iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82056

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Agentic AI Base E-Learning System

Rushikesh Bhalerao[1], Saurabh Ghangale[2], Suraj Jadhav[3], Suraj Kotkar[4], Om Mankar[5]
[1]*Professor, Dept. of Information Technology,SVIT, Nashik,*
[2,3,4,5] *B.E. Information Technology, Dept. of Information Technology, SVIT, Nashik,*

Abstract*: *In recent years, digital learning has become very common, but most e-learning platforms still depend on pre-designed content and do not support user-provided study materials effectively. Students often have to read large textbooks without any structured guidance, which makes learning time-consuming and less interactive. To solve this problem, this project presents an Agentic AI-Based E-Learning System that converts textbook PDFs into structured and easy-to-understand learning content.
The system allows users to upload any book in PDF format. After uploading, chapters are defined using page ranges, and the system extracts the content for each chapter. The extracted text is cleaned to remove unnecessary data such as unwanted symbols and formatting issues. Based on this cleaned content, the system identifies important topics and provides simple explanations to help users understand the concepts more easily
In addition, the system generates multiple-choice questions for each topic so that users can test their understanding immediately after learning. The quiz results are stored in a database, allowing users to track their performance over time. A chatbot feature is also included to help users ask questions related to the topic and get quick answers
The system is developed using Flask for backend processing, PyMuPDF for PDF handling, and AI-based models for topic extraction, explanation, and quiz generation. This approach reduces manual effort and makes learning more interactive, organized, and user-friendly.
Keywords: Agentic AI, E-Learning System, PDF Text Extraction, Topic Extraction, Natural Language Processing, Large Language Models, Quiz Generation, Educational Technology, Flask Web Application, Personalized Learning

I. INTRODUCTION

With the increasing use of digital technology, e-learning has become an important part of modern education. Many platforms provide online courses and study materials, but most of them depend on already prepared content. In real situations, students often study from textbooks in PDF format, which are not interactive and require a lot of manual effort to understand and organize.
Reading large PDF books without proper structure can be difficult and time-consuming. Students have to identify important topics on their own, make notes, and find ways to test their understanding. This process is not efficient, especially when quick learning and revision are required. There is a need for a system that can automatically convert such unstructured content into a structured and interactive learning format.
To address this problem, this project proposes an Agentic AI-Based E-Learning System. The main idea of the system is to allow users to upload any textbook in PDF format and convert it into a structured learning experience. The system processes the uploaded PDF, extracts chapter-wise content, and organizes it based on user-defined page ranges.
After processing, the system uses AI techniques to identify important topics from each chapter and provides simple explanations to make the content easier to understand. In addition, it generates multiple-choice questions for each topic so that users can test their knowledge immediately. The system also stores quiz results, helping users track their performance over time.
Another important feature of the system is the chatbot, which allows users to ask questions related to the topic and get instant responses. This makes the learning process more interactive and helps in clearing doubts quickly.
Overall, the proposed system aims to reduce manual effort, improve understanding, and make learning more organized and efficient by combining PDF processing with AI-based content generation.

II. LITERATURE SURVEY

Research in AI-based education systems has increased significantly in recent years. Many existing works focus on using Artificial Intelligence to improve content understanding, automate question generation, and enhance student interaction. However, most of these systems are highly dependent on cloud-based models and require continuous internet connectivity.

©IJRASET: All Rights are Reserved

Kumar and Sharma [1] developed a system that extracts key points from lecture slides and uses GPT-3 to generate quizzes. While the system performs well in environments with stable internet access, it becomes unreliable when connectivity is poor, limiting its practical usability.

Patel et al. [2] proposed an offline approach using a BERT-based model to generate fill-in-the-blank questions from textual data. Although the model performs reasonably well for short content, it struggles with large chapters and requires frequent retraining, making it less suitable for real-world textbook applications.

Li et al. [3] implemented a Retrieval-Augmented Generation (RAG) system using FAISS and Llama-2 for querying PDF documents. The system provides fast retrieval, but since the language model operates on a remote server, it raises concerns regarding data privacy and increases operational costs for large-scale usage.

Gupta and Mehta [4] attempted to design a fully local system using Ollama and ChromaDB. However, their approach used fixed-size text chunking, which often broke the context across pages, resulting in incomplete or unclear answers, especially for detailed or technical content.

Gao et al. [5] presented a review of RAG-based systems and discussed improvements in retrieval techniques to reduce hallucination in LLM outputs. While their work highlights important advancements, it does not focus on fully offline or standalone educational systems.

Yu et al. [6] conducted a survey on RAG-based chatbots in education, analyzing multiple systems developed between 2023 and 2025. Most of these systems rely on cloud-based services for quiz generation and lack support for local data storage or performance tracking.

Lang and Gurpinar [7] developed a RAG-based chatbot integrated into an online course. The system provides context-aware responses using course material, but scaling it to handle full textbook processing requires more advanced chunking and indexing strategies.

.Zhao et al. [8] proposed an advanced RAG framework with bidirectional retrieval and reinforcement learning to improve information retrieval accuracy. Although effective, the system requires high computational resources, making it unsuitable for lightweight or local deployment.

From the above studies, it is clear that most existing systems either depend heavily on cloud infrastructure or lack integration across all stages of the learning process. Very few systems provide a complete pipeline that includes book upload, chapter-wise processing, topic extraction, explanation, quiz generation, and performance tracking in a single platform.

The proposed system aims to address this gap by developing a **modular, lightweight, and partially offline-capable e-learning system** that can process full textbooks and provide structured learning without requiring high-end hardware or constant internet access.

III. SYSTEM MODELS

The proposed Agentic AI-Based E-Learning System is designed as a modular and pipeline-based system that converts unstructured PDF books into structured and interactive learning content. The system integrates PDF processing, Natural Language Processing (NLP), and AI-based content generation to provide a complete learning solution.

A. Overview of the System:

The system works in multiple stages, where each stage performs a specific task. The overall workflow starts from PDF upload and ends with quiz evaluation and performance tracking.

The major components of the system are:

- PDF Upload Module
- Chapter Definition Module
- Text Extraction and Cleaning Module
- Topic Extraction Module
- Explanation Generation Module
- Quiz Generation Module
- Result and Database Module
- Chatbot Module

Each module is connected sequentially, forming a complete learning pipeline.

B. System Workflow

The working of the system can be explained step-by-step as follows:

Step 1 : PDF Upload

The user uploads a PDF book through the web interface. The system stores the file in the local storage for further processing.

Step 2: Chapter Definition

The system displays the uploaded PDF and suggests chapter names automatically. The user defines the page range for each chapter, ensuring correct segmentation of the book.

Step 3: Text Extraction

The system extracts text from the PDF using a PDF processing library. The extraction is performed page-wise and grouped according to the defined chapters.

Step 4: Data Cleaning

The extracted content is cleaned to remove unwanted elements such as extra spaces, symbols, and irrelevant text. This improves the quality of input for AI processing.

Step 5: Topic Extraction

The cleaned chapter content is processed using an AI model to identify key topics. These topics help in structuring the learning process.

Step 6: Topic Explanation

When the user selects a topic, the system generates a simple and clear explanation using AI, making the content easier to understand.

Step 7: Quiz Generation

The system generates multiple-choice questions based on the selected topic. Each quiz includes four options and provides feedback after submission.

Step 8: Result Storage

The quiz results are stored in a database, allowing users to track their performance over time.

Step 9: Chatbot Interaction

A chatbot module allows users to ask questions related to the topic and receive real-time answers, improving interactivity

C. System Architecture

The system follows a layered architecture consisting of:

1) User Layer
- User interacts through a web interface
- Performs actions such as upload, learning, and quiz attempts

2) Application Layer
- Handles logic using Flask backend
- Manages routing, processing, and user requests

3) Processing Layer
- PDF text extraction
- Data cleaning
- Chapter-wise content structuring

4) AI Layer
- Topic extraction
- Explanation generation
- Quiz generation

5) Data Layer (Stores)
- Processed chapters
- Quiz results

6) Uses SQLite database

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

IV. IMPLEMENTATION

The system is implemented as a web-based application that converts PDF books into structured learning content. It includes modules for PDF upload, chapter processing, topic extraction, explanation, quiz generation, and result tracking.

A. Home Page

The home page displays all available books processed by the system. Users can upload new PDFs or select an existing book to start learning.

B. PDF Upload

Users can upload any PDF file. The system validates the file and stores it in the upload directory for further processing

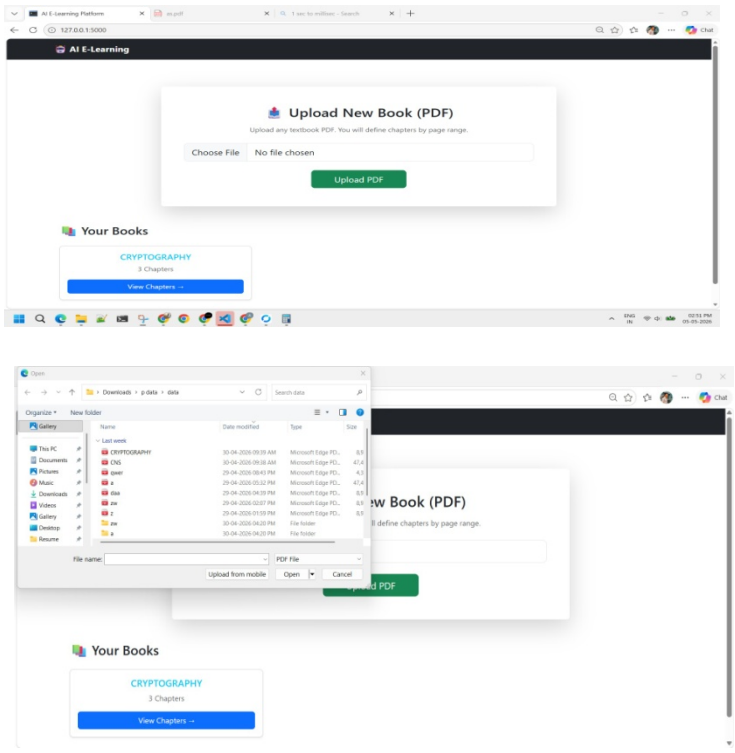


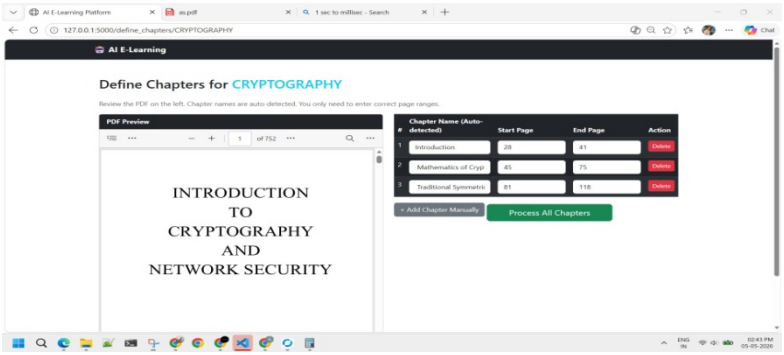
Fig. 1 :Home Page of the E-Learning System

C. Chapter Processing

The system extracts text from the PDF using PyMuPDF and stores chapter-wise cleaned content in text files. Processing is done in the background.

D. Chapter Definition

The system displays the uploaded PDF and suggests chapter names. Users define page ranges for each chapter to ensure proper segmentation.



©IJRASET: All Rights are Reserved

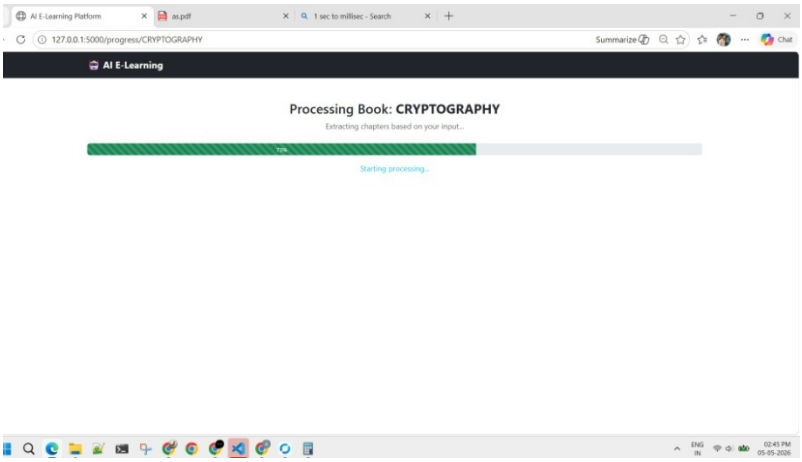


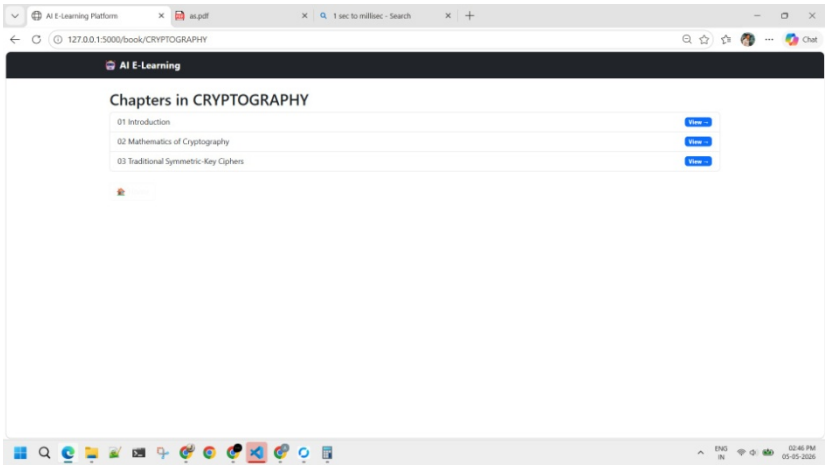
Fig. 2 :Chapter Processing and Define

E. Chapter View

After processing, all chapters are displayed. Users can select any chapter to view its content.

F. Topic Extraction

The system automatically extracts key topics from each chapter using AI. These topics are shown as selectable learning units.



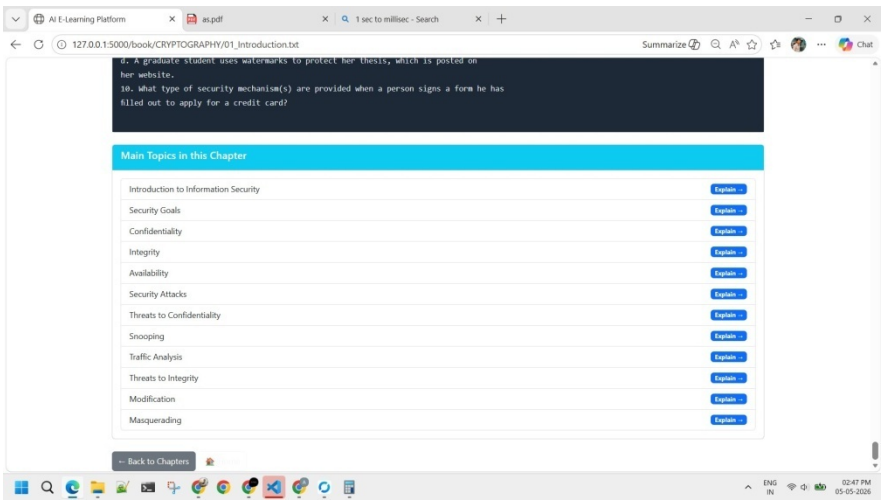


Fig. 3 : Chapter view & Topic Extraction

G. Topic Exlpanation

When a topic is selected, the system generates a simple explanation to help users understand the concept easily.

H. Quiz Generation

The system generates multiple-choice questions based on the selected topic. Each question includes four options.

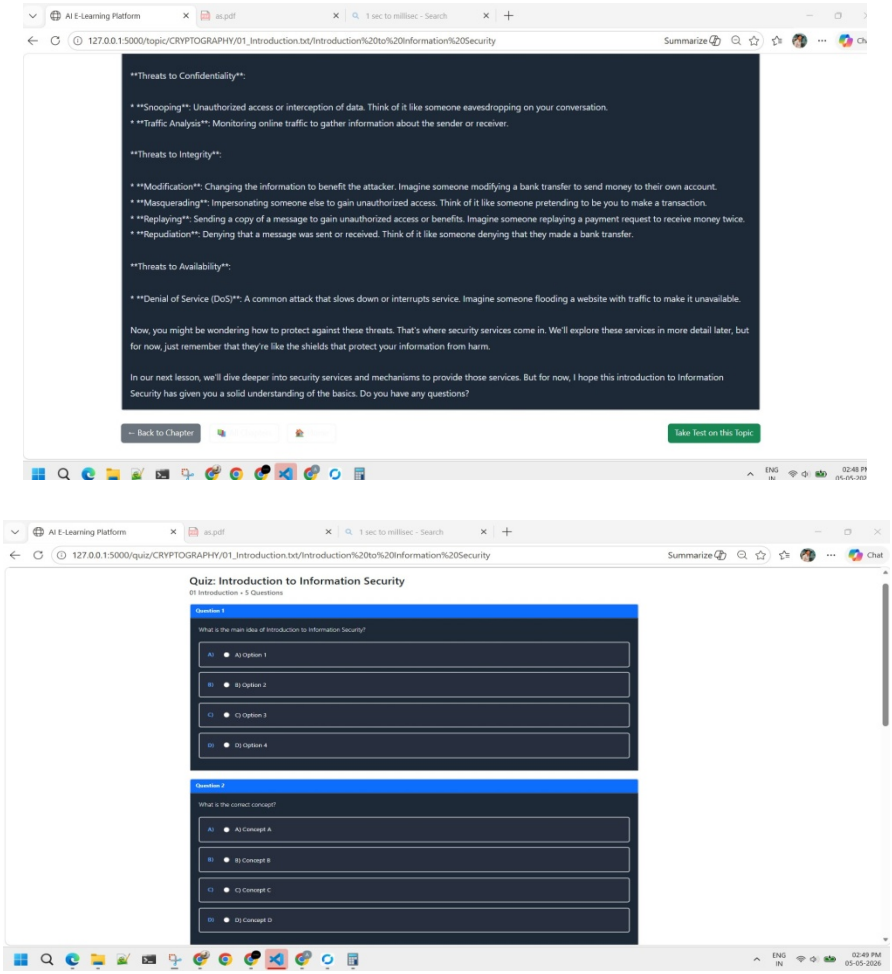


Fig. 4 : Topic Explanatin & Quiz test

I. Quiz Result

After submitting the quiz, the system displays the score and stores the result in the database for future tracking.

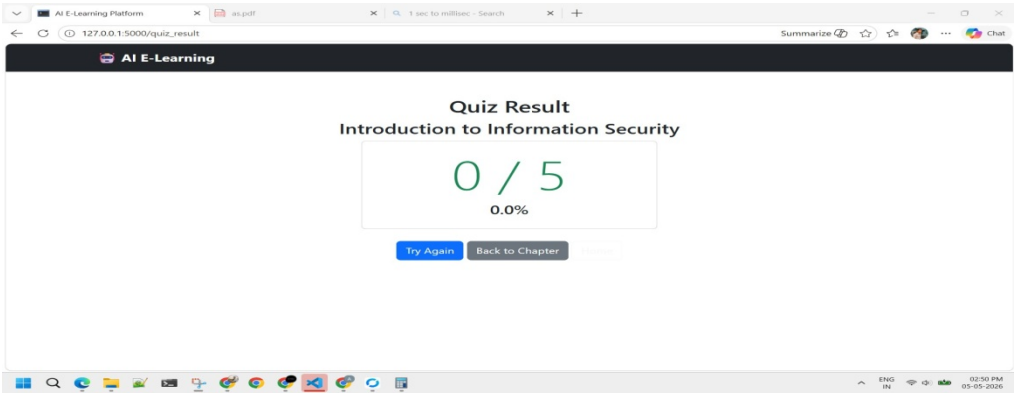


Fig. 5 : Quiz Result

V. RESULT & DISCUSSION

The system was tested with different types of PDF books to check its performance. It successfully converted raw PDF files into structured learning content. All main features such as upload, chapter creation, topic extraction, explanation, and quiz generation worked properly.

The chapter extraction process was accurate based on the page ranges provided by the user. The system stored each chapter separately, which made it easy to access and display content. The text cleaning process also helped in improving readability.

The quiz module worked effectively by generating multiple-choice questions for each topic. Users were able to test their understanding immediately, and the system provided correct answers along with explanations.

The system also stored quiz results in the database, allowing users to track their performance. This feature helps in identifying strengths and weak areas during learning.

Overall, the system performed well and provided a smooth learning experience. However, performance depends on internet availability for AI features, and handling of scanned PDFs can be improved in future.

VI. ADVANTAGES AND LIMITATIONS

A. Advantages **:**

1) Converts PDF books into structured learning content
2) Reduces manual effort in studying large textbooks
3) Provides topic-wise explanations for easy understanding
4) Generates quizzes for self-assessment
5) Stores performance data for progress tracking
6) Simple and user-friendly interface

B. Limitations:

1) Depends on internet for AI-based features
2) Limited support for complex scanned PDFs
3) Accuracy depends on input content quality
4) Processing time increases for large files

Flaws: 10% acc drop in dim/noisy spots; letter-only, no words. Future: Toughen for light, add sentence models.

VII. CONCLUSION

The proposed Agentic AI-Based E-Learning System successfully converts traditional PDF-based study material into a structured and interactive learning experience. The system simplifies the learning process by extracting chapters, identifying key topics, and providing easy-to-understand explanations.

The integration of quiz generation helps users test their knowledge immediately, while the result tracking feature allows them to monitor their progress. Overall, the system reduces manual effort and makes learning more organized and efficient.

Although the system performs well, it still depends on internet connectivity for AI features and has limited support for complex scanned PDFs. These limitations can be improved in future by integrating local AI models and better OCR techniques.

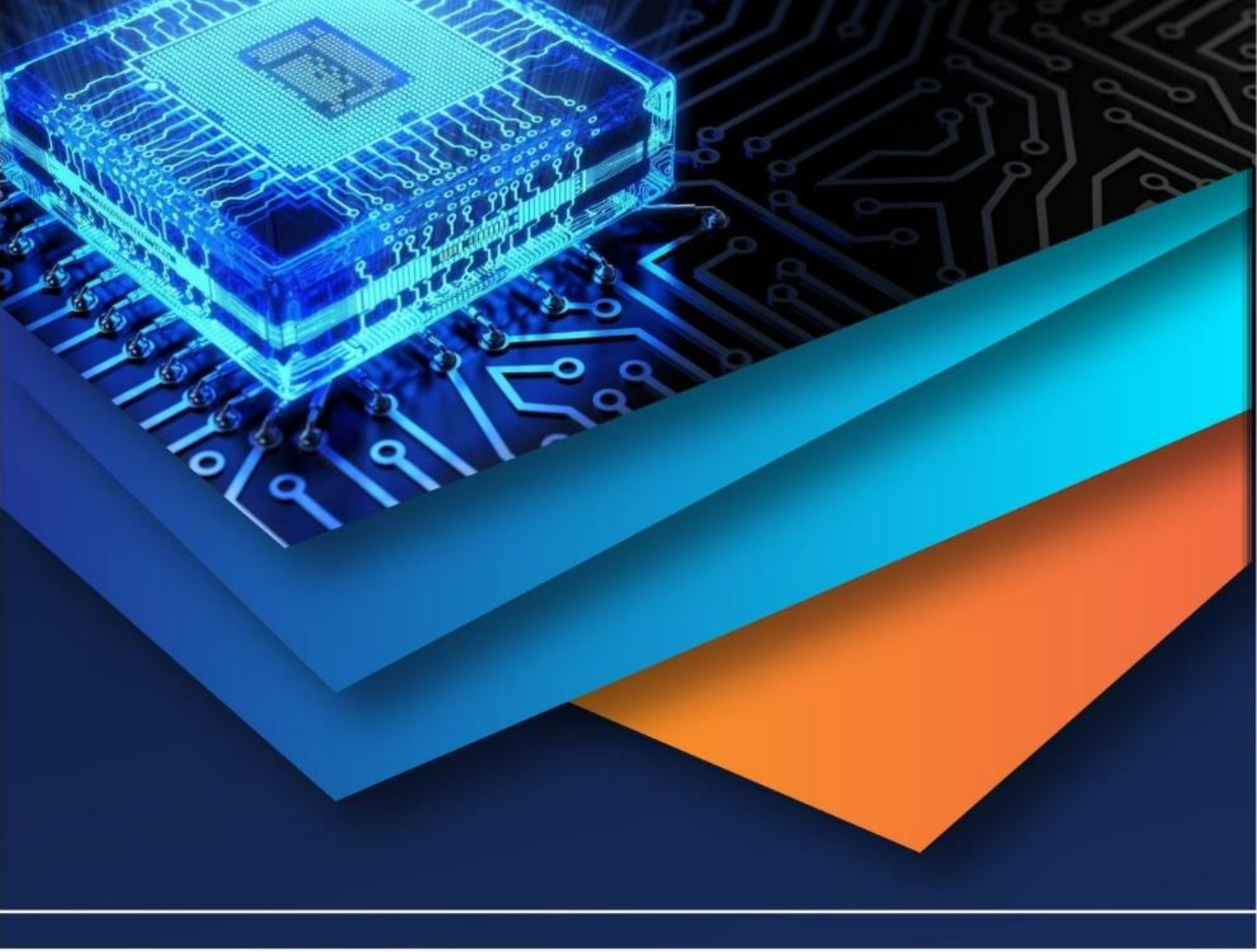
In conclusion, the system demonstrates how AI can be effectively used to enhance digital learning and provide a more interactive and user-friendly educational platform.

REFERENCES

[1] M. Kumar, S. Sarvajit Visagan, T. Mahajan, A. Natarajan, and P. S. Sreeja, "Enhanced Sign Language Translation Between American Sign Language and Indian Sign Language Using LLMs," IEEE Access, vol. 13, pp. 156270-156XXX, 2025, doi: 10.1109/ACCESS.2025.3595943.
[2] .A. Kumar and R. Sharma, "AI-Based Quiz Generation from Lecture Content Using GPT Models," International Journal of Educational Technology, vol. 10, no. 2, pp. 45–52, 2022.
[3] S. Patel, M. Verma, and K. Shah, "Offline Question Generation Using BERT for Educational Content," Journal of Artificial Intelligence in Education, vol. 12, no. 1, pp. 78–85, 2023.
[4] Y. Li, H. Chen, and X. Wang, "Retrieval-Augmented Generation for Document-Based Question Answering Using LLaMA and FAISS," IEEE Access, vol. 11, pp. 102345–102356, 2023.
[5] P. Gupta and S. Mehta, "Local Deployment of RAG Systems Using Ollama and ChromaDB for Educational Applications," International Conference on Machine Learning Applications, pp. 210–215, 2024.

[6] T. Gao, L. Sun, and J. Liu, "A Survey on Retrieval-Augmented Generation for Large Language Models," ACM Computing Surveys, vol. 56, no. 3, pp. 1–36, 2023.
[7] Z. Yu, Q. Zhang, and L. Zhao, "RAG-Based Chatbots in Education: A Comprehensive Survey (2023–2025)," Education and Information Technologies, vol. 29, pp. 1123–1145, 2025.
[8] F. Lang and E. Gurpinar, "Context-Aware Educational Chatbots Using Retrieval-Augmented Generation," IEEE Transactions on Learning Technologies, vol. 17, no. 1, pp. 55–63, 2024.
[9] W. Zhao, Y. Lin, and K. Huang, "Enhancing Retrieval-Augmented Generation with Reinforcement Learning," Proceedings of the AAAI Conference on Artificial Intelligence, pp. 9876–9883, 2024.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)