent LLM-driven interpretation. By integrating semantic retrieval with guided LLM generation, SigBot ensures that even moderately high similarity values lead to accurate, context-aware, and meaningful responses, unlike the existing system.

TABLE III. COMPARISON OF COSINE SIMILARITY SCORES

Question	Existing Model and Answer	SIGBOT Answer	Cosine Similarity
WHAT IS ALIASING	Aliasing in digital signal processing (DSP) occurs when a continuous-time signal is sampled at a rate that is too low to accurately capture its variations…	Aliasing occurs when a continuous-time signal is sampled at a rate too slow to accurately represent its frequency components….	71.90
WHAT IS DFT	The Discrete Fourier Transform is a mathematical operation that converts a discrete signal into its frequent components…	DFT stands for discrete Fourier Transform. It is a way to turn the sequence of number into a picture of its frequencies….	63.10

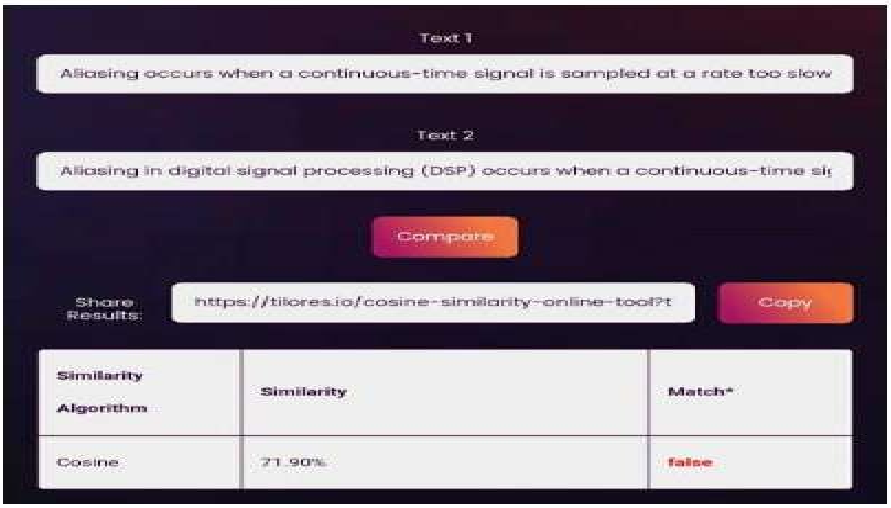

Figure. 6: Cosine Similarity Analysis: Existing Model vs. SIGBOT for Question No: 1

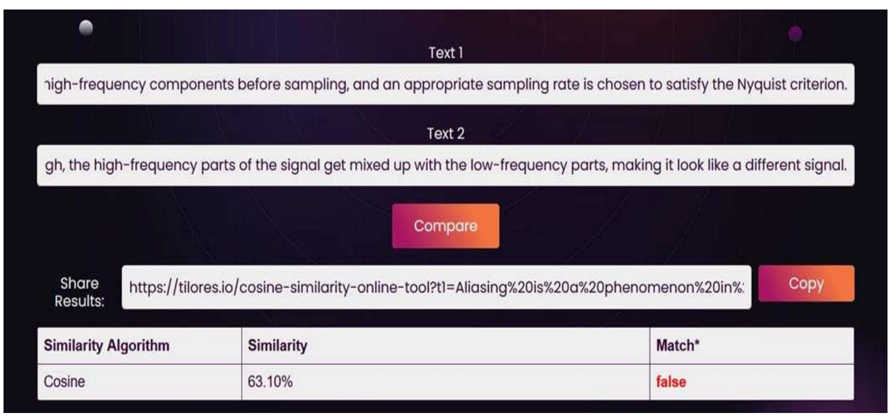

Figure. 7: Cosine Similarity Analysis: Existing Model vs. SIGBOT for Question No: 2

Figure. 6 and Figure.7 shows the cosine similarity comparison between the existing retrieval approach and the proposed SigBot model. The comparison shows the that cosine similarity scores which is moderately high (above 60%).

F. User Interface of Sigbot for Real-Time Query Processing

The below Figure.8 and 9 shows the deployed web-based dashboard of the SIGBOT system, where users can interact with the chatbot through a clean and intuitive interface. The dashboard displays essential components such as the chat window, chat history panel, control settings, and quick-action buttons, enabling seamless communication between the user and the system. When a user enters an academic query through this interface, the system processes the input using the integrated NLP preprocessing pipeline, semantic retrieval engine, and Ollama-based response generator to produce accurate and context-aware answers.

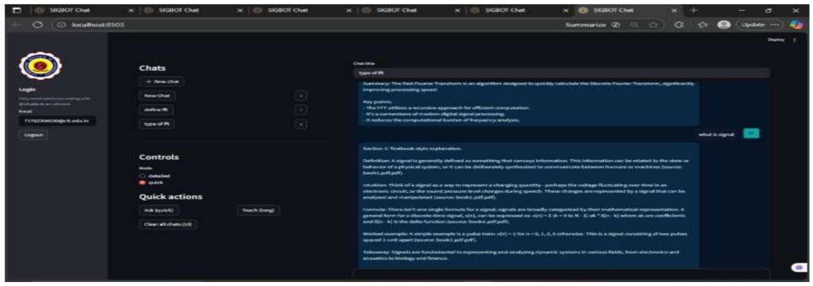

Figure. 8: Web-Based SIGBOT Interactive Dashboard

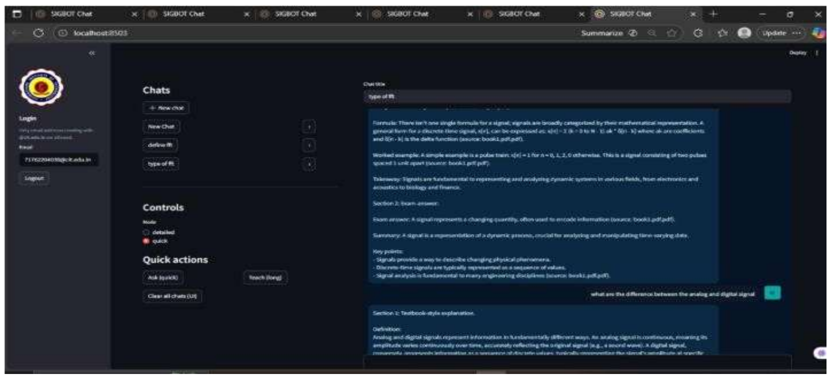

Figure. 9: Web-Based SIGBOT Interactive Dashboard

During the backend processing, the SigBot system leverages OCR-based preprocessing, embedding generation, FAISS indexing, and cosine similarity search to retrieve the most relevant academic content. OCR extraction, performed using Tesseract and PyMuPDF, allows the system to identify both textual information and embedded figures within the academic PDFs. Extracted information is indexed and stored so that the dashboard can display explanations enriched with accurate references to source content.

This integrated interface and retrieval framework enhances user engagement by combining semantic text understanding with a responsive, visually organized dashboard. The system not only provides clear and accurate answers but also supports interactive learning, making SigBot a reliable academic assistant for students and educators.

X. Conclusion

The development of the domain-specific chatbot, SigBot, demonstrates the effectiveness of combining Natural Language Processing, semantic retrieval, and locally deployed language models to provide accurate and context-aware academic support. By integrating document preprocessing, embedding generation, cosine-similarity-based retrieval, and intelligent response generation through the Ollama framework, SigBot successfully transforms raw academic material into an interactive knowledge system.The system consistently delivers relevant answers, retrieves both text and image content from PDFs, and adapts to user queries through semantic understanding rather than keyword dependence. This ensures higher accuracy, improved contextual interpretation, and reduced ambiguity compared to traditional rule-based or generic chatbots.Overall, the chatbot provides a reliable, efficient, and secure learning assistant capable of supporting subject-specific academic queries. The results validate that a domain-focused NLP architecture can greatly enhance learning experiences by offering real-time explanations, improved conceptual clarity, and interactive support aligned with the academic curriculum. Future enhancements such as GUI development, advanced rephrasing, and model fine-tuning will further strengthen SigBot's capability as an intelligent educational tool.

Future Work

Although the current implementation of SigBot successfully demonstrates accurate semantic retrieval, context-aware response generation, and efficient local deployment, several enhancements can further improve its performance, usability, and scalability. To enhance the naturalness, fluency, and pedagogical clarity of responses, advanced rephrasing models will be integrated. Domain-specific fine-tuning using academic datasets will further improve precision in technical interpretations and explanations. A speech-to-text and text-to-speech module will be integrated to enable hands-free interaction. This transforms SigBot into an accessible tool for visually impaired users and supports conversational learning through natural voice communication. Future versions will incorporate emotional and intent analysis to adjust the teaching style—providing hints, simplifying explanations, or offering advanced details depending on the student's learning needs.

ACKNOWLEDGMENT

This work was supported by Coimbatore Institute of Technology, Electronics and Communication Engineering Department. The authors would like to express their sincere gratitude to the Principal of Coimbatore Institute of Technology and Dr. A. Rajeswari, Head of the Department of Electronics and Communication Engineering, for their continuous support and encouragement throughout this research work.

REFERENCES

[1] B.R. Ranoliya, N. Raghuwanshi, and S. Singh, "Chatbot for university related FAQs," in Proc. IEEE Int. Conf. Big Data (Big Data), Boston, MA, USA, 2023, pp. 1525–1530.

[2] S. Mendoza, L. M. Sánchez-Adame, J. F. Urquiza-Yllescas, B. A. González-Beltrán, and D. Decouchant, "A Model to Develop Chatbots for Assisting the Teaching and Learning Process," Sensors, vol. 22, no. 15, p. 5532, 2022.

[3] I. A. Pap and S. Oniga, "eHealth Assistant AI Chatbot Using a Large Language Model to Provide Personalized Answers through Secure Decentralized Communication," Sensors, vol. 24, no. 18, p. 6140, 2024.

[4] Y. Dan et al., "EduChat: A Large-Scale Language Model-based Chatbot System for Intelligent Education," arXiv preprint arXiv:2308.02773, Aug. 2023.

[5] A. Chaidrata et al., "Intent Matching based Customer Services Chatbot with Natural Language Understanding," 2021 5th International Conference on Communication and Information Systems, ICCIS 2021, pp. 129–133, 2021

[6] Munira Ansari, Anupam Singh, et al., "Intelligent Chatbot", International Journal of Engineering Research & Technology (IJERT), Volume 9, Issue 4, Special Issue-2021, PP 79–82.

[7] Alturaiki A. M., Banjar H. R., Barefah A. S., Alnajjar S. A., Hindawi S. (2022). A smart Chatbot for interactive management in beta thalassemia patients. Int. J. Telemed. Appl. 2022.

[8] M. Ansari and A. Singh et al., "Intelligent chatbot," Int. J. Eng. Res. Technol. (IJERT), vol. 9, no. 4, Special Issue, pp. 79–82, 2021.

[9] M. D. L. Roca, M. M. Chan, A. Garcia-Cabot, J. Garcia-Lopez, and A. Salvatierra, "The impact of a chatbot working as an assistant in a course for supporting student learning and engagement," Comput. Appl. Eng. Educ., vol. 32, 2024.

[10] A. Sharma, R. Gupta, and S. Kumar, "AI-Driven Chatbot for Academic Assistance Using NLP,"in Proc. IEEE Int. Conf. Computing, Power and Communication Technologies (GUCON), 2023, pp. 980–985

[11] M. Al-Zubaidi and I. Aljarah, "A Smart University Chatbot Using Deep Learning-Based Intent Classification,"in IEEE Access, vol. 11, 2023, pp. 45010–45022.

[12] M. I. Hossain and F. M. Salahuddin, "Design of a Virtual Assistant for University Services Using Dialogflow and NLP," in Proc. IEEE Int. Conf. Sustainable Technologies for Industry 4.0 (STI), 2021, pp. 1–6

[13] A. Dey and S. Chatterjee, "Educational Chatbot for E-Learning Support Using Hybrid NLP Approach," in Proc. IEEE Int. Conf. Inventive Communication and Computational Technologies, 2021.

[14] R. Thomas and J. Mathew, "Multilingual FAQ Chatbot Using Transformer Models for Higher Education Institutions," in Proc. IEEE Int. Conf. Emerging Smart Computing & Informatics (ESCI), 2024.

[15] P. S. Raj et al., "Conversational Agent for Student Support Using Transformer-Based Language Models," in Proc. IEEE Int. Conf. Intelligent Systems and Green Technology (ICISGT), 2023.

[16] K. E. Lee and S. Park, "Intent Detection and Slot Filling for University Service Chatbots Using Joint BERT Models,"in Proc. IEEE ICMLA, 2022.

[17] J. Das and F. Wong, "Chatbots Development Using Natural Language Processing: A Review," in Proc. IEEE Int. Conf. on Computer Science and Information Technology (CSIT), 2023

[18] L. Chen and M. Kumar, "Emerging Trends in Chatbot Development: A Survey of Conversational AI Systems," in Proc. IEEE Int. Conf. on Computational Intelligence & Communication Technology (CICT), 2023

[19] H. Patel and K. Singh, "A Systematic Review of Recent Research on Chatbots," in Proc. IEEE Int. Conf. on Artificial Intelligence and Knowledge Engineering (AIKE), 2024

[20] S. Verma and R. Iyer, "Mental Health Chatbot for Addressing Psychological Well-being in India," in Proc. IEEE Int. Conf. on Healthcare Informatics (ICHI), 2023.