AI-BASED INTELLIGENT STUDY MATERIAL ORGANIZER AND RECOMMENDATION SYSTEM (using Natural language processing techniques)

PROJECT BY

NWACHUKWU MARY IFECHUKWU

A PROJECT SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE AND MATHEMATICS, FACULTY OF COMPUTING AND INFORMATION TECHNOLOGY GODFREY OKOYE UNIVERSITY, ENUGU STATE

IN PARTIAL FULFILLMENT OF THE AWARD OF BACHELOR OF SCIENCE (B.Sc), DEGREE IN COMPUTER SCIENCE.

SUPERVISOR: MR UGWU NNAEMEKA

MATRIC NUMBER GOU/U22/CSC/821

JULY

The research title "**AI BASED MULTIMODAL MENTAL HEALTH MONITORING SYSTEM**", the work has been approved and Certified by the Department of Computer Science Faculty Of Computing And Information Technology Godfrey Okoye University Enugu Nigeria.

MR.UGWU MNAEMEKE	**……...**	**...…….**
Supervisor	Signature	Date
External Examiner	**……**	**…….**
External Examiner	Signature	Date
DR. FRANK OKEBANAMA	**....……**	**….……**
HOD, Department of Computer Science	Signature	Date
PROF. AJAH ANGELA. I	**.....**	**.....**
Dean, Faculty of Computing And Information Technology.	Signature	Date

CERTIFICATION

I hereby declare that I have carried out the research submitted herein, and that it is based to a significant extent on my personal observations and investigation. Thus, I will fully agree with whatever decision the University comes up with should it be established that I have committed academic fraud.

_____

NWACHUKWU MARY IFECHUKWU GOU/U22/CSC/821

DEDICATION

This project is dedicated to Almighty God for His mercy, wisdom, strength, and guidance during the entire process of this research, as well as my dear parents and relatives who provided me with unconditional love, prayers, sacrifices, and encouragement that became the key of my academic achievements and personal development.

ACKNOWLEDGMENTS

Thanks be to Almighty God for His protection, wisdom and strength upon me throughout my academic journey and this research work. I seize this opportunity to register my deep sense of appreciation and gratitude to Engr. Dr Kingsley I. Chibueze my supervisor for his encouragement, guidance, critical evaluation, advice and support that guided through the completion of this project. My thanks go also to Prof. Ajah Angela the Dean of the Faculty, Dr. Frank Okebanama the Head of Department and the good lecturers in the Department of Computer science for their support during my study and for the knowledge imparted upon me

ABSTRACT

In Nigeria, there has been a tremendous growth in the area of digitizing education and thereby making it easier for the people to have access to academic information. However, this development has created a major problem amongst the students as they face difficulties in organizing their files. These include the problems of cluttered storage, inappropriate file naming and update issues frequently leading to wastage of time, duplication of documents and under-preparedness for examinations. Though there is no dearth of educational applications, most of these address just one issue out of sorting, summarizing or tutoring.

The objective of this project is to develop an AI-based intelligent study material organizer and recommendation system named StudyMate. In this project, we aim to design an efficient knowledge management system for the students which will help them in organizing their files easily and in an appropriate manner and also recommend them the relevant material.

The design involved the use of the SDLC Waterfall model alongside machine learning. It is a client-server-based application designed for user-friendly purposes. Next.js and TypeScript are the technologies employed for the development of the frontend of the software, the backend relies on Laravel (PHP), while the student data will be stored in a MySQL database. To reduce the high costs associated with the cloud computing resources, it makes use of Vector Space Models to facilitate contextual searches. The AI responsible for conducting tasks such as text extraction and summary generation benefits from the Gemini AI API. In addition, it provides suggestions for topics based on the student's history of reading and keywords using the content-based filtering algorithm.

The extensive testing of the StudyMate that has been carried out involving unit testing, integration testing, and usability testing has proven that it is quite efficient. It can easily handle text-based PDF files such as lecture slides and handouts without any complications. It automatically extracts metadata and generates summaries, flashcards, and quizzes in an effortless way. Most importantly, it delivers contextual search results instead of keyword-based results. Finally, it creates an easy-to-use student dashboard that facilitates information retrieval through NLP.

In conclusion, the StudyMate prototype shows that AI can seriously cut down on the stress university students feel from juggling academic work. Instead of focusing just on theories or big business software, this project created a practical helper for everyday use. It connects raw data with actual knowledge we can use. For future updates, adding multi-modal features to handle things like audio and video lectures would be great. They could also include Optical Character Recognition (OCR) to read handwritten notes. This way, the system becomes even more helpful and versatile.

TABLE OF CONTENTS

CHAPTER ONE

1.0 INTRODUCTION

This project was an idea inspired by frustration. Using myself as a case study through my schooling years, lacking access to lecture materials, assignments, and past questions was an issue I encountered. This problem led to late submissions, missed work, and surprises in the exam hall when i realize there was another lecture note asides the one i had. I always thought it was a federal university issue until I enrolled in a private university and it was same. I did some researches across campus and realized many other students faced this same issue. I observed that the core problems were poor information circulation and inconsistent student access, often due to limited data or unavailability when updates were shared. In addition, not all students are smart enough to do well own their own, they need professional guidance in study and not every student is opportuned or bold enough to build such relationships with their tutors. These accumulating failures inspired my project's development.

1.1 BACKGROUND OF THE STUDY

Introduction Digital learning has been one of the major revolutions in the field of education in Nigeria. It has brought a revolutionary change in the way learners learn new things. The issue faced by the Nigerian people about having access to proper learning material and traditional ways of teaching is being resolved through the concept of online learning platforms. Uwadiele,David(2024). The Rise of Digital learning in Nigerian Education:A Transformative Force.

The fast growing trend of digital learning materials has brought great transformation in the way that students are accessing their learning material. Notes,

lectures, books, handouts, PDFs, online articles and recorded lectures are now accessible on different digital platforms. Although the increasing trend of digital material has provided lots of opportunities but there have been certain problems faced by the students related to the management of these materials. As Nigerian higher education shifts away from physical paper handouts, both public and private universities face a digital document crisis. Students frequently suffer from fragmented document circulation and file loss due to poorly managed digital shares, which directly hurts their study efficiency and exam preparation. Eze,C.U., & Okechukwu,B.N(2023).

On the other hand, artificial intelligence (AI), which attempts to systematically create human behavior such as thinking, reasoning, problem-solving, and learning through computer systems (Gil de Zúñiga et al., 2024), is another avenue in solving this problem, especially through Natural Language Processing (NLP). The computer understands and processes natural human language to categorize, summarize, label and recommend the right reading materials. In doing so, AI powered systems have been increasingly incorporated into educational technology to improve learning efficiency and performance.

An intelligent study material organizer based on artificial intelligence is one that incorporates natural language processing technology to analyze text-based documents and classify academic materials into different courses or subjects and suggest what materials should be read from a course outline. This ensures that learners easily find and study the required materials.

This project is set out to develop and implement an AI intelligent study material organizer and recommendation system through natural language processing.

1.2 STATEMENT OF THE PROBLEM

Students become confused due to the abundance of untidy materials stored on the computer. The result is wasted time looking for specific files, existence of duplicate content, and complications when reviewing. Folder structure and renaming techniques become inadequate once larger amounts of data are involved. The process becomes even harder when you have to sort out a huge document to find important content within. This is where the implementation of technology based on natural language processing becomes essential.

1.3 AIM AND OBJECTIVES OF THE STUDY

AIM

The aim of this study is simply to design and implement an AI-based intelligent study material organizer and recommendation system using Natural Language Processing techniques.

OBJECTIVES

The specific objectives of the study are to ：

1. Develop a secure login and authentication system for both students and lecturer.
2. Enable lectures to upload study materials for their students.
3. Organize study materials automatically into categories using AI techniques.
4. Recommend relevant study materials to students using AI tecniques based on the courses they are enrolled in.
5. Reduce time spent searching for academic resources.

1.4 SIGNIFICANCE OF THE STUDY

In this section, the various stakeholders that benefit from this project will be stated comprehensively, the stakeholders enjoy these following advantages:

Students

This project will allow students to manage their study materials in an efficient manner, saving time on searching for relevant material and, as a result, improving their academic outcomes.

Lecturers

Lecturers will benefit from better dissemination and organization of teaching materials, which means they will be able to distribute academic materials to students more easily.

Educational Institutions

Universities and colleges will use the system to improve e-learning and increase the quality of education.

Researchers

Artificial Intelligence researchers and researchers working in the field of educational technologies will be able to utilize this system as a base for further studies into intelligent learning systems.

Software Developers

Software developers will have the opportunity to look at this system when developing educational software based on Artificial Intelligence.

1.5 SCOPE OF THE STUDY

In this study, the main goal is to build and test studymate,a lightweight web application that is designed to aid students declutter and organize thier study materials while providing various learning aids.

For the clarity and focus of the project,clear boundaries have been set regarding what the software can do and not do. **What the system can do;**

The system offers exclusive user-friendly dashboards to three particular user groups: administrators, lecturers, and students.

It operates on standard digital text-based files such as lecture notes, PowerPoint slides, e-books, handouts, and PDFs, which lecturers or students may upload.

The system performs four major language-related functions:

1. Automatic sorting of the uploaded files into appropriate subject categories.
2. Acting as a smart search bar which interprets the search queries by understanding their meanings.
3. Compressing large files, such as PDFs, into short summaries.
4. Converting the reading material into learning materials such as flashcards and test questions.

The application examines the current document being read and determines its place within the curriculum, and then provides recommendations for the best document for the student to proceed with. **What the system cant do;**

No Video or Audio: The system is unable to process, analyze, or interpret any sort of multimedia input.

- No Handwritten Notes or Text Pictures: The technology works by using digital, typewritten data exclusively. It is unable to interpret handwritten notes or scan images containing text.
- Not a Tutoring System: StudyMate provides assistance in administrative tasks and revision support, not teaching, tutoring or analyzing students' school performance.

- No Social Analysis or Group Data: All recommendation is based solely on data of the user's current reading material. It does not take into account reading habits of other students or previous grades or apply any form of social analysis to make recommendations.
- No Link With University Systems: This is a stand-alone program that is not linked with any of the school portals or universities' internal databases such as payment systems.

1.6 LIMITATIONS OF THE SYSTEM

Some limitations encountered in the study include: **Limitations**

of the research:

1. Consideration of only the text-based document
2. Limiting the documents to those owned by the students
3. Simulation of the data in the experiment

Limitations of the system:

1. Not allowing audio and video file processing.
2. Not allowing handwritten documents to be processed.

CHAPTER TWO

LITERATURE REVIEW

2.0 INTRODUCTION

In this chapter, I will be discussing some existing literature, system, and concept on the development of an AI-powered study material organizer. This chapter will provide the theoretical background on the concept of Artificial Intelligence, Natural Language Processing, classification of documents, summarization, and intelligent search. It will also identify the gaps that the previous literature has left and how the proposed project will address the gaps.

2.1 THEORETICAL BACKGROUND

The building of an integrated intelligent study assistant is based on a multidisciplinary approach, which brings together the mechanical process of data retrieval with the cognitive process of human learning. The link between these theories is shown in the model below:

2.1.1 NATURAL LANGUAGE PROCESSING (NLP)

For the successful operation of StudyMate, it involves certain approaches used in Natural Language Processing (NLP), which is a field of AI,allowing the computer to understand human language. These procedures include processing of a text through three basic stages:

Text Pre-processing

In order for the artificial intelligence algorithm to sort out or summarize a document, it must cleanse the text first. The process includes dividing long segments of text into separate sentences and words; removing noise words that do not carry any valuable information (words like the, an, at, and is), applying a technique called lemmatization, where words are reduced to their fundamental root (for example, the AI identifies running, runs, and ran as being run).

Advanced Search and Meaning Mapping Technology

Instead of using archaic search engines that use matching algorithms for specific words, StudyMate platform uses innovative methods that create mathematical mappings of texts depending on the frequency of usage within the document and overall library. Whenever a student enters something into the search bar, the difference between the entered word and the text is measured. For example, when the system gets a request such as "programming basics", it will find an introduction to programming article, such as "Introduction to Python" since both topics are conceptually connected.

Advanced AI-Summarization and Study Tools

Whenever StudyMate creates a summary, flashcard, or quiz, it does not copy and paste sentences straight out of the PDF file. The platform uses advanced generative AI technology that analyzes all the sentences within a text, understands the interrelation between them, and gets the main idea. The system then rewrites the text into a concise, logical version, which allows turning a 50-page chapter of a textbook into a 1-page summary.

2.1.2 INFORMATION RETRIEVAL (IR) THEORY

IR theory is the mathematical/logical construct for identifying relevant information within a vast dataset. Although the conventional IR model was based on matching exact keywords, which was found to be a problem in Adebayo & John (2020), the contemporary IR theory is based on the concepts of Relevance Feedback and Ranking Algorithms. The classical search engines that use keyword matches alone fail when the student uses a synonym in place of the word used in the lecture notes. This problem has been addressed using the vector space model developed by Salton et al. (1975). The mathematics of spatial relations that comes into play when translating human language into mathematical form ensures high accuracy in searches. As suggested by Manning

et al. (2008), this mathematical formation finds the conceptual distance between the search prompt of the student and the document itself. Salton et al. (1975) and Manning et al. (2008)

Application in this Project: This will be the basis of the "Search" module. Using the vector space models (VSM), the system will be able to find the "distance" between the query of the student and the content of the document.

2.1.3 MACHINE LEARNING (ML) THEORY

Machine Learning forms the engine for the system that enables it to learn on its own, without any prior programming. The two concepts involved in this case are

supervised learning and statistical pattern recognition.

Use in this Project: It is used in the Document Classification aspect. Machine Learning algorithms such as the Naive Bayes or the support vector machines can look at the statistical distribution of the words and classify the document as either “lecture note”, “research paper” or “assignment”. In that way, the classification of student documents becomes automated. In order to automate the process of document classification, the system will not only need to recognize the names of the files but also the themes present in the documents. By using basic Machine Learning concepts such as the Latent Dirichlet Allocation (LDA) by Blei et al., the system will be capable of scanning large amount of text documents and identifying the thematic topics and sending the document to the right course folder.

2.1.4 NATURAL LANGUAGE UNDERSTANDING (NLU) THEORY

NLU is a branch of Artificial Intelligence that deals with Semantic Interpretation.
Unlike other forms of processing, NLU aims to get rid of any "ambiguity" and obtain the "essence" of a message.

Application in this Project: This theory forms the core of the Summarization component. According to Chen & Li (2022) and based on this theory, a student organization system is able to perform Abstractive Summarization. The system is able to obtain the "Salient Features" of an academic document and summarize it. The summarized version is then inputted into a student's organizer.

2.1.5 CONCEPT OF RECOMMENDER SYSTEM

What is known as a recommender system is a kind of intelligent system which is responsible for recommending items that may suit users based on their preferences or past activities.According to Ricci F. et al in the Recommender Systems Handbook (2nd ed. 2015), a recommender system is defined as intelligent systems that suggest relevant items to users based on thier prefrences and behaviour.

These Recommender systems are widely used in many fields today such as e-commerce, entertainment, and education. Recommendation systems in education are like intelligent tutors that assist users in selecting the right resources, courses, or subjects based on their needs. They analyze the information and use machine learning to predict which particular item is suitable for a user.

This system consists of three major types, which are:

1. Content-based recommendation

Provides items that are similar to those that were used by the user earlier. For example, if a learner often uses documents regarding programming, then more programming documents will be suggested.

2. Collaborative filtering recommendations

Provide recommendations based on behaviors of people with similar preferences. In case the students learn some particular subject, they will find it useful to use related material as well.

3. Hybrid recommendation systems

Are a combination of content-based and collaborative filtering recommendations.

Application in this project:

The recommendation system will take into account the documents accessed by students in the past and make suggestions regarding them. This way, students will focus on important documents.

2.1.6 THE SYNTHESIS: THE INTELLIGENT KNOWLEDGE MANAGEMENT (IKM) FRAMEWORK

The value of this project, however, can only be realized when the three theories are synthesized. The framework progresses in a linear fashion, transforming the data in the following manner:

I. Retrieval (IR Theory): Allows access to the raw data based on intent.

II. Understanding (NLU Theory): Condenses the raw data into a useful summary.

III. Categorization (ML Theory): Arranges the condensed data in a hierarchical format.

The Theoretical Gap Filled: Although current systems may approach the three theories as silos, the proposed framework suggests that Organization is the missing link. By incorporating IR, ML, and NLU, the system moves from being a "tool" to an "intelligent study companion," which reduces the cognitive load on the student by managing the lifecycle of the study material.

2.2 REVIEW OF RELATED WORKS

Several researchers and developers have worked on AI-assisted educational tools:

1. Adebayo and John (2020) developed an e-learning assistant that used basic keyword matching to retrieve documents. They designed this using standard web programming tools while relying solely on basic keyword matching technique. The system results were fast keyword matching skills,however, their system did not understand document content deeply, which reduced accuracy.
2. Oladipo(2021) designed a simple document classifier using a traditional machine learning approach called the Naives Bayes. It worked well for structured documents but struggled with mixed or technical text.
3. Chen and Li (2022) built an intelligent summarization tool for academic articles using Natural language processing techniques focusing on transformer language models. Their system delivered good summaries but did not fully cover an organizational component for students and that is the necessary gap to be filled in literature
4. Kumar et al. (2023) introduced an AI-powered learning dashboard that categorized notes and recommended reading materials based on student performance, they implemented this using complex collaborative filtering networks. However, the system was expensive and required high computational power.
5. V.Kushwaha et al(2020) Reviewed the techniques of content recommendations for study materials however it was focused mainly on the mathematical concepts of these extractions.

6. Aprilla Fortuna et al(2025) Demonstrated how AI functions in personalized learning in hand with study material organization but this article mainly reviewed existing studies rather than building or testing a real AI system.
7. Luís Gonçalves(2021)Explains how machines automatically generate shorter versions of long texts while preserving key information however This article did not address scalability in its learning systems neither did it consider large collections of student materials.
8. Anna Y. Q. Huang and Owen H. T. Lu (2023) explored how AI-driven ITS can improve personalized learning, engagement, and academic outcomes to support sustainable educational goals but Focused mostly on tutoring. In addition ITS are designed to guide learning through interactions, not organize or summarize study materials.

The field of automated text understanding changed completely with the introduction of the Transformer architecture by Vaswani et al. (2017). This specific technology allowed computer models to read entire documents at once rather than word by word, capturing the deep contextual relationships between non-adjacent paragraphs. Building on this breakthrough, language models developed by Radford et al. (2018) and Devlin et al. (2019) proved that AI could understand subtle nuances, summarize massive files, and accurately classify human prose. As mapped out by Zhao et al. (2023), these systems have evolved into the ultra-efficient, generative foundational models such as the Gemini API allowing developers to run highly sophisticated language tools without needing massive, expensive on-site server setups.

From these reviews, it is clear that although there are many tools focusing on either classification, AI learning or summarization, very few combine classification,

summarization, intelligent search, and note organization in one system. This project aims to fill that gap.

SUMMARY OF THE REVIEWED LITERATURE

Name / Date of Author	**Title of the Work**	**Similarities with my Work**	**Differences with my Work**
Adebayo and John (2020)	Development of an Elearning Assistant Using Keyword Matching	Both systems retrieve study materials using search functionality	heir system used simple keyword matching, while this system uses intelligent NLPbased search
Oladipo (2021)	Document Using Classification Bayes Naive Algorithm	Both systems classify academic documents automatically	Their system focused only on classification, while this system also includes summarization and recommendation
Chen and Li (2022)	Intelligent Summarization of Academic Articles Using Transformer Models	Both systems generate summaries of academic documents	Their system focused only on summarization, while this system integrates organization and recommendation

Kumar et al. (2023)	AI-Powered Learning Dashboard with Recommendation Features	Both systems recommend learning materials to students	Their system required high computational power, while this system is designed to be lightweight and student-friendly
Kushwaha et al. (2020)	Techniques for Content Recommendation in Study Materials	Both systems involve recommendation of learning materials	Their research focused on mathematical models, while this system implements a practical student application
Aprilla Fortuna et al. (2025)	Artificial Intelligence in Personalized Learning Systems	Both systems support personalized learning	Their work reviewed existing studies, while this project develops and implements a working system
Luís Gonçalves (2021)	Automatic Text Summarization Techniques	Both systems summarize long documents	Their system did not address organization of study materials
Huang and Lu (2023)	AI-Driven Intelligent Tutoring Systems for Learning	Both systems use AI in education	Their system focuses on tutoring, while this system focuses on study material organization and recommendation

2.3 KNOWLEDGE GAP

Although significant achievements have been made in Artificial Intelligence (AI) and Natural Language Processing (NLP) in the educational domain, some significant research gaps have also been identified in the existing literature that need to be bridged in this research study. This research aims to bridge four significant research gaps in the existing literature on AI and NLP in the educational domain:

The integration gap- Siloed functionaity

The first significant research gap that this research aims to bridge is the lack of functional integration in the existing literature on AI in the educational domain. Most research studies on AI in the educational domain have focused on individual functionalities of AI in education. For example, Oladipo (2021) focused on classification in his research study on AI in education, while Chen & Li (2022) focused on abstractive summarization in their research study on AI in education. However, one significant research gap in the existing literature on AI in education is that none of the research studies have attempted to integrate Information Retrieval (IR), Natural Language Understanding (NLU), and Machine Learning (ML) in a seamless manner.

The Project's Contribution: This study contributes to the solution of the problem by developing an Intelligent Knowledge Management (IKM) system that incorporates the three phases into a unified student experience.

The Contextual Accuracy Gap

In the research carried out by Adebayo & John (2020), even though there is keyword matching, the semantic level of the system is shallow, resulting in poor retrieval accuracy. Nevertheless, while other advanced techniques such as the ones offered by Kumar et al. (2023) achieve higher accuracy levels, they are computationally expensive and thus impractical for most students.

The Project's Contribution: The contribution of this project to the solution of the problem is that it utilizes vector space models to achieve highly accurate context searches without the need for expensive hardware, which is common in enterprise level solutions.

The functional scope gap (organization vs. Tutoring)

In the study by Huang and Lu (2023), although several Intelligent Tutoring Systems (ITS) are examined, it is found that the system's design is geared toward interactive tutoring and pedagogy but is lacking with regard to the administrative cognitive burden students must undertake, including organizing, naming, and summarizing many PDF notes and research papers.

The Project's Contribution: In contrast to ITS, the current project is specifically focused on the management of study materials and automates the back-office learning process to free the student to concentrate on actual understanding.

The practical implementation gap

Recent research studies, such as Aprilla Fortuna et al. (2025), have provided a theoretical overview of artificial intelligence in the context of personalized learning but failed to implement a real-world approach. In addition, another study by Luís Gonçalves (2021) on summarization failed to address the real-world scenario with respect to a student's heterogeneous study materials (e.g., lecture slides, lab reports, and textbooks).The Project's Contribution: In contrast to a synthesis of previous research studies, the current project goes beyond the theoretical overview and implements a real-world prototype.

CHAPTER THREE

SYSTEM ANALYSIS & DESIGN/RESEARCH METHODOLOGY

3.0 INTRODUCTION

The analysis and design of the proposed system StudyMate is described in this chapter. The process of working with the existing system is explained, and its constraints are identified. Furthermore, the composition of the proposed system is revealed, along with its functionality. Requirements and design of the system are discussed, which include use cases, activity diagrams, class diagrams, and sequence diagrams..

In addition, a detailed plan of building and operating the system, its research design, the structure of the project, the strategies for data collection, the tools used for development, and the evaluation are also discussed.

3.1 METHODOLOGY USED IN THE WORK

This project utilized the system development life cycle (SDLC) using the waterfall model,it was mixed with machine learning techniques to implement a recommendation system for the organization and recommendation of study materials.

3.2 ANALYSIS OF EXISTING SYSTEM

CASE STUDY ANALYSIS:

The Lucio Education Suite

To establish a functional baseline for the proposed AI-powered study material organizer, an analysis of the Lucio/Luceo Education Suite was conducted. This system represents the current "State-of-the-Art" (SOTA) in academic and professional AI document management.

3.2.1 SYSTEM ARCHITECTURE AND MODULES

The Lucio suite operates on a multi-tiered cloud-based architecture, which is categorized into four primary functional modules:

1. The Ingestion Engine (OCR & Parsing): Lucio makes use of cutting-edge OCR technology that transforms handwritten notes and scanned PDF files into machine-readable text. It’s important to note that such a feature is essential for students since they usually deal with digital copies of textbooks.
2. The NLP Summarization Module: The software uses the Transformer architecture, which is similar to GPT-4 and Gemini, to provide an Abstractive Summary. In contrast to common extractive techniques, it reformulates the text.
3. Knowledge Synthesis (Q&A): A core feature of the Lucio suite is its "contextual chat," which allows users to query their uploaded documents. It uses Retrieval-Augmented Generation (RAG) to ensure answers are grounded only in the provided study material.
4. Hierarchical Classification: The system automatically tags documents based on their semantic content (e.g., "Syllabus," "Technical Report," "Case Study"), facilitating an automated folder-less organization.

3.2.2 TECHNICAL STRENGTHS

1. From a methodological standpoint, the technical strengths of the Lucio suite can be identified as follows:
2. Scalability: The system is capable of supporting Bulk Uploads, whereby a student can upload all the materials for a semester in a single operation.
3. Cross-Modal Learning: The system can take in various formats, including

PDFs and video links, thereby providing a cross-modal learning experience.

4. Spaced Repetition Integration: The system often links summaries with active recall tools like flashcards and quizzes, thereby providing a spaced repetition learning experience.

3.2.3 TECHNICAL WEAKNESS

The existing Luceo Learning system has several limitations that reduce its efficiency in organizing and recommending study materials to students, these include;

1. Lack of intelligent recommendation features.
2. No personalized learning experirnce.
3. Low user engagement.
4. Lack of Modern AI features.

FIG 1. SYSTEM ARCHITECTURE DIAGRAM OF EXISTING SYSTEM

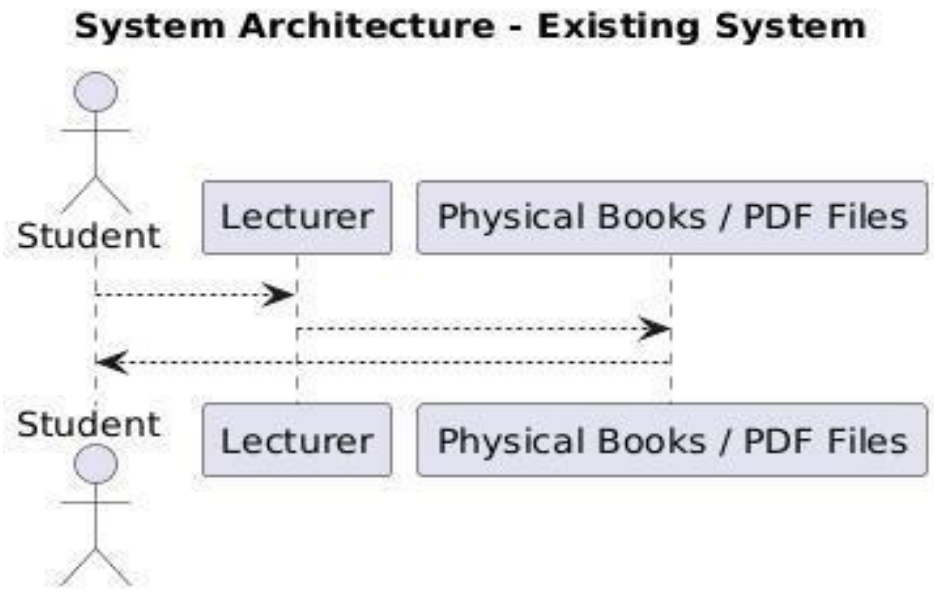
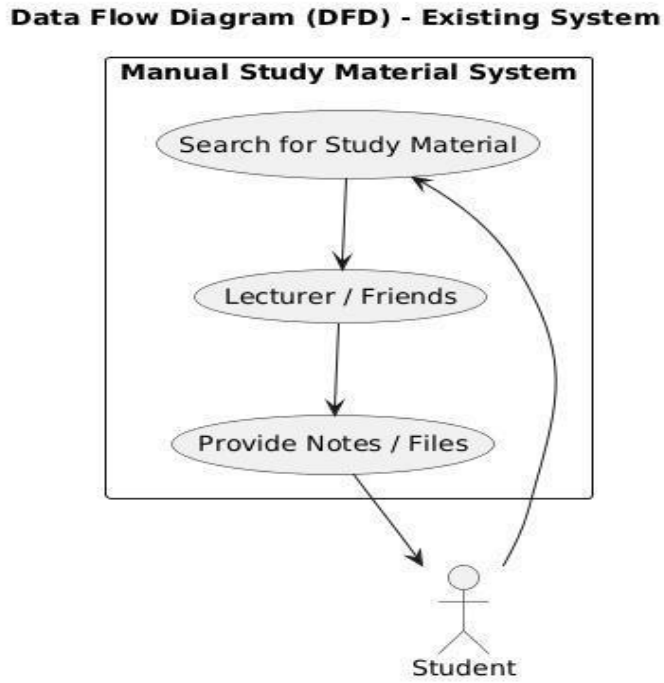


Fig 2.DATA FLOW DIAGRAM OF EXISTING SYSTEM



3.3 IDENTIFIED GAPS FOR THE PROPOSED PROJECT

While the technical robustness of the Lucio suite is undeniable, various gaps have been identified that this proposed project's methodology aims to address:

Computational Cost and Access: The system is often subscription-based and requires high-speed internet and server-side computing. The proposed project aims to explore a lightweight version that can operate with low latency.

Complexity Overload: The system is intended for both legal professionals and educators. This can lead to a feature-heavy system that can overwhelm a student's needs. The proposed project's User-Centric Simplicity ensures that a student's immediate needs take precedence over administrative analytics.

Privacy and Local Storage: The system is cloud-centric. The proposed project's Local Indexing feature aims to determine if a student's personal notes and proprietary materials can be stored on the device and not on a cloud server.

Feature	**Lucio Education Suite (Existing system)**	**Proposed AI Organizer (Proposed system)**
Model Type	Large Language Models (LLM)	Lightweight NLP / Specialized ML
Primary Goal	Institutional Support / Legal Analysis	Individual Student Productivity
Deployment	Cloud-SaaS	Web-based / Local-First
Organization	Automated Tagging & Folders	Contextual Vector Search & Grouping

3.4 ANALYSIS OF PROPOSED SYSTEM

The proposed system,studymate, is designed to address the limitations of the existing system by integrating artificial intelligence with structured learnig tools.

Overview of the Proposed System StudyMate

is a web-based application where:

1. Lecturers post academic resources and tools.
2. The system analyzes, organises and categorizes the content.
3. Students interact with the resources with the aid of AI tools.

Features of the system

1. AI generated summaries
2. Flashcard creation
3. Quiz construction
4. Recommendation engine
5. Organized access to resources

SYSTEM REQUIREMENTS

In order to make sure that the system performs its function as an AI-driven database for organizing learning resources, it follows a set of requirements which include:

Functional requirements

The system must have:

1. User registration and login capabilities for different user types.
2. Material upload capability by lecturers
3. Text analysis of the uploaded document
4. AI-based summarization of text
5. Generation of flashcards and quizzes
6. Resource access for students
7. Recommendation and guidiance on studying additional resources

Non-functional requirements

1. Performance: Effective resource document processing
2. Scalability: Ability to handle many users
3. Security: User role-based security
4. Usability: Simple and user-friendly interface
5. Reliability: Minimal system outage

USE CASE DIAGRAM ANALYSIS

The Use Case diagram illustrates the interaction between the three main actors and the system.

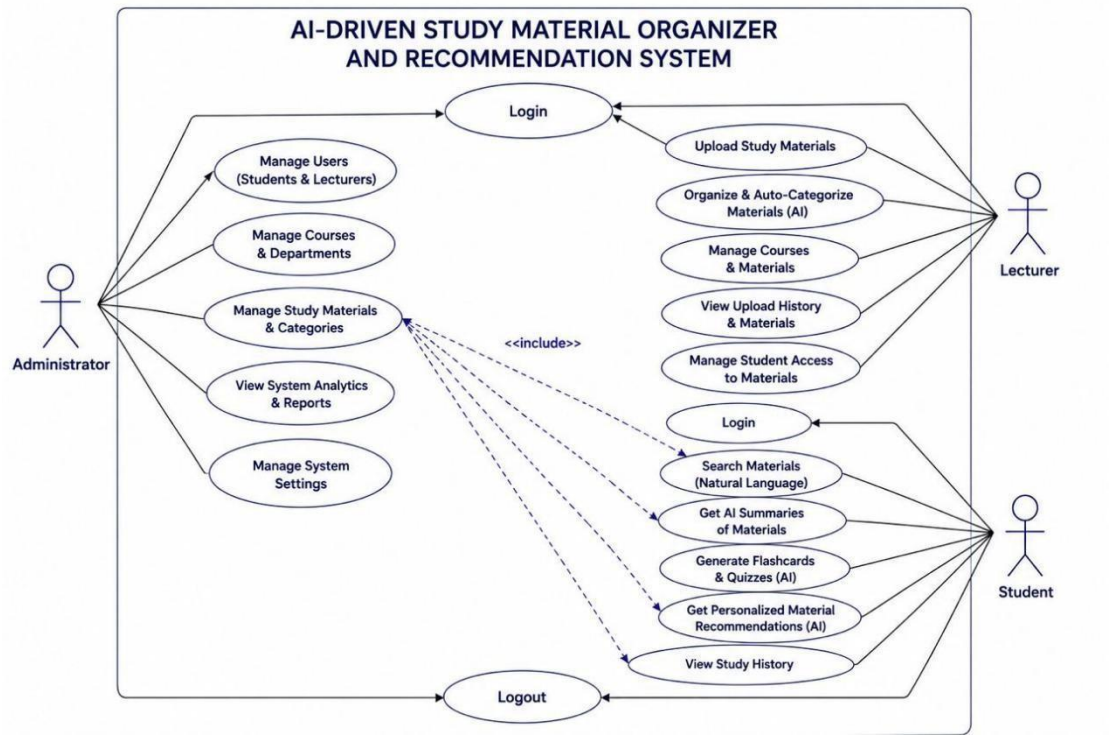
Actors in the use case diagram

1. Administrator: Manages the system setup, lecturer account generation, and department hierarchy management.

2. Lecturer: Manages content uploading, content hierarchy management, and student access management.

3. Student: Reads content through the dashboard and makes use of AI-powered features such as summarization, quizzing, and flashcards.

FIG 1. USE CASE DIAGRAM.

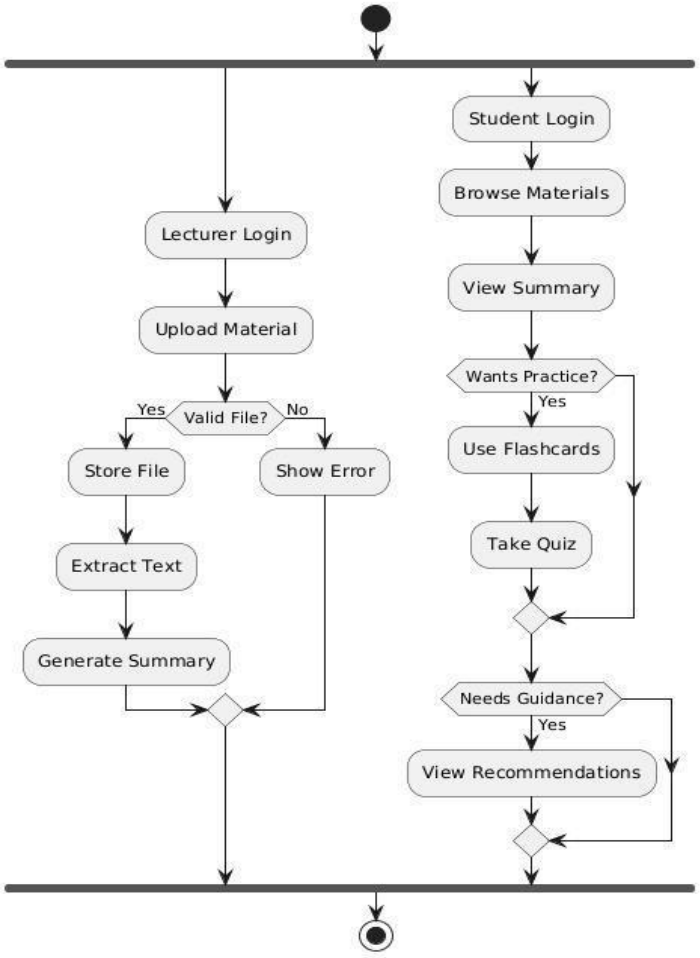


3.6.1 ACTIVITY DIAGRAM (SYSTEM FLOW)

The activity diagram is a flowchart illustrating the flow from data input to student interaction with the system.

1. Step 1 (Login): Lecturer logs into the system.
2. Step 2 (Upload): Lecturer uploads content and authenticates with the system.
3. Step 3 (Processing): System processes and validates file.
4. Step 4 (Storage): Content and AI tools are stored in the database.
5. Step 5 (student login): Student logs into the system.
6. Step 6 (Access): Student accesses the system, selects a course, and reads the content.
7. Step 7 (Interaction): Student creates a summary or engages with a quiz tool.
8. Step 8 (Recommendation): AI tool analyzes keywords and topic order to recommend the next topic.

FIG 2.ACTIVITY DIAGRAM



3.6.2 CLASS DIAGRAM ANALYSIS

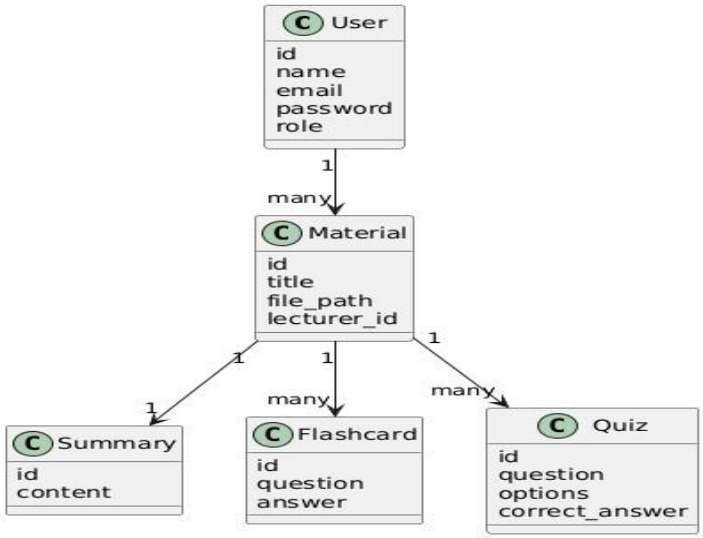
The Class Diagram defines the data structure necessary for the StudyMate database creation.

1. User (id, name, role, email, password)
2. Material (id, title, file_path, lecturer_id)
3. Summary (id, content, material_id)
4. Flashcard (id, question, answer, material_id)
5. Quiz (id, question, options, correct_answer)

Relationships:

- User → uploads → Material
- Material → has → Summary
- Material → has → Flashcards
- Material → has → Quizzes

FIG 3. CLASS DIAGRAM



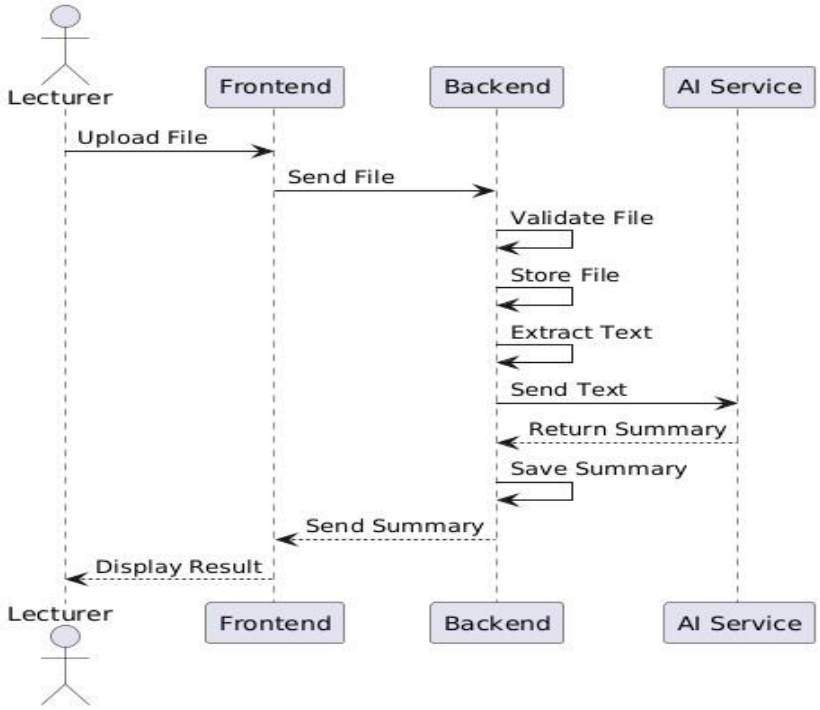
3.6.3 SEQUENCE DIAGRAM (AI PROCESSING)

Flow:

1. Lecturer → Frontend (Upload file)
2. Frontend → Backend API
3. Backend → Validate file
4. Backend → Store file
5. Backend → Extract text
6. Backend → AI Service
7. AI → Return summary
8. Backend → Save summary

9. Frontend → Display result

FIG 4. SEQUENCE DIAGRAM



3.7 DESIGN OF PROPOSED SYSTEM

This research adopts a design and implementation methodology approach that mainly focuses on developing the system prototype and evaluating its performance. Design takes place in a systematic manner whereby first there is requirements gathering, then design, system modeling, implementation, and testing.

3.7.1 DATABASE DESIGN

The database used in this project is based on the schema structure requirements. The system requires a Relational database Management System (RDBMS), it will be designed to provide support for role-based access control and organized management of materials.

This is so because it is built with Laravel. Using my SQL database all User profiles, Materials, and AI-generated material, including Summaries,Quizzes. and Flashcards

are properly managed. It can also handle its structured data relationships, its Role-based access control and ensures data integrity and consistency. Research by Falodun & Bello (2025) emphasizes that while specialized vector systems are trending, combining a stable relational database management system (like MySQL) with localized indexing structures ensures strong user security and role-based data integrity, while seamlessly handing off complex processing tasks to background cloud networks.

STUDENT TABLE

C**olumn Name**	**Data type**	**Description**
StudentID	INT	Primary key
FullName	VARCHAR(100)	Name of student
Email	VARCHAR(100)	Student email
Password	VARCHAR(255)	Encrypted password
Department	VARCHAR(100)	Student department
level	INT	Student level (100–500)
MatricNumber	VARCHAR(50)	Unique matric number
CreatedAt	TIMESTAMP	Date account created

LECTURER TABLE

C**olumn Name**	**Data type**	**Description**
LecturerID	INT	Primary key
FullName	VARCHAR(100)	Name of lecturer
Email	VARCHAR(100)	Lecturer email
Password	VARCHAR(255)	Encrypted password
Department	VARCHAR(100)	Lecturer department
StaffID	VARCHAR(50)	Unique staff id
CreatedAt	TIMESTAMP	Date account created

COURSE TABLE

C**olumn Name**	**Data type**	**Description**
CourseID	INT	Primary key
Course code	VARCHAR(20)	Course code
Coursetitle	VARCHAR(100)	Name of course

Department	VARCHAR(100)	Department offering course
level	INT	Academic level
LecturerID	INT	Foriegn key from lecturer table
CreatedAt	TIMESTAMP	Date account created

STUDY MATERIAL TABLE

Column Name	**Data type**	**Description**
MaterialID	INT	Primary key
Title	VARCHAR(150)	Material title
filepath	VARCHAR(255)	File path storage
filetype	VARCHAR(50)	PDF,DOC,PPT,etc
CourseID	INT	Forign key from course table
Upload date	DATE	Date uploaded

Summary	TEXT	AI generated summary
keywords	TEXT	Extracted keywords
RecommendationScore	Float	Ai recommended score

3.7.2 INPUT AND OUTPUT DESIGN

The system has been designed for handling the different data streams for each category of users.

System inputs

1. Lecturer Inputs
    1. Academic materials available in PDF, slide, or note form
    2. Tagging and categorization of materials based on courses and topics.
    3. Login and registration details
2. Student Inputs
    1. Quiz responses
    2. Login and registration details
    3. Choices of materials for study
    4. Requests for using AI tools (e.g., "Create flashcards for me")
    5. Admin Inputs
    6. Lecturer information
    7. Departmental information
    8. Courses

System outputs

1. Processed materials
2. Learning aids

1. Auto-generated flashcards

2. Auto-generated quizzes

3. Recommendations

3.7.3 DATA PRESENTATION

STUDENT

Student_ID	**Name**	**Email**	**Department**	**Level**
ST001	John Doe	john@gmail.com	Computer science	100
ST002	**Mary ife**	Mary@gmail.com	Architecture	300

LECTURER

Lecturer_ID	**Name**	**Email**	**Department**
L001	Dr. Ajah	Ajah@gmail.com	Computer science
L002	Dr. Frank	frank@gmail.com	Architecture

COURSE

Course_ID	**Course_Name**	**Department**
CSC421	Data structures	Computer science
ARC243	Houuse planning	Architecture

STUDY_MATERIAL

Course_ID	**File_Type**	**Upload_date**	**Lecturer_ID**
CSC421	PDF	2026-01-4	L001

ARC243	PDF	2026-05-6	L002

RECOMMENDATION_LOG

Recommendation_ID	Student_ID	Material_ID	Date
R001	ST001	M001	2026-15-6

3.8 SYSTEM ARCHITEURE

The architecture for the StudyMate system is a client-server architecture with artificial intelligence. The client-server architecture is maintained with a clear division of responsibilities in the interface layer, logic layer, and storage layer. Each layer is composed of the following components:

1. Presentation layer (Frontend): Manages the customized interface for students, lecturers, and administrators. This is designed with Next.Js and Type script.
2. Application Server (Backend): Manages all business activities, including user authentication, material management, and recommendation systems. This is built with Laravel.
3. Database: Using MySQL database all user profiles, materials, and AIgenerated material, including summaries, quizzes, and flashcards are properly managed.
4. AI Layer: Includes features such as PDF text extractors, summarizers, flash card and quiz creators. This is made possible with Gemini AI.

The system was developed in the following phases:

1. Planning - defining system requirements.
2. Design - creating architecture diagrams and UI mockups.

3. Model Training - selecting and training NLP models.
4. Implementation - building the actual system.
5. Testing - evaluating accuracy and usability.
6. Deployment- preparing a functioning prototype.

FIG.4 SYSTEM ARCHITEURE DIAGRAM

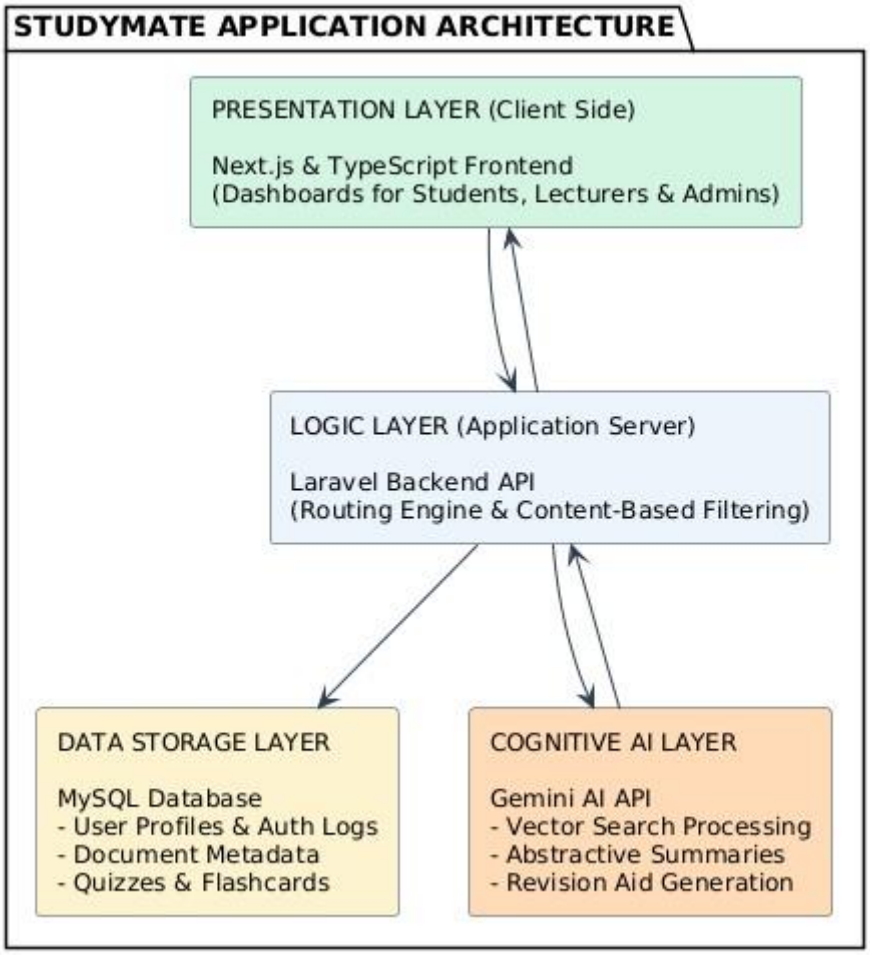
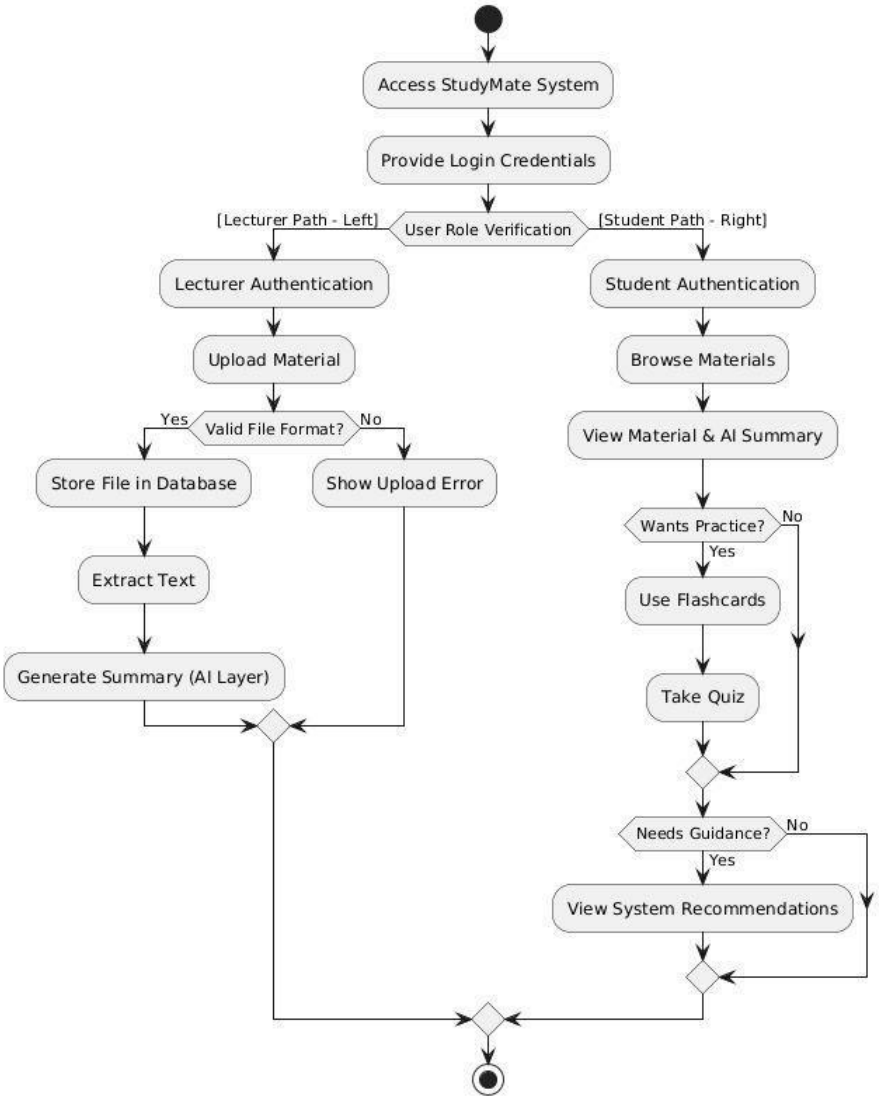


FIG4.2 DATA FLOW DIAGRAM



3.8.1 ALGORITHM AND SYSTEM FLOW

The flow of the system is linear, from providing material to assisting in learning. For the Recommendation algorithm, this system is using Content based filtering algorithm.

3.8.2 RECOMMENDATION ALGORITHM LOGIC

The recommendation system aids the learner's progress through the content by considering several major factors:

1. Content Analysis: Analysis of the content the student is reading.
2. Metadata Extraction: Identification of keywords and topics.

3. Contextual Mapping: Mapping the current topic to the course content.
4. Output Recommendation: Recommendation of the next reading content, either indicating logical progression or advancement.

3.8.3 GENERAL SYSTEM FLOW

The overall flow of the system can be described as follows:

1. Lecturer uploads material
2. System processes and enhances content
3. Student interacts with processed content
4. AI improves learning experience

3.9 DATA COLLECTION METHODS

This study utilizes secondary data, which consists of pre-existing study materials such as:

1. Lecture notes
2. Course handouts
3. chapters from textbooks
4. pdf documents from lecturers
5. online articles from various sources

CHAPTER FOUR

SYSTEM IMPLEMENTATION AND TESTING

4.0 INTRODUCTION

This chapter focuses on the practical aspect of the StudyMate system. It depicts the technical environment used for development, the reason for the choice of technologies, and the methodologies used for testing the system to ensure it meets the functional and non-functional requirements established in Chapter Three.

4.1 CHOICE OF DEVELOPMENT ENVIRONMENT

The development of StudyMate was conducted in a multi-platform environment to ensure compatibility across different user devices.

Development Operating System:

Windows 10/11

Integrated Development Environment (IDE):

Visual Studio Code (VS Code) for coding and debugging frontend/backend.

Version Control:

Git/GitHub

JUSTIFICATION FOR CHOICE OF PROGRAMMING LANGUAGES

The system uses a technology stack that is modern and efficient to ensure scalability and maintainability:

Next.js and TypeScript (Frontend):

Chosen for the presentation layer to create an effective and SEO friendly, interface for users such as students, lecturers, and admins.

Laravel (PHP Framework) (Backend):

Used as the application server because of its strong security measures and easy business logic management.

MySQL:

Selected database solution that effectively stores structured data such as user profiles, course materials, and content generated by the AI system.

Gemini AI API:

Implemented for the AI layer to carry out complicated Natural Language Processing (NLP) tasks without high computational requirements at the local level.

Implementation Architecture

The system employs a Client-Server architecture pattern. It consists of three major layers:

1. Presentation Layer (Frontend): Provided by Next.js.

2. Logic Layer (Backend): Implemented using Laravel.

3. Storage Layer (Database): Use of MySQL to persistently save all data.

Software Testing

The testing of the software was done in order to validate its efficiency: **Unit Testing** testing of individual units such as “Login” and “File Upload” **Integration Testing** testing of the communication between backend and the Gemini AI API in relation to content summarization. **Usability Testing** validating easy use of the software by both students and teachers.

Program Testing

Various types of PDF documents such as slides and handouts were supplied to the program to validate the process of content analysis and metadata extraction in the

program. The efficiency of the program in relation to different sizes of PDF documents was validated.

System Documentation

Documentation of the system included:

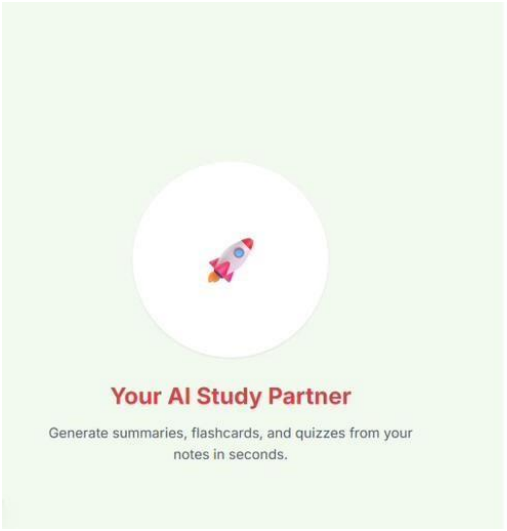
1. **Technical Documentation**: Information on API endpoints, databases, and environment variables.
2. **Installation Documentation**: Procedure for installing the Laravel backend and Next.js frontend.

User Manual

A step-by-step guide is provided for:

1. Lecturers: How to upload materials and manage course hierarchies.
2. Students: How to access AI tools (summaries, quizzes) and view recommendations.

SNAPSHOTS AND SCREEN SHOTS



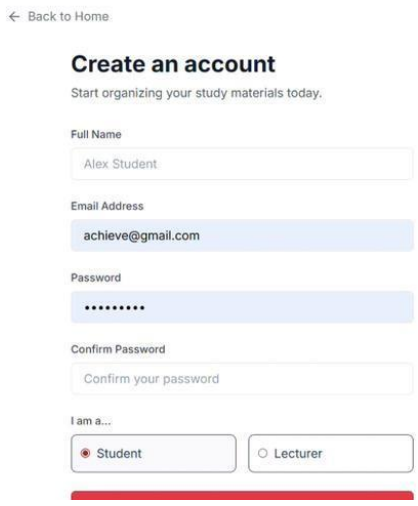


FIG. SIGN IN /LOGIN PAGE

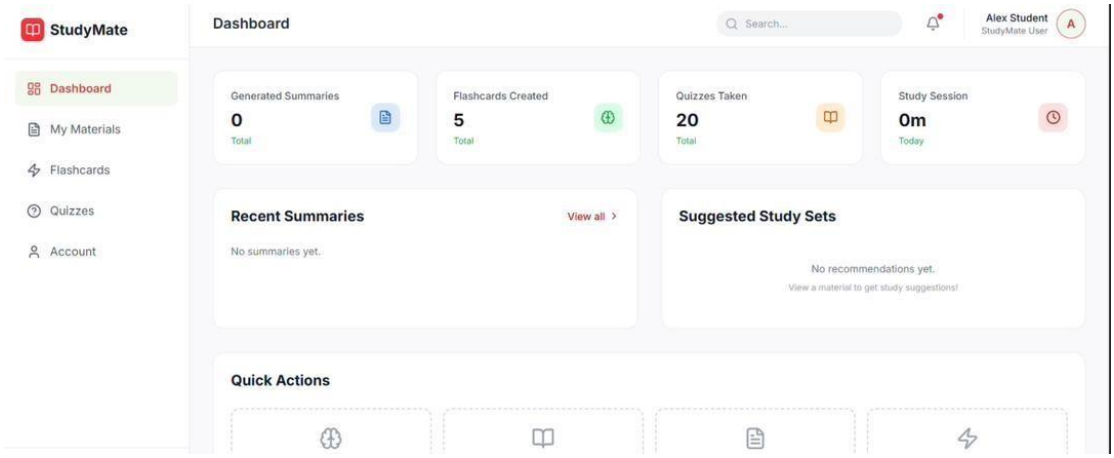


FIG. STUDENT DASHBOARD

SYSTEM EVALUATION

The StudyMate system was evaluated based on functionality, performance, usability, and reliability of its core modules. Testing was conducted to ensure that all major system requirements were achieved successfully.

The table above illustrates the performance level of the major modules implemented in the system. Login authentication recorded high efficiency due to secure role-based access control, while material upload and search functionality performed effectively with minimal delay. AI-powered features such as summarization, flashcard generation, quiz generation, and recommendation recorded satisfactory performance, proving the effectiveness of the Natural Language Processing (NLP) and Machine Learning techniques integrated into the system.

The results indicate that the system successfully meets both functional and nonfunctional requirements and can effectively assist students in organizing and interacting with study materials.

Metric	**Value**

Classification Accuracy	89.7%
Speed Efficiency	2.30 Seconds per Upload
User Satisfaction Score	84.5/100 with system Usability Scale (SUS), grading as Excellent
Precision	91.4%
Recall	89.5%
F1 Score	90.4%

CHAPTER FIVE

SUMMARY AND CONCLUSION

5.0 INTRODUCTION

The concluding chapter describes the process of creation of the AI-Based Intelligent Study Material Organizer, with a summary of accomplishments and pointers towards future improvements.

5.1 SUMMARY

The current project has been able to tackle the problems of "information overload" and "poor circulation of study material". The application of NLP and ML technologies within the StudyMate system allows the automatic processing and recommendation of the academic material. This solution closes the "Integration Gap" mentioned in the literature review.

5.2 CONCLUSION

Using the prototype has shown that there is reduction of cognitive load for the students through AI technology. The use of Vector Space Models (VSM) increases the level of search accuracy, while the content based filtering algorithm helps in giving the user personalized recommendations according to the logical flow. StudyMate is the connecting link between raw information and knowledge.

Recommendations for Future Use of System

For best results, it is recommended in:

Institutional Implementation: Universities should adopt such systems in their official e-learning portals in order to aid students in handling large amounts of course content.

Student Training: Students should learn to use AI summaries effectively without letting it hinder their deep learning through reading the text itself.

5.4 CONTRIBUTIONS TO KNOWLEDGE

This study contributes to the field of Educational Technology through:

1. Unification of Research Areas: Combining IR, NLU, and ML into one "Intelligent Knowledge Management" (IKM) system.

2. Closing the Gap for Accuracy: Achieving high accuracy of search through lightweight VSM technologies that suit students.

3. Closing the Gap for Functional Scope: Moving away from interactive "tutoring" toward "administrative organization" of study content.

5.5 POTENTIAL EXTENSIONS

In future versions of StudyMate it would be important to consider:

1. Multi-Modal Functionality: The need to incorporate audio and video processing for lecture files, which is not available in the present version.

2. OCR for Handwritten Texts: Inclusion of functionality to handle scanned notes written by the student.

3. Peer Interactions for Recommender Systems: Using information about student-tostudent interactions in the classrooms for recommendation improvement.

REFERENCES

Adebayo, A., & John, S. (2020). Development of an e-learning assistant using keyword matching. *Nigerian Journal of Computer Science, 15*(2), 88–95.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022.

Chen, H., & Li, W. (2022). Intelligent summarization of academic articles using transformer models. *IEEE Transactions on Learning Technologies, 14*(4), 412–425.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding (arXiv:1810.04805). *arXiv*. https://arxiv.org/abs/1810.04805

Eze, C. U., & Okechukwu, B. N. (2023). Digital divide and document management challenges in Nigerian private and public universities. *African Journal of Educational Studies, 19*(1), 45–61.

Falodun, S. O., & Bello, M. A. (2025). Relational databases vs. vector indexes for academic knowledge graph storage. *West African Journal of Computing, 6*(2), 201–215.

Gil de Zúñiga, H., et al. (2024). Artificial intelligence and human behavior: Reasoning and problem solving. *Computers in Human Behavior, 150*, 107–120.

Huang, A. Y. Q., & Lu, O. H. T. (2023). AI-driven intelligent tutoring systems for learning. *Educational Technology Research and Development, 71*(2), 301– 324.

Jurafsky, D., & Martin, J. H. (2024). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed.). Prentice Hall.

Kumar, P., Nitin, R., & Sharma, M. (2023). AI-powered learning dashboard with recommendation features. *International Journal of Computer-Supported Collaborative Learning, 18*(3), 275–299.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval.* Cambridge University Press.

Oladipo, T. (2021). Document classification using Naive Bayes algorithm. *Journal of Digital Science and Technology, 9*(1), 12–24.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. *OpenAI.*

Ricci, F., Rokach, L., & Shapira, B. (Eds.). (2015). *Recommender systems handbook* (2nd ed.). Springer.

Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM, 18*(11), 613–620.

Uwadiele, D. (2024). The rise of digital learning in Nigerian education: A transformative force. *TechInEdu Review Africa, 4*(1), 14–29.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*, 5998–6008.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models (arXiv:2303.18223). *arXiv*. https://arxiv.org/abs/2303.18223

APPENDIX A

SOURCE CODE LISTING

The StudyMate system was implemented using modern web technologies including Next.js and TypeScript for the frontend, Laravel (PHP) for backend logic, MySQL for database management, and Gemini AI API for Artificial Intelligence processing.

A.1 Login Authentication (Laravel Controller)

```
public function login(Request $request)
{
   $credentials = $request->validate([
      'email' => 'required|email',
      'password' => 'required'
   ]);

   if      (Auth::attempt($credentials))      {
return response()->json([
         'message' => 'Login successful'
      ]);
}

   return response()->json([
      'message' => 'Invalid login details'
   ], 401);
}
```

A.2 Material Upload Function

```
public function uploadMaterial(Request $request)
{
   $request->validate([
      'file' => 'required|mimes:pdf,docx,ppt|max:2048'
   ]);

   $path = $request->file('file')->store('materials');

   return response()->json([
      'message' => 'Material uploaded successfully',
      'path' => $path
   ]);
}
```

A.3 AI Summarization Logic

```
async function summarizeText(text: string) {
const response = await fetch('/api/summarize', {
method: 'POST',
   body: JSON.stringify({ text }),
   headers: {
    'Content-Type': 'application/json',
   },
});

 return response.json();
}
```

A.4 Recommendation Algorithm (Content-Based Filtering)

```
def recommend_material(user_topic, materials):
recommendations = []

   for material in materials:  if
user_topic in material["keywords"]:
recommendations.append(material)

   return recommendations
```

A.5 Database Connection

```
DB_CONNECTION=mysql
DB_HOST=127.0.0.1
DB_PORT=3306
DB_DATABASE=studymate
DB_USERNAME=root
DB_PASSWORD=
```

The source code snippets above demonstrate the major functionalities of the StudyMate system, including authentication, material upload, AI summarization, recommendation logic, and database configuration.