iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81205

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 IV Apr 2026- Available at www.ijraset.com

Study Sync - Group Study Platform with RAG based Assistance

Kongari Bhuvana[1], Mullapati Kushitha[2], Dr. Shaik Irfan Babu[3], Ms. S. RajiniParimala[4]
[1, 2]*Students,* [4, 5]*Assistant Professor, Department of Emerging Technologies (CSE-AI&ML), Mahatma Gandhi Institute of Technology, Hyderabad, Telangana – 500 075*

Abstract: ***Students today juggle multiple challenges in their academic journey—scattered study materials, lack of personalized learning guidance, difficulty retaining information from vast course content, and the struggle to maintain focus during group study sessions. StudySync: AI-Enhanced Collaborative Group Study Platform addresses these real-world challenges by integrating retrieval-augmented generation (RAG), real-time data processing, and collaborative intelligence. The platform lets students upload their study materials—PDFs, PowerPoint presentations, notes—and intelligently retrieves the most relevant content to generate personalized quizzes, concept maps, and learning summaries. With its live group chat feature, role-based access, and multi-format document support, StudySync makes studying together feel natural and productive, even when students are miles apart. Whether you're preparing for exams or working through complex topics as a group, the platform learns from your materials and adapts to what you actually need to understand.***

The magic of StudySync lies in its data-driven methodology that combines RAG with machine learning to turn raw study materials into intelligent learning tools. When you upload documents, the system extracts and chunks the content using specialized processors—textract for presentations, pdf-parse for PDFs—creating a searchable knowledge base. Your study materials are then indexed and made available through our RAG pipeline, which retrieves contextually relevant passages when generating quiz questions and concept visualizations. Mistral AI's language models process this retrieved data to craft personalized questions that focus on what you actually studied, not generic trivia. PostgreSQL stores your group's learning history, while Socket.IO ensures real-time synchronization across all study members. This end-to-end pipeline transforms how groups learn together—producing accurate, material-specific insights that stick, enabling deeper understanding, and making collaborative studying feel less like a chore and more like having a smart study buddy who really knows your coursework.

Keywords: Retrieval-AugmentedGeneration(RAG),Artificial Intelligence (AI), Natural Language Processing (NLP), Collaborative Learning Platform, AI-Powered Educational Technology, Real-Time Knowledge Retrieval, Intelligent Quiz Generation

I. INTRODUCTION

Student collaboration faces a critical challenge: study materials remain scattered across devices, cloud drives, and notes, making it difficult to access relevant information during group learning sessions. Traditional study platforms lack intelligent retrieval mechanisms and rely on generic AI responses disconnected from actual course content, forcing students to manually search through documents or repeat information. StudySync addresses this by introducing an AI-powered collaborative study platform that combines Retrieval-Augmented Generation (RAG) with real-time group collaboration. By intelligently indexing and retrieving from student-uploaded materials—PDFs, PowerPoint presentations, notes—the platform generates contextually accurate quizzes, concept maps, and learning summaries. This ensures every response is grounded in the student's actual coursework, not generic information, transforming how groups learn together effectively and efficiently.

A. Motivation

The motivation behind StudySync is to make intelligent, personalized learning accessible to collaborative study groups by leveraging their own study materials through RAG technology. Students spend hours searching through scattered documents and struggle to synthesize information from multiple sources during group sessions. StudySync solves this by creating a unified knowledge base from uploaded materials and using RAG to retrieve relevant content, allowing the AI to generate quiz questions and concept maps that directly align with what students are actually studying. This empowers students to make faster progress, retain information better, and study smarter by having an AI assistant that truly understands their coursework through intelligent data retrieval and augmentation.

©IJRASET: All Rights are Reserved

B. ProblemDefinition

The main problem faced by students is the fragmentation of study materials and lack of intelligent content retrieval during group learning. Most existing study platforms offer basic collaboration features but lack RAG capabilities—they cannot intelligently retrieve and process student-uploaded materials to generate context-aware learning aids. Text-based search is inefficient, AI responses are often generic and disconnected from actual coursework, and students waste time manually finding relevant content. StudySync solves this by implementing a comprehensive RAG pipeline that indexes all uploaded documents (PDFs, presentations, notes), retrieves contextually relevant passages, and augments AI responses with this retrieved data to deliver accurate, personalized, and material-specific guidance.

C. ExistingSystem

Existing collaborative study platforms provide basic group chat and document sharing features but suffer from significant limitations. Most lack intelligent content processing—documents are simply stored without being indexed or made searchable through RAG. There is no retrieval mechanism to connect quiz generation or concept mapping to actual student materials; AI responses remain generic and disconnected from coursework. These systems don't extract knowledge from uploaded documents, offer no way to retrieve specific information for learning purposes, and fail to augment AI responses with student-specific data. Students must manually navigate through folders and re-read materials to find relevant information, defeating the purpose of intelligent collaborative learning.

D. ProposedSystem

StudySync is an AI-powered collaborative study platform designed to overcome limitations through advanced RAG technology integrated with group collaboration. The system implements a multi-stage RAG pipeline: when students upload educational materials (PDFs,notes), the platform automatically extracts text using specialized processors (pdf-parse for PDFs). This content is chunked, embedded, and indexed into a searchable knowledge base stored in PostgreSQL. When students request quizzes or concept maps, the RAG system retrieves the most relevant passages from the knowledge base, which Mistral AI then processes to generate contextually accurate learning materials grounded in actual coursework. Real-time Socket.IO communication enables seamless collaboration, while the RAG backend ensures every AI-generated response is augmented with retrieved data from student materials. Role-based permissions and secure JWT authentication protect collaborative integrity. This end-to-end RAG-enhanced system ensures students receive intelligent, material-specific guidance that directly supports their actual learning needs.

II. LITERATURE SURVEY

Research on intelligent collaborative learning systems is rapidly expanding with a focus on improving student comprehension and academic performance. Earlier approaches mainly provided basic document sharing and generic communication, but recent advancements in Retrieval-Augmented Generation, large language models, real-time collaboration technologies, and semantic vector indexing have enabled more accurate, material-grounded learning support and intelligent educational decision-making for study groups.

TableIhighlightsselectedliteratureonthis topic.

S. No	Year	Author	Title	Methodology
1	2024	D.Oreškiet al.	Retrieval Augmented Generationin Large Language Models	Integration of RAG withLLMsforchatbots at university level; use case&feedback-driven improvements
2	2025	Saad-Falcon et al.	ASurveyonKnowledge-Oriented Retrieval-Augmented Generation	Comprehensivereview & taxonomy of RAG for NLP, including educational use

3	2024	Jacobs & Jaschkeetal.	Designinga student-friendly RAG-based chatbot	Personalized,domain- specific assistant, providing QA for students
4	2024	Velazquezet al.	Enabling Educators to Build Specialized AIChatBots with RAG	RAG-based chatbot helping educators manageinstructional content

S. No	Year	Author	Title	Methodology
5	2024	Digvijay Singhetal.	AI-enabled Virtual Collaborative Learning Classroom	AnalyticalreviewofAI interventions in online classrooms
6	2025	SindhuMet al.	Enhancing student supportwith a RAG owered chatbot	ImplementationofRAG fortextbookQ&Atasks for students
7	2025	AttilaKovari et al.	A systematic reviewofAI- powered collaborative learning	Review of current status, impact of AI collaborativetoolsin higher education
8	2024	EdTech Startups (Disco, 360Learning, Kahoot! AI)	AI-Powered Collaborative Learning Platforms	AIusesanalytics,RAG, peer-driven content sharing, context suggestions,discussions

III. METHODOLOGY

A. Dataset Collection

StudySync collects learning data from multiple sources including student-uploaded documents (PDFs, PowerPoint presentations, text notes), real-time chat messages, quiz interactions, and group metadata. These materials represent actual coursework and study patterns, helping the system understand what students are learning and how they collaborate. The platform stores raw document content, extracted text, embeddings for semantic search, and interaction logs in a structured PostgreSQL database. This diverse dataset enables accurate content retrieval, personalized quiz generation, and adaptive concept mapping while maintaining a comprehensive knowledge base of each study group's learning materials.

B. Data Preprocessing

The core functionality of StudySync relies on processing diverse educational materials from student uploads and real-time collaboration data. Raw documents undergo specialized extraction—textract for PowerPoint presentations, pdf-parse for PDFs, and direct parsing for text notes. Extracted content is cleaned through noise removal, normalization, and chunking into semantically meaningful passages. Text embeddings are generated for semantic similarity matching in the RAG pipeline. Chat messages and quiz responses are preprocessed to extract learning patterns and engagement metrics.

This preprocessing ensures accuracy and reliability, enabling the system to provide precise content retrieval, intelligent quiz generation, and meaningful concept visualizations grounded in actual student materials..

C. Model Selection

The intelligent engine of StudySync combines rule-based retrieval with advanced language models to deliver context-aware learning tools. The RAG pipeline uses vector embeddings and semantic similarity algorithms to retrieve relevant passages from the knowledge base when processing student queries. For quiz generation, the system analyzes retrieved content to identify key concepts and formulate questions. Mistral AI's language models process both retrieved data and query context to generate coherent, pedagogically sound quizzes and concept maps.

D. Backend Development

The backend of StudySync is developed using Node.js with Express.js framework, managing core functions including RAG pipeline orchestration, document processing, real-time communication, and quiz generation. It follows a modular and scalable architecture with specialized routes for authentication, group management, file uploads, message handling, and quiz endpoints. The RAG system integrates with PostgreSQL for document storage and embedding vectors, uses Socket.IO for real-time collaboration, and connects to Mistral AI for content augmentation.

E. Frontend Development

The frontend of StudySync is designed for simplicity, accessibility, and intuitive collaboration, especially for students new to AI-powered learning tools. Built using React.js with TypeScript and Vite, it features modules like document upload, real-time chat, quiz display, concept map visualization, and group management. Interactive elements such as file drag-and-drop, live member tracking, responsive quiz interfaces, and visual concept maps enhance usability. The AuthContext manages session state, Axios handles API communication with the backend's RAG pipeline, and Socket.IO client ensures real-time synchronization of messages and collaborative updates across all study members.

F. AI Integration

The system includes an intelligent augmentation module that enhances responses by grounding them in retrieved study materials. When students request quizzes, the RAG system retrieves relevant passages from uploaded documents, which Mistral AI then analyzes to generate context-specific questions directly aligned with coursework. For concept mapping, retrieved content is processed to identify key topics, relationships, and learning objectives, creating visual representations that reflect actual material. The AI integration provides personalized recommendations for learning focus areas based on chat patterns and quiz performance, while multilingual support ensures clear guidance. This RAG-augmented AI approach ensures every response is evidence-based and grounded in the student's actual materials, not generic knowledge.

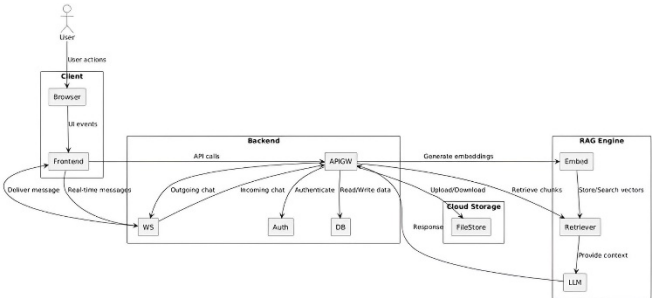
G. Deployment

StudySync is deployed on cloud platforms such as AWS, Azure, or Render to ensure scalability, availability, and high-performance real-time collaboration. Deployment involves containerizing the Node.js backend with Docker, configuring the PostgreSQL database with optimized indexing for RAG vector searches, and setting up CDN for frontend assets. Socket.IO is configured for real-time communication across distributed server instances using Redis adapters. The RAG pipeline is optimized for low-latency retrieval and embedding generation. Monitoring tools track system performance, API response times, and RAG retrieval accuracy. Regular updates improve model quality, add new document processing capabilities, and enhance the overall learning experience, ensuring reliable and continuous support for collaborative study groups.

IV. SYSTEM DESIGN

A. Dataflow Diagram

Figure 1 illustrates the complete data flow architecture of StudySync, showcasing how user interactions flow through multiple system layers to deliver intelligent, material-grounded learning responses.

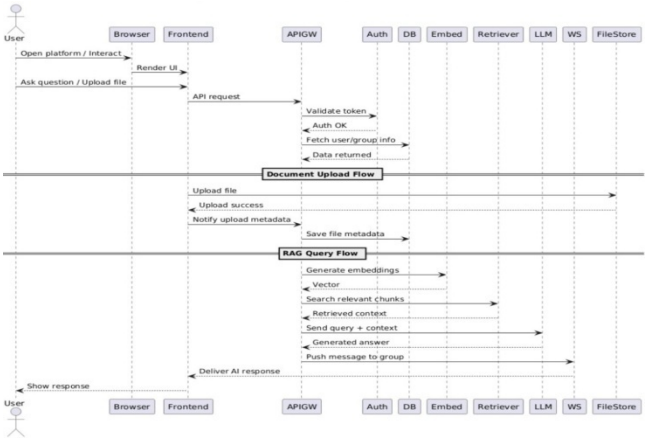


[Fig.1:DataflowDiagram]

- User uploads documents → stored in FileStore → immediately embedded and indexed in RAG Engine
- Student requests quiz → Retriever fetches relevant passages → LLM generates questions grounded in actual materials
- Real-time chat → distributed through WS to all connected members instantly
- Every response is augmented with retrieved context, ensuring accuracy and relevance to actual coursework

B. Sequence Diagram

Figure 2 ,the sequence diagram depicts the complete interaction flow between system components during two primary workflows: document uploads and intelligent query processing through the RAG pipeline.



[Fig.2:Sequence Diagram]

• User initiates action – Opens platform via browser, uploads document or requests quiz
• Authentication – Frontend sends API request; Auth validates JWT token and verifies group membership
• Data retrieval – System fetches user and group metadata from Database
• Embedding generation – Embed module processes documents into semantic vector embeddings
• Vector search – Retriever searches indexed embeddings for relevant passages matching query
• Context augmentation – Retrieved passages combined with user query
• LLM processing – Mistral AI generates material-grounded responses
• Real-time delivery – WebSocket pushes generated answer to user and all group members
• Display – Frontend shows AI response to user

C. Workflow Diagram

Figure 3 describes about StudySync's end-to-end user journey and system interactions. It illustrates how users authenticate, log in, and access the dashboard, then either upload study documents to Cloud Storage or submit queries. For queries, the RAG Pipeline generates embeddings, retrieves relevant material passages, and uses the LLM to create answers. WebSocket delivers real-time communication.

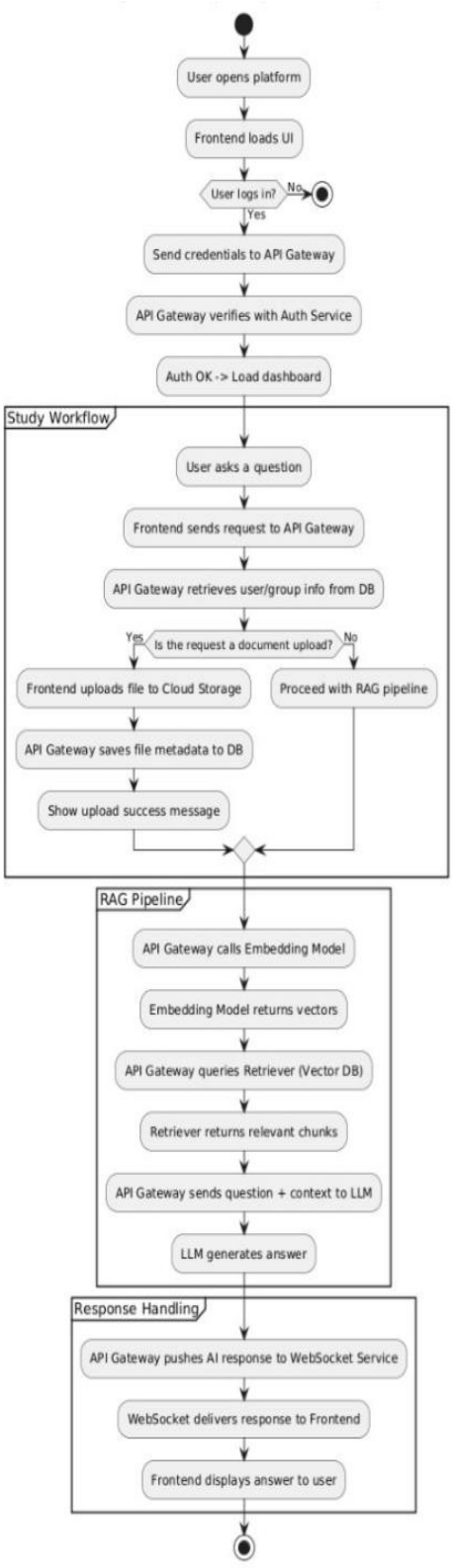
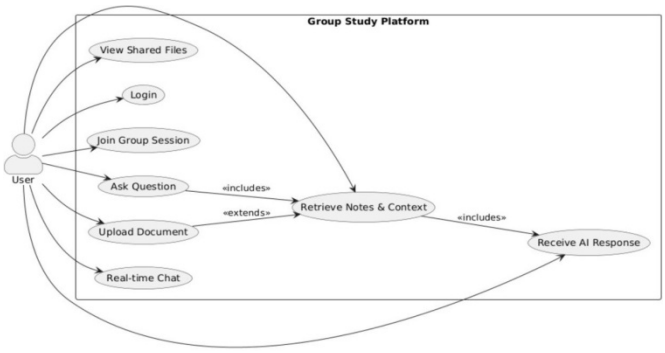


Fig.3:Workflow Diagram]

D. UseCase Diagram

Figure 4 illustrates all primary user interactions and system functionalities within the Group Study Platform. The diagram centers on the User actor and outlines six core use cases that define the platform's capabilities as shown in it.



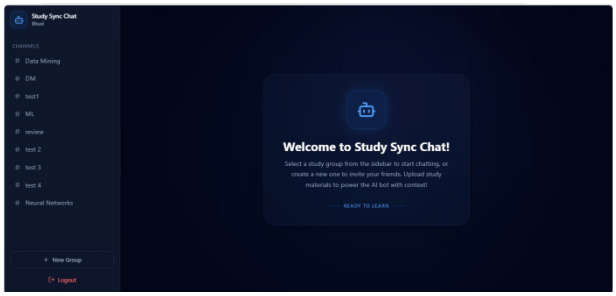
[Fig.4:Use case Diagram]

V. TESTING

A. *Testing*

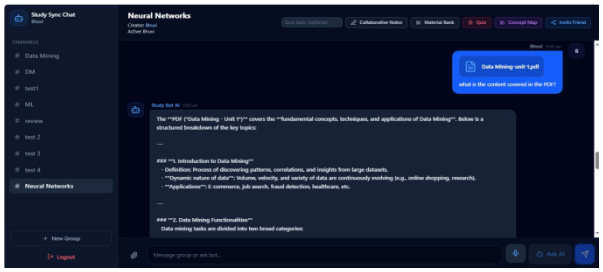
• Functional Testing: Document processing, RAG pipeline, quiz generation, concept mapping, and AI response delivery work correctly.
• Unit Testing: Individual modules (document extraction, embedding, retrieval, LLM integration, auth, database) tested independently.
• Integration Testing: Smooth interaction between document upload, embedding generation, vector search, LLM processing, and WebSocket communication.
• UI Testing: Document upload interface, real-time chat, quiz display, concept maps, and group management functionality.
• Data Validation Testing: Document formats, file sizes, embedding vectors, and invalid input handling.
• Response Accuracy Assessment: Generated quizzes, concept maps, and recommendations are accurate and grounded in actual uploaded materials.
• Performance Testing: Document embedding speed, vector search latency, LLM response time, and API efficiency under concurrent loads.

B. *Results*



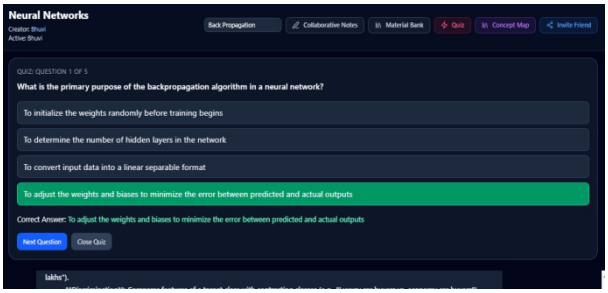
[Fig.5: Main UserInterface]

AsshowninFig. 5,The StudySync main page welcomes users with a clean, intuitive interface featuring a sidebar of available study groups. The central welcome message guides users to select or create study groups, upload materials to empower the AI-powered RAG bot, and start collaborative learning.



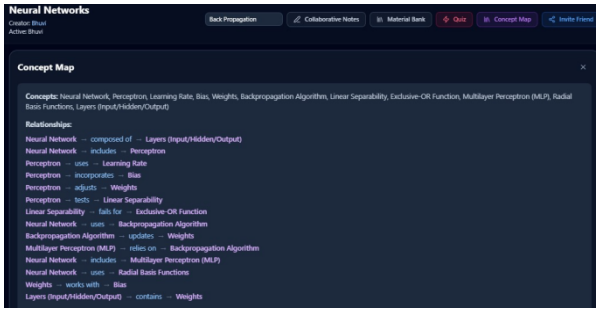
[Fig.6:Group Chat Creation]

Fig. 6 describes about a collaborative group study platform interface displaying an active study session. It shows real-time chat functionality where users query uploaded study materials and receive AI-generated responses grounded in the document content. The interface provides integrated features including quiz generation, concept mapping, material repository access, and collaborative note-taking, enabling seamless peer learning with intelligent AI support augmented by retrieved course materials.



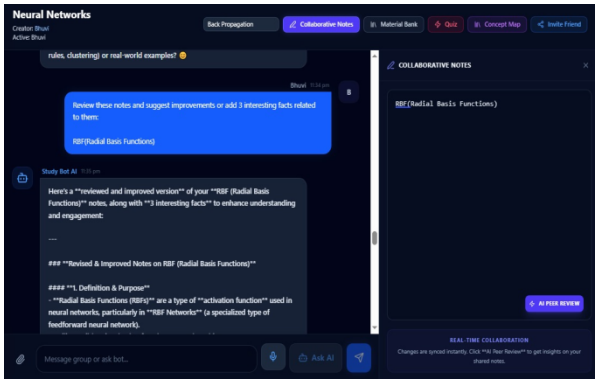
[Fig.7:Quiz Generation]

Fig. 7 describes about StudySync's AI-generated quiz feature within a study group session. It displays an interactive quiz question with multiple-choice answers extracted from uploaded course materials. The correct answer is highlighted in green, providing immediate feedback to students. This RAG-powered quiz generation automatically creates assessment questions grounded in actual study materials, enabling students to reinforce learning and test knowledge on course content in real-time.



[Fig. 8:Concept Map Generation]

Fig. 8 describes about StudySync's AI-generated concept map feature within a study group session. It visualizes key concepts and their interconnected relationships from course materials. The concept map organizes complex subject matter into a visual knowledge structure, helping students understand topic hierarchies and relationships to facilitate deeper comprehension of study materials.



[Fig. 9:Collaborative Notes Generation]

Fig. 9 describes about StudySync's Collaborative Notes feature enabling group members to annotate and improve study materials in real-time. It shows students reviewing tagged notes on course topics, suggesting improvements, and adding relevant insights to the shared knowledge base.

VI. CONCLUSION

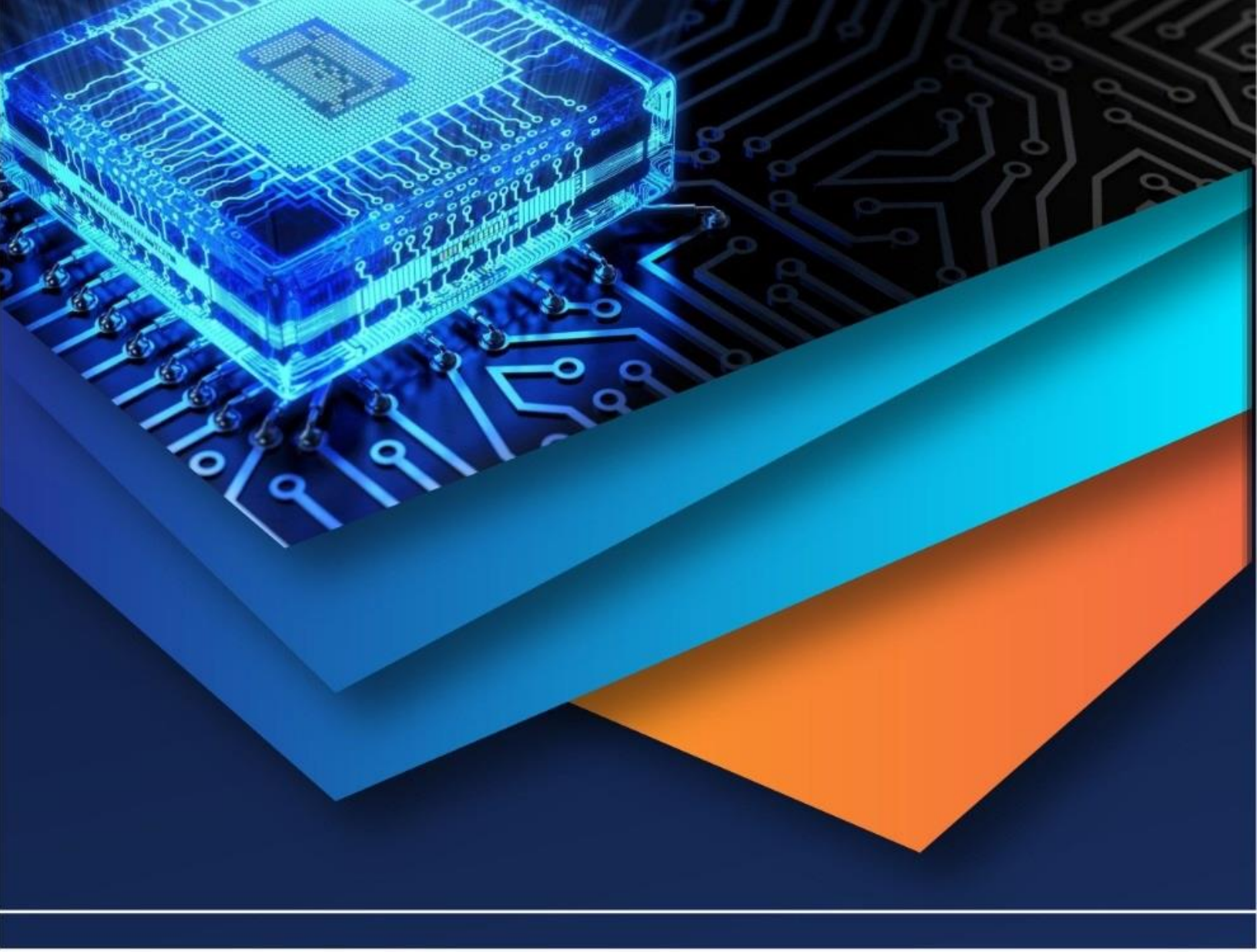
StudySync successfully demonstrates RAG technology's power in collaborative learning. By intelligently indexing uploaded materials and grounding AI responses in actual coursework, the platform delivers material-specific quizzes, concept maps, and recommendations. Real-time communication, secure authentication, and scalable architecture ensure a robust learning environment that addresses fragmentation of study materials and lack of intelligent content processing in existing platforms.

The platform ensures evidence-based learning support through RAG-augmented AI responses grounded in uploaded materials. Students receive material-specific guidance rather than generic information, promoting deeper understanding through concept mapping and faster knowledge retention through grounded quiz generation. This RAG-enhanced approach demonstrates how AI can revolutionize collaborative education by connecting all learning support directly to actual course content.

Future improvements include advanced embedding models for better retrieval accuracy, multi-modal RAG support for images and videos, learning analytics dashboards, adaptive quiz difficulty, offline-first architecture, and mobile applications. LMS integration and educator-focused features will enable broader institutional adoption, positioning StudySync as a comprehensive solution for modern collaborative learning environments.

REFERENCES

[1] D. Oreki, et al. (2024). "Retrieval Augmented Generation in Large Language Models."Presented at the International Symposium on AI Chatbots and University Education.
[2] Saad-Falcon, et al. (2025). "A Survey on Knowledge-Oriented Retrieval-Augmented Generation." Proceedings of the Annual Conference on Natural Language Processing and Education Technologies.
[3] Jacobs, Jaschke, et al. (2024). "Designing a Student-Friendly RAG-Based Chatbot." Journal of Educational AI and Personalized Learning.
[4] Velazquez, et al. (2024). "Enabling Educators to Build Specialized AI Chat Bots with RAG." International Journal of Instructional Technology and Teacher Education.
[5] Digvijay Singh, et al. (2024). "AI-Enabled Virtual Collaborative Learning Classroom." AI in Education Review.
[6] Sindhu M, et al. (2025). "Enhancing Student Support with a RAG Powered Chatbot." International Conference on Educational Data Mining.
[7] Attila Kovari, et al. (2025). "A Systematic Review of AI-Powered Collaborative Learning." Higher Education Technology Journal.
[8] EdTech Startups (Disco, 360Learning, Kahoot! AI) (2024). "AI-Powered Collaborative Learning Platforms." EdTech Innovations Conference Proceedings.



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)