Visvesvaraya Technological University, Belagavi – 590018.

PROJECT SYNOPSIS
ON

A Retrieval-Augmented Generation Model for Academic and Administrative Workflows

Submitted in partial fulfillment for the award of degree of

BACHELOR OF ENGINEERING
in
ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

Submitted by

Abhijna	**4CB22AI002**
Akshata Krishna Hegde	**4CB22AI007**
Karishma B	**4CB22AI024**
Shriram Udaya Kumar Shenoy	**4CB22AI054**

Under the Guidance of

Mr. Deepak D
Assistant Professor, Department of AIML

DEPTARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

CANARA ENGINEERING COLLEGE

(Affiliated to VTU Belagavi, Recognized by AICTE)

Sudhindra Nagar, Benjanapadavu, Mangaluru - 574219, Karnataka.

A.Y 2024-25.

CANARA ENGINEERING COLLEGE

(Affiliated to VTU Belagavi, Recognized by AICTE)

Sudhindra Nagar, Benjanapadavu, Mangaluru - 574219, Karnataka.

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

CERTIFICATE

Certified that the Synopsis for project work entitled **"A Retrieval-Augmented Generation Model for Academic and Administrative Workflows"** is submitted by

Abhijna	**4CB22AI002**
Akshata Krishna Hegde	**4CB22AI007**
Karishma B	**4CB22AI024**
Shriram Udaya Kumar Shenoy	**4CB22AI054**

the bonafide students of VI semester Artificial Intelligence & Machine Learning in partial fulfillment for the award of Bachelor of Engineering in Artificial Intelligence & Machine Learning of the Visvesvaraya Technological University, Belagavi during the year 2024-2025. It is certified that all corrections/suggestions indicated for Internal Assessment as indicated during internal assessment. The Synopsis has been approved as it satisfies the academic requirements in respect of project work-Phase I prescribed for the said degree.

Mr. Deepak D	**Dr. Sujatha M**	**Prof. Basappa B Kodada**
Project Guide	Project-Coordinator -AIML	HOD

Abstract

Retrieval-Augmented Generation (RAG) technology enhances both academic support and administrative efficiency in educational institutions. By processing diverse academic materials such as lecture notes, presentations, and supplementary resources, these systems enable efficient data retrieval through a vector database. A fine-tuned Large Language Model (LLM) generates contextually relevant and personalized responses, improving student learning experiences. Additionally, by integrating student data, administrative workflows such as tracking academic progress and automating routine tasks are streamlined, reducing manual effort and enhancing institutional operations. Implementing AI-driven methodologies in educational environments optimizes learning outcomes and improves operational efficiency, offering a scalable framework for future advancements in academic and administrative management.

Keywords: Retrieval-Augmented Generation (RAG), Vector Database, Large Language Model (LLM), Academic Assistance, Administrative Automation, Artificial Intelligence in Education, Context-Aware Responses,Facebook AI Similarity Search(FAISS).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Managing vast academic resources and administrative tasks efficiently is a growing challenge in educational institutions. Retrieval-Augmented Generation (RAG) technology offers a powerful solution by enabling intelligent access to academic materials while automating various institutional processes. By integrating a fine-tuned Large Language Model (LLM) with a vector database, this approach ensures accurate and context-aware information retrieval, enhancing both learning experiences and operational efficiency. [1][2]

For students, this system provides personalized academic support, offering instant access to lecture notes, presentations, and supplementary materials. By generating contextually relevant responses, it fosters self-paced learning and deeper subject understanding. Unlike traditional systems, which rely on static content delivery, this AI-driven framework adapts dynamically to user queries, making learning more interactive and accessible. [3][4]

The role of an LLM-driven chatbot in higher education, particularly for databases and information systems, highlighting how AI chatbots can enhance student learning by providing personalized assistance, answering complex queries, and improving engagement. Similarly,investigated the integration of Retrieval-Augmented Generation (RAG) with LLMs, demonstrating how retrieval mechanisms improve the accuracy and reliability of AI-generated responses by incorporating relevant external knowledge. Their research emphasized the importance of factually and contextually appropriate answers, making AI-driven educational tools more interactive, accessible, and capable of providing real-time, high-quality information to students.[5][6][7]

Beyond academic support, the framework streamlines administrative operations by automating routine processes such as tracking student progress, managing institutional records, and handling course-related data. API integration with existing educational platforms ensures smooth functionality, reducing manual workload for administrators and enabling data-driven decision-making.[8][9][10]

A user-friendly web interface facilitates seamless interaction, allowing students, faculty, and administrators to access relevant information effortlessly. Designed for scala-

bility, this AI-powered system can be adapted to various academic environments, from universities to e-learning platforms, enhancing institutional efficiency and educational outcomes.

By combining advanced retrieval mechanisms with generative AI, this approach redefines how academic institutions manage knowledge and administrative workflows, fostering a smarter and more efficient learning ecosystem.

Chapter 2

Motivation and Problem statement

2.1 Problem Statement

Educational institutions struggle with efficient academic retrieval and administrative management. This project develops a RAG-based LLM with a vector database to deliver instant, context-aware responses, enhancing learning and institutional efficiency.

2.2 Objectives

1. To develop a system capable of providing users with contextually relevant responses based on academic materials.
2. To design an efficient storage and retrieval mechanism for educational resources, ensuring seamless access to information.
3. To RAG framework that enhances student learning by delivering personalized and context-aware academic support.

2.3 Scope and limitations

The scope of this system encompasses both academic and administrative functionalities within educational institutions. It leverages RAG to enhance student learning by providing contextually relevant, personalized responses based on academic resources such as lecture notes, presentations, and supplementary materials. Additionally, it streamlines administrative processes by automating routine tasks, tracking student progress, and managing institutional records. The system is designed to integrate seamlessly with existing educational platforms through API endpoints, ensuring scalability and adaptability to various academic environments, including universities and online learning platforms.

Limitations

The accuracy and relevance of generated responses depend on the quality and completeness of the academic materials stored in the vector database. Additionally, while automation improves administrative efficiency, it may require periodic human oversight to handle complex cases or ensure data integrity. The model's performance is also influenced by computational resources, which may affect response times and scalability in resource-constrained environments. Furthermore, maintaining data privacy and security remains a challenge, necessitating robust measures to prevent unauthorized access and ensure compliance with institutional policies.

Chapter 3

Proposed Methodology

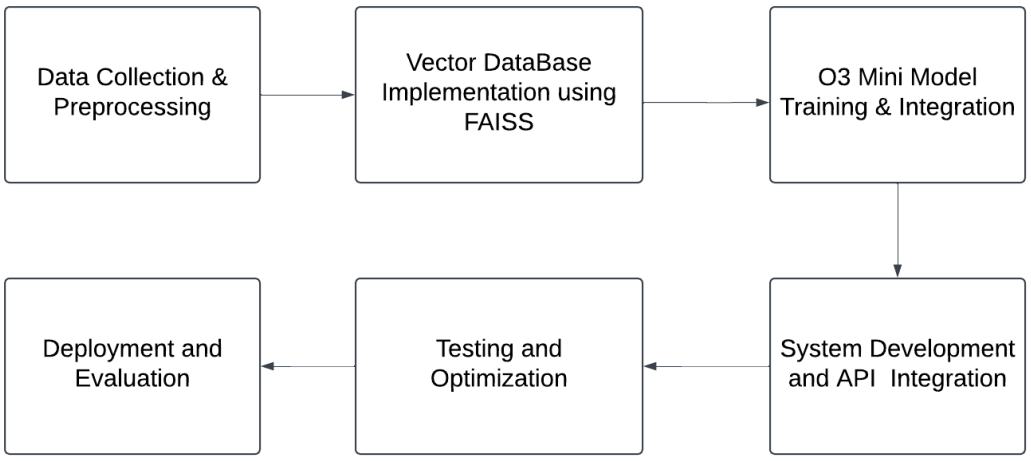

Figure 3.1: AI-Powered Academic RAG System Workflow

The Figure 3.1 outlines the AI-powered academic RAG system workflow with six key stages:

1. Data Collection & Preprocessing: All academic resources such as lecture notes, PPTs, PDFs, and textbooks are collected. Text is extracted, cleaned, and preprocessed through normalization. This process prepares the data for efficient embedding and retrieval.

2. Vector Database Implementation using FAISS: Each document chunk is converted into numerical vectors using embedding models. These vectors are stored in a FAISS-based vector database, which enables fast and accurate semantic search, where queries are matched based on meaning rather than just keywords.

3. o3 Mini Model Training and Integration: A lightweight and efficient LLM, the o3 Mini model, is fine-tuned or configured. This model is integrated with the retrieval system to generate context-aware answers based on the retrieved academic content.

4. System Development and API Integration: A backend system is developed using frameworks like Flask or FastAPI. APIs are implemented to connect the user interface with the database and language model, forming a seamless pipeline from user query to intelligent response.
5. Testing and Optimization: The system undergoes rigorous testing to ensure the accuracy of responses, the speed of retrieval and response generation, and the overall quality of the user experience. Based on the test results, optimizations are performed on the embeddings, model outputs, and system workflow.
6. Deployment and Evaluation: The final system is deployed on a cloud platform or an institutional server.Post-deployment, real-time performance metrics and user feedback are used for evaluation and continuous improvement of the system.

3.1 Relevance and Type

Educational institutions benefit from enhanced learning support and streamlined administration through AI-driven automation. Leveraging RAG, the system delivers personalized, context-aware responses, improving students' access to academic resources and enriching their learning experience.

Faculty and administrative staff also benefit from automated workflows, cutting down on time spent on repetitive tasks like answering common queries, tracking student progress, and managing institutional data. The AI-driven knowledge retrieval ensures quick and accurate access to information, helping both students and educators stay informed effortlessly.

Beyond its immediate impact, this project aligns with the growing trend of AI integration in education, showing how intelligent systems can enhance learning, simplify operations, and optimize institutional resources.

This is an AI-powered educational and administrative support system, designed to improve both academic assistance and institutional management. It leverages machine learning, NLP, and knowledge retrieval to provide efficient, scalable solutions.

The project functions as both a research-driven initiative exploring the use of RAG in education—and a practical tool that can be deployed in real-world settings. It is also interdisciplinary, combining computer science, educational technology, and administrative automation to create a smart, adaptable solution.

Chapter 4

Resource Requirements

1. Hardware
Computer/Server with
4-core CPU (minimum)
8GB RAM(minimum)
12GB free disk space
Optional GPU(for better performance)

2. Software
Model: o3 Mini
Frameworks and Libraries: Ollama 3 (for running LLMs), LangChain(for RAG-based applications) FAISS(for vector database and Similarity search)
Programming Language: Python

3. Materials
Training Data: Any text-based dataset used for retrieval-augmented generation (RAG)
Pre-trained Model Files: o3 Mini model weights
Embeddings and Vectorized Data: For FAISS database

4.1 Applications

Academic Assistance: The system provides personalized, context-aware responses to student queries based on academic materials, lecture notes, and institutional resources, enhancing the learning experience.
Faculty Support: Educators can use the system to quickly retrieve relevant course materials, generate summaries, and assist in research-related queries, improving productivity.
Administrative Automation: Routine tasks such as tracking student progress, man-

aging institutional data, generating reports, and responding to common inquiries can be automated, reducing workload for administrative staff.

Student Advisory and Guidance: By integrating with student support systems, the AI can assist in academic counseling, course selection, and career guidance, ensuring students receive timely and informed recommendations.

Exam and Assignment Assistance: The system can help students understand complex topics, provide explanations, and suggest study materials, acting as a virtual tutor.

Integration with Learning Management Systems (LMS): The AI can enhance existing LMS platforms by providing intelligent search capabilities, contextual recommendations, and automated assistance for coursework.

Institutional Help Desk: The AI-powered chatbot can handle common queries related to admissions, fees, schedules, policies, and other institutional services, improving response times and efficiency..

4.2 Budget

Table 4.1: Project Budget

No.	**Component**	**Cost (INR)**
1	Software	3000/-
2	Conference Paper Publishing	8000/-
3	Miscellaneous	2000/-
	Total	**13200/-**

Chapter 5

Time schedule

Table 5.1: Project Timeline Schedule

Sl. No.	**Duration**	**Activity/Task to be Completed**
1	11-2-2025 to 23-2-2025	Finalization of Project Titles
2	24-2-2025 to 6-3-2025	Literature Survey
3	7-3-2025 to 9-3-2025	Framing Problem Statement
4	10-3-2025 to 12-3-2025	Framing Objectives
5	14-3-2025 to 16-3-2025	Draft Methodology
6	17-3-2025 to 20-3-2025	Writing Synopsis
7	20-3-2025	Submission of Draft Synopsis
8	24-3-2025 to 3-4-2025	Amendment of Synopsis
9	2-4-2025	Synopsis Submission to Guide
10	8-4-2025	Synopsis Submission to Coordinator
11	8-4-2025 to 15-4-2025	Gathering Requirements
12	16-4-2025 to 24-4-2025	Application-based: Gather Requirements
13	25-4-2025 to 9-5-2025	Project Phase-I Report Writing
14	6-5-2025	Submission of Draft Report to Guide
15	7-5-2025 to 13-5-2025	PPT Preparation or Amendment of Report
16	14-5-2025	Submission of Draft Report to Coordinator

References

[1] Y. Chang and H. Kim. "Enhancing Student Performance Prediction with LLMs". In: *IEEE Education Forum*. 2024.

[2] D. Garcia and X. Lopez. "Automating Administrative Processes Using AI Chatbots". In: *Smart Education Journal* 7.1 (2024), pp. 30–50.

[3] M. Gomez and R. Rivera. "Bridging the Digital Divide with AI-Powered Chatbots". In: *Tech in Education* 6.2 (2024), pp. 112–130.

[4] P. Joshi et al. "Robust Multi-Model RAG Pipeline for Documents Containing Text, Table Images". In: *IEEE ICAAI Conference*. 2024. DOI: 10.1109/ICAAIC60222.2024.10574972.

[5] P. Kumar and M. Sharma. "A Survey on RAG-Based AI Systems in Higher Education". In: *International Journal of AI in Education* 9.2 (2024), pp. 88–105.

[6] H. C. Lee et al. "Development of an RAG-Based LLM Chatbot for Enhancing Technical Support Service". In: *IEEE TENCON Conference*. 2024. DOI: 10.1109/TENCON61640.2024.10902801.

[7] A. T. Neumann et al. "An LLM-Driven Chatbot in Higher Education for Databases and Information Systems". In: *IEEE Transactions on Education* 68.1 (2025), pp. 103–120. DOI: 10.1109/TE.2024.3467912.

[8] S. Patel and A. Singh. "Improving Online Learning Through AI Recommendations". In: *International Conference on AI for Learning*. 2024.

[9] J. Smith and L. Brown. "Exploring LLMs for Student Assistance". In: *AI in Education Journal* 12.3 (2024), pp. 45–60.

[10] T. Wang and R. Li. "Using AI to Improve University Record-Keeping". In: *AI in Academia Conference*. 2024.