JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Course Crafter: Recommendation System Using AI

Prof. Manasi Kulkarni
dept. of Computer Engineering
Pillai College Of Engineering
New Panvel, India
manasikulkarni@mes.ac.in

Anushka Menon
dept. of Computer Engineering
Pillai College Of Engineering
New Panvel, India
anushkam22comp@student.mes.ac.in

Akash Satpute
dept. of Computer Engineering
Pillai College Of Engineering
New Panvel, India
sakash22comp@student.mes.ac.in

Hemant Waghmare
dept. of Computer Engineering
Pillai College Of Engineering
New Panvel, India
whemant22comp@student.mes.ac.in

Abhishek Nambiar
dept. of Computer Engineering
Pillai College Of Engineering
New Panvel, India
abhishekn22comp@student.mes.ac.in

***Abstract*— The proliferation of educational content on video platforms like YouTube presents learners with unprecedented access to knowledge, yet the quality variance and overwhelming volume create significant navigation challenges. Traditional recommendation algorithms typically prioritize engagement metrics over educational value and fail to consider individualized learning objectives. This paper presents the successful implementation of *Course Crafter*, an AI-driven personalized course recommendation and generation system. The project integrates Google Gemini API for natural language understanding and YouTube Data API for curated video selection. The system automatically generates a structured course outline, fetches high-quality educational videos, summarizes content, and provides quizzes for continuous learning evaluation. Learners can save their progress and resume later, creating a complete self-paced, adaptive learning platform.**

Keywords*— Intelligent recommendation systems, online education, content enhancement, knowledge graphs, natural language understanding, personalized learning, educational video analysis, AI in education, learning pathways

I. Introduction

In the last decade, the rapid expansion of online learning platforms has transformed the educational landscape. Platforms such as YouTube, Coursera, Udemy, and edX have made knowledge accessible to millions of learners around the world. However, this digital abundance has created a new challenge — *information overload*. Learners are often faced with thousands of tutorials, lectures, and articles for even the simplest topics. Selecting appropriate and sequential learning material becomes a time-consuming and confusing process, especially for beginners who lack the expertise to evaluate content quality and relevance.

Despite the availability of powerful search engines and recommendation algorithms, these systems generally focus on popularity metrics such as views, likes, or trending status rather than pedagogical value or conceptual continuity. As a result, learners are left with unstructured content, repetitive explanations, and incomplete understanding. There is an urgent need for intelligent systems that can curate structured learning experiences automatically, just as a human instructor would.

To address this gap, our research introduces Course Crafter — an AI-based personalized course generation and learning platform that utilizes the capabilities of modern artificial intelligence models to create tailored educational experiences for every learner. The primary goal of Course Crafter is to combine AI understanding, content recommendation, and interactive learning assessment into a single automated system.

Unlike traditional e-learning tools, Course Crafter allows users to start with a simple input prompt, such as “Learn Python Programming,” and instantly generates a complete, logically sequenced course structure. This structure is composed of topics and subtopics, each linked to high-quality educational videos sourced directly from YouTube using the YouTube Data API. The platform then uses the Google Gemini API, an advanced generative AI model, to produce concise topic summaries, highlight key concepts, and even create short quizzes for self-evaluation.

Furthermore, the system is designed for self-paced learning. Learners can track their progress, store completed lessons, and resume from where they left off — all through an intuitive, interactive web interface. This combination of personalization, automation, and adaptive content generation allows Course Crafter to deliver a unique learning experience that mimics the logical structure of a professional online course but without any human curation.

The novelty of this research lies in the integration of two intelligent APIs — Gemini and YouTube — to form a cohesive educational workflow. The Gemini API handles text-based intelligence: understanding user intent, generating outlines, summarizing videos, and creating quizzes. The YouTube API ensures that the chosen learning resources are current, authentic, and popular among learners. Together, they form the foundation of an AI-driven educational ecosystem capable of dynamically generating courses in any subject area.

Moreover, Course Crafter represents a shift from passive content consumption to active learning. By including quizzes and progress tracking, the system engages the learner, ensuring retention and understanding. The model not only saves time spent searching and filtering information but also enables scalable, on-demand education for individuals, institutions, and organizations.

In a broader context, this work contributes to the ongoing movement toward AI-assisted education and adaptive e-learning systems. With increasing adoption of large language models, there is potential to create intelligent digital tutors that personalize content in real time. Course Crafter demonstrates how combining natural language processing (NLP), video curation, and interactive learning design can simplify and enrich the educational process.

The objectives of this implementation can be summarized as follows:

- To reduce the cognitive load on learners by automating course structuring and content selection.
- To use AI language models for generating comprehensive yet concise learning outlines and summaries.
- To leverage YouTube educational resources for multimedia learning support.
- To incorporate interactive quizzes for knowledge retention and self-assessment.
- To create a user-centric system that remembers progress, offering flexibility and continuity.

II. Literature Review

A. Natural Language Processing and Input Processing

To ascertain preferences and purpose, these models examine user interactions, including text inputs, clicks, and browsing patterns. Natural language processing (NLP) techniques aid in the extraction of valuable insights from user input.

- To effectively interpret user inquiries regarding course preferences, difficulty levels, and topic interests, BERT (Bidirectional Encoder Representations from Transformers) leverages extensive pre-training to grasp context in both directions. Its key innovations include the Masked Language Model (MLM) and Next Sentence Prediction (NSP), which enable the development of deep, bidirectional contextual representations without directional limitations [1].
- To increase language understanding, RoBERTa, an optimized version of BERT, eliminates the NSP aim and makes use of bigger training datasets and batch sizes. This makes it ideal for applications like suggesting relevant courses based on user-specified criteria, where accurate understanding of user input is crucial [2].
- Reimers and Gurevych's Sentence-BERT adapts BERT with a Siamese network structure to generate semantically significant sentence embeddings. This is especially helpful in educational settings where semantic similarity—rather than just keyword matching—is more significant between user inquiries and course descriptions. Even when different language is utilized, Sentence-BERT allows for the precise and efficient retrieval of pertinent content that is consistent with user intent [3].

B. Video Retrieval Models

Based on verbal or visual searches, these models index, search, and extract pertinent video information using deep learning techniques. These models use multimodal learning, semantic comprehension, and feature extraction to enhance user experience and retrieval accuracy.

- Using Explicit Negative Examples to Enhance Ranking Function Learning Agarwal examined in this work how adding negative examples to information retrieval systems might improve their ranking functions. This study is especially pertinent to the discovery of instructional content, because it is essential to discern between pertinent and irrelevant resources. Retrieval systems can have a better understanding of what makes educational information suitable for certain user demands by learning from both good and negative examples. This method aids in eliminating courses that may match keywords but are unsuitable for the user's learning goals or skill level[4].

- Utilizing Natural Language Query to Search a Video Database discusses the increasing demand for efficient video retrieval systems in light of the rapidly expanding amount of video content. The researchers note that content and rule-based models, which are domain-dependent and necessitate structured queries, are frequently used in traditional video databases. Two distinct models are proposed in the paper. An image captioning method is used to implement the first model. The second model employs a method of audio processing. By offering a more flexible and user-friendly interface for video retrieval that adjusts to users' natural language patterns instead of requiring them to learn specific query syntax, the system marks a substantial improvement over earlier methods [5].

- According to Integrated Semantic–Syntactic Video Modelling, the database and information retrieval communities use entity-relation or object-oriented models to capture semantics, whereas the video processing and computer vision communities usually use structural video models (shot-based or object-based) with low-level descriptors like color, texture, shape, etc. In order to bridge the gap between the two conventional approaches, they propose a novel generic integrated semantic-syntactic video model that incorporates semantic entities (video objects and events) and their relations in an organized way. This model combines both textual and low-level descriptors within a single comprehensive framework [6].

C. Recommendation and Ranking Models

Based on user interactions, these models dynamically modify recommendations by utilizing machine learning approaches including collaborative filtering, fuzzy logic, and deep learning.

- Collaborative Filtering is an advanced method of course suggestion that uses data mining techniques to find significant trends in the courses that students choose. The algorithm creates individualized course recommendations with anticipated performance outcomes by integrating clustering, association rule mining, and sequential pattern analysis.To boost suggestion quality, their suggested approach mainly uses collaborative filtering that is supplemented with data mining tools [7].

- By incorporating several coefficients like Pearson correlation and chi-square correlation, among others, Graph Neural Network and Social Network Analysis integrated Social

Network Analysis with Graph Neural Network to thoroughly examine the relationship between courses and user preferences. This method provides a fresh viewpoint for understanding the complexities of the MOOC learning environment [8].

- A course recommendation system that employs the fuzzy logic technique. The fuzzy logic framework offers a suitable approach for managing the complexity and unpredictability that come with choosing a course. Four functional blocks make up the system's design, which adheres to a traditional fuzzy logic process: fuzzification, rule evaluation, aggregation, and defuzzification. The author's work improves on earlier recommendation systems while addressing significant drawbacks found in existing methods, especially by combining subjective interest evaluations with objective performance measurements [9].

- Deep Learning illustrates how deep learning may significantly boost large-scale recommendation systems' performance. It offers useful insights into the planning, execution, and upkeep of such systems, especially when considering YouTube's peculiar difficulties. Through the inclusion of both objective performance measures and subjective interest assessments, deep learning enables the system to learn intricate, non-linear user item interactions and adjust to the dynamic nature of YouTube's content [10].

- Session-based suggestions, Recurrent Neural Networks (RNNs) are suggested. This approach is suitable for situations in which users engage with the system continuously. By modeling the order of user interactions, this method captures the temporal interdependence in user behavior. In terms of course suggestions, this allows the system to take into account the user's educational background and suggest material that expands on previously learned material. Since learning is progressive and ideas frequently build upon one another, the model's sequential structure fits in nicely evaluations of ST [11].

- Neural Collaborative Filtering (NCF), which integrates the advantages of neural networks and matrix factorization for recommendation tasks, was presented in the study. Complex, non-linear user-item interactions can be efficiently modeled by NCF. When it comes to course recommendations, NCF can identify minute trends in the ways that users engage with different types of instructional materials, making it possible to forecast more precisely which courses would be most helpful for certain students. When past interaction data from several users is available, this method is especially beneficial [12] .

- Re-ranking for recommendation using Mask Pretraining reveals that current re-ranking models are not very good at taking advantage of local mutual impact across item groups, which might cause recommendation lists to shift categories suddenly. By proposing a novel re-ranking framework based on the idea of "scenes" as the fundamental unit of analysis, the authors address the problem. According to formal definitions, a scene is a collection of objects that includes a key item that establishes the scene's characteristics. All of the objects in the scene are distinct from one another but preserve a high preference similarity with the key item allows the model to avoid abrupt category transitions and produce more logical recommendation sequences. They mostly employed models like Matrix Factorization, Graph Convolutional Networks (GCNs), and Mask Pretraining Module [13].

D. Video Segmentation Models

These models use actions, objects, or scene changes to segment films into understandable chunks. Video footage can be better indexed and retrieved by using techniques like Topic Modelling Based Segmentation, Semantic Segmentation, and Visual Transition detection.

- Deep Reinforcement Learning uses deep reinforcement learning to generate excellent video summaries without the need for human annotations. It proposes an encoder-decoder design for a Deep Summarization Network (DSN). A Convolutional Neural Network (GoogLeNet pretrained on ImageNet) serves as the encoder, extracting visual features from video frames. Frame selection probabilities are generated by the decoder, which is a bidirectional Long Short-Term Memory (LSTM) network with a fully linked layer on top. This architecture facilitates end-to-end training and allows the model to recognize long-term dependencies in video frames [14].

- Segmentation Models with Per-frame Inference show that current approaches usually compromise at least one of these components: multi-frame methods that process temporal information introduce computational overhead that is inappropriate for mobile applications, while keyframe-based approaches (which process some frames completely and propagate results to others) produce unbalanced latency and cumulative errors. They offer a novel approach that introduces temporal consistency through specific training techniques while preserving per-frame inference, which processes each frame independently during testing. This method eliminates extra inference overhead and guarantees balanced processing time across frames [15].

- Semantic video segmentation is a three-step procedure that includes categorization, form identification, and object detection. They demonstrate how closely related these processes are to one another and how each stage's execution directly affects the others. Major issues in this subject are highlighted in the study, such as how angle, direction, scale, blurring, object camouflage, and overlapping objects can change an object's appearance. The authors stress how deep neural networks have lately transformed methods for semantic segmentation, despite the fact that similar problems have existed for more than 20 years. They recommend a number of models, including CNNs, SVM, Markov Random Fields, Random Decision Forest, and others [16].

E. Educational Content Generation

These models use AI and NLP techniques to produce interactive content, quizzes, and learning materials that are customized to the needs of the learners.

A. Google Gemini The study emphasizes how Gemini's improved knowledge sharing, automated evaluation, and tailored learning could revolutionize education. Several large language models (LLMs) and natural language processing (NLP) technologies are integrated into Gemini. It is shown as a system that makes use of underlying AI models. The technical details of Gemini's underlying models are not covered in detail in this work. Rather, it emphasizes the potential made possible by those models [17].

B. Networks of Attention-Based Encoder-Decoders In order to replicate the human process of choosing important shots for video summarization, the study aims to create an attention-based encoder-decoder network. By assigning varying relevance weights to distinct frames in the input video sequences of Gemini's underlying models, this system seeks to produce precise video summaries. Rather, it emphasizes the potential made possible by those models. It suggests a framework for video summarization called Attentive encoder-decoder networks, which is made up of two attention-based LSTM networks: Additive Attention Mechanism (A-AVS) and Multiplicative Attention Mechanism (M-AVS). The encoder is a Bidirectional Long Short-Term Memory (BiLSTM) network [18].

C. RNN model for generating questions In order to automatically produce quiz questions from instructional content, the study presents QGNet, a reader-generator system based on a recurrent neural network (RNN). The system's goal is to produce factual queries that are both fluid and pertinent. The model encourages the reader to concentrate on particular aspects of the input context by explicitly encoding answer information as an extra input. To produce questions that highlight particular passages in the supplied text, the question generator uses a pointer network [19].

Research Gap Identification

After analyzing prior works, a clear research gap is evident — while existing systems support personalization and adaptive learning, most depend on manual content curation or limited automation. There is a lack of a unified system capable of end-to-end automated course generation that dynamically retrieves, summarizes, and evaluates educational content.

Course Crafter fills this gap by integrating generative AI (Gemini API) with video-based content retrieval (YouTube Data API) to deliver a comprehensive, self-sustaining learning ecosystem. This unique combination of text-based intelligence and multimedia integration marks a significant contribution to the evolution of AI in education.

III. Proposed Model

The proposed system in Fig. 1 is an AI-powered educational recommendation engine designed to address the limitations of current video-based learning platforms. It transforms unstructured educational videos into structured, personalized learning journeys. The architecture leverages Natural Language Processing (NLP), deep learning, video analysis, and content generation techniques to enhance learner engagement, personalization, and overall learning effectiveness.

A. System Architecture

The architecture of *Course Crafter* is designed to ensure modularity, scalability, and efficient communication between various system components. The system follows a client-server model, where the front-end interacts with the back-end through API calls, and the back-end coordinates between the AI engines (Gemini and YouTube APIs) and the database. The overall design is divided into four key components — Frontend Interface, Backend Server, Intelligent APIs, and Database System.

1) Frontend:The front-end, developed using HTML, CSS, and JavaScript, provides an intuitive and visually appealing interface for user interaction. It includes pages for login, course creation, video viewing, summary display, quiz participation, and saved-course gallery. The responsive layout ensures accessibility across both desktop and mobile devices.
2) Backend:The backend is implemented using the Python Flask framework, which handles all logic related to authentication, API integration, and data storage. Flask provides lightweight yet powerful routing mechanisms to manage communication between the user interface and AI modules.
3) APIs:The Google Gemini API is responsible for generating course outlines, summarizing video content, and producing quizzes. The YouTube Data API v3 retrieves the most relevant videos corresponding to each topic generated by Gemini. Together, they form the "intelligence layer" of the system, enabling automatic content creation and selection.
4) Database:A PostgreSQL database is used to store user details, course outlines, video metadata, summaries, quiz questions, and progress history. This allows users to resume learning from where they left off, maintaining a persistent learning state.

Fig. 1 Course Crafter System Architecture

B. Workflow Pipeline

The system operates in a modular pipeline designed to guide the learner from content discovery to personalized learning and outcome evaluation:

1) User Input Collection: Learners specify their educational objectives, topics, time availability, and desired difficulty level.
2) Intent Interpretation: Advanced NLP models extract the user's goals, interpret context, and generate a semantic representation of the learning needs.
3) Content Retrieval: The system performs a targeted search for relevant videos from YouTube and other platforms using semantic filtering.
4) Metadata and Semantic Analysis: Retrieved content is analyzed for educational quality, based on both user engagement data and semantic relevance to the input.
5) Recommendation and Ranking: Content is recommended and ordered using a combination of similarity-based, content-aware, and knowledge-driven ranking models.
6) Video Segmentation: Long-form videos are divided into logical, topic-specific chapters, enhancing content digestibility and focus.
7) Content Enhancement: AI-generated summaries, notes, and quizzes are added to reinforce learning and promote active recall.
8) Learning Pathway Assembly: All resources are compiled into a coherent learning journey tailored to the learner's goals and existing knowledge.
9) Progress Monitoring and Feedback Loop: The system tracks learning activities and collects feedback to adapt future content and pathway recommendations.

This structured flow ensures a guided, engaging, and effective learning experience from discovery to knowledge reinforcement.

C. Functional Flow

The functional workflow of the system can be summarized as follows:

1) User Login / Registration: Users create an account or log in securely. Authentication ensures personalized tracking and progress management.
2) Topic Input: The learner enters a desired topic or keyword (e.g., "Learn Machine Learning").
3) Course Outline Generation (Gemini API): The system sends the topic to the Gemini API, which returns a structured, multi-level course outline with modules and subtopics.
4) Video Fetching (YouTube API): For each subtopic, the backend sends a search request to the YouTube API. It fetches 3–5 top educational videos filtered by relevance, language, and view count.
5) Summarization and Quiz Generation (Gemini API): For every selected video, the Gemini API generates a concise summary and quiz questions to assess comprehension.

6) Display and Interaction: The front-end displays the selected video alongside the summary and quiz. Learners can interactively switch between modules and track scores.
7) Progress Saving and Resume: The learner's completion status, quiz results, and course metadata are saved in the database. Users can return anytime and continue from the last completed video.

IV. METHODOLOGY

The methodology adopted for developing *Course Crafter* follows a modular, iterative, and phase-wise approach, ensuring each functional component is built, tested, and integrated systematically. The overall development strategy aligns with the Software Development Life Cycle (SDLC), particularly the iterative and incremental model, where each phase contributes to the final fully functional AI-based learning platform.

The system was developed in five main phases, from authentication to progress tracking:

A. Phase 1 – Login and Authentication

The first phase focuses on building a secure user management system that enables learners to create accounts, authenticate their credentials, and access personalized dashboards.

User registration includes collecting information such as name, email address, and password. For data security, passwords are hashed using cryptographic algorithms (SHA-256) before being stored in the database. This ensures that even if the database is compromised, raw credentials remain unreadable.

Once registered, users can log in using their credentials, which triggers a session-based authentication mechanism. Flask's session management system assigns a unique token to each user session, ensuring personalized access without repeated logins.

Each user has a personalized dashboard that stores their previously created courses, learning progress, and quiz scores. This personalization allows returning users to continue their studies seamlessly, supporting the system's adaptive learning principle.

B. Phase 2 – Topic Analysis and Course Outline Generation

The second phase is the foundation of the AI-driven component. It begins when a user enters a learning topic (for example, *"Learn Python Programming"*). This input is converted into a prompt and sent to the Google Gemini API for processing.
The Gemini model interprets the topic and generates a hierarchical course outline consisting of multiple levels of detail:
1) Modules (broader topics)
2) Subtopics (specific areas within each module)
3) Learning Objectives (clear, measurable outcomes for each subtopic).

This structured outline provides a clear learning pathway, ensuring that concepts are introduced progressively — from basic fundamentals to advanced applications.

Example prompt to Gemini API:
"Generate a beginner-friendly 5-module course outline for learning Python programming. Include module titles, subtopics, and key learning goals."

Example Output:

1) Module 1: Introduction to Python
   a. What is Python and why it is used
   b. Setting up the environment
   c. Basic syntax and data types
2) Module 2: Control Structures
   a. Conditional statements
   b. Loops and iterations

The Gemini-generated outline is returned in JSON format, which is then parsed and stored in the database for use in subsequent stages.

C. Phase 3 – Video Retrieval

The API fetches a list of videos that match the subtopic using parameters such as relevance, view count, and content duration. To ensure quality and educational value, videos are filtered using keyword heuristics such as *"tutorial," "lecture," "lesson,"* and *"explained."*

This process ensures that learners receive content that is educationally appropriate, concise, and popular within the learning community.

The algorithm gives preference to:

1) Educational category videos (videoCategoryId=27)
2) Duration under 15 minutes for better focus
3) Higher view-to-like ratios for credibility

The top 3–5 videos are stored for each topic, allowing learners to choose the most suitable option from within the Course Crafter interface.

D. Phase 4 – Summarization and Quiz Generation:

After video retrieval, the Gemini API is invoked again to enhance the learning material. This phase focuses on content summarization and assessment creation.

For each selected video, Gemini uses its NLP capabilities to analyze video metadata, titles, and descriptions to produce a concise, text-based summary. These summaries help learners quickly grasp the essence of each video before watching or revising afterward.

In addition, the system generates multiple-choice quizzes (MCQs) or short conceptual questions related to each topic. The quizzes serve as a formative assessment tool, allowing learners to test comprehension and reinforce knowledge.

For instance, after watching a Python loop tutorial, the system might automatically generate:

1) Question: Which Python statement is used to exit a loop prematurely?
a) Stop
b) Break
c) Exit
d) Skip
i) *Correct Answer:* (b) break

All generated summaries and quizzes are stored in the database and dynamically linked to their respective videos for quick retrieval.

E. Phase 5 – Storage and Progress Tracking

The final phase ensures that each learner's journey is tracked, saved, and restorable.

Whenever a learner completes a module or video, their progress percentage, quiz score, and completion status are recorded in the database. This enables a “Save and Resume Later” feature, allowing users to pick up from their last active session.

The data structure for this functionality includes the following key tables:

1) Users (id, name, email, password)
2) Courses (id, user_id, title)
3) Videos (id, course_id, title, url, summary, quiz)
4) Progress (user_id, course_id, video_id, status, score)

This architecture supports multi-user scalability, enabling several learners to use the platform simultaneously without data conflicts.

V. IMPLEMENTATION

The implementation stage combines both the technical development and integration of intelligent APIs to bring the Course Crafter platform to life. The process was guided by modular coding principles, security considerations, and user-centric design practices.

A. Front-End Design

The front end of *Course Crafter* serves as the visual and interactive layer where learners engage with the system. Developed using React and Tailwind CSS, it is designed for both responsiveness and simplicity.

Key features include:

1) Modular Cards: Each course or video is represented as a card, displaying title, thumbnail, and summary.
2) Video Player Section: Integrated YouTube player to view tutorials directly on the platform.
3) Side Panel: Displays AI-generated summaries and quizzes beside each video, encouraging interactive learning.
4) Progress Buttons: “Mark as Completed” and “Save to Gallery” options help maintain learning continuity.

The interface follows a clean design philosophy, ensuring learners can focus on content rather than navigation.

B. Backend Integration

The backend, powered by the Node JS, manages all communication between the user interface, AI APIs, and the database. Flask’s lightweight architecture enables rapid data processing and API handling through defined routes such as:

1) /create-course – handles topic input and Gemini outline generation
2) /fetch-video – fetches YouTube videos for each topic
3) /generate-quiz – retrieves Gemini-generated quiz and summary data

The system employs asynchronous request handling, ensuring that even while APIs process data, the user interface remains responsive. Proper error handling is implemented for API timeouts, key failures, or rate-limit exceedances.

C. API Usage Details

1) Gemini API (Google Generative Language Model)

a. Endpoint:https://generativelanguage.googleapis.com/v1beta/models/gemini-pro:generateContent

b. Functions: Generate hierarchical course outlines which produce concise summaries and quiz questions

c. Output Format: JSON (structured for parsing)

2) YouTube Data API v3

a. Endpoint:https://youtube.googleapis.com/youtube/v3/search

b. Functions: Retrieve educational video metadata using keywords and prompts.

This dual-API approach combines AI text generation and real-time educational video retrieval, making Course Crafter both intelligent and resourceful.

D. Security Measures

Given the system’s dependence on user data and third-party APIs, security is a primary concern.

1) Encrypted API Keys: Stored as environment variables to prevent exposure.
2) Password Hashing: All user passwords hashed before storage using secure algorithms.
3) HTTPS Protocol: All API communications encrypted to prevent data interception.
4) Session Tokens: Each logged-in user assigned a secure token to prevent unauthorized access.
5) Error Handling: The system automatically handles API key expiration, rate-limit errors, and invalid responses gracefully.

E. Database Schema

The database is the backbone of Course Crafter, responsible for persisting all user data, course content, and learning progress. The schema was designed using Prisma ORM, which provides a clear, relational structure that is both scalable and maintainable.

The schema is composed of two main parts: authentication models required by NextAuth.js and the core application models that define the course structure and content.:

1) Authentication and User Models:

a. User: The central model for learners, storing essential profile information like email and name

b. Account / Session: Standard models used by NextAuth.js to manage user sessions and link OAuth provider information (like Google) to a User account.

2) Course Structure Models:

a. Course: The top-level entity representing a full course. It contains the course name, a cover image, and links to its constituent units.

b. Unit: A logical grouping of chapters within a Course, acting as a module or section (e.g., "Module 1: Core Concepts").

c. Chapter: The most granular content piece, representing a single lesson. It stores the name, the videoId for the associated YouTube lesson, and the AI-generated summary.

3) Interactive Content and Progress Models:

a. Quiz: A collection of questions that can be associated with a Chapter, Unit (as a unit test), or the entire Course (as a final exam).

b. Question: Represents a single multiple-choice question, holding the question text, an array of options, and the correct answer.

c. User Progress: A simple but crucial table that links a User to a Chapter, tracking completion status with a completed boolean flag.

This structure ensures that data is well-organized, with clear relationships between users, the courses they create, and their progress through the educational content.

VI. Result

The *Course Crafter* project was successfully implemented and tested to evaluate its functionality, performance, and usability. The results confirm that the system effectively integrates artificial intelligence with multimedia content retrieval to provide a seamless learning experience. This section presents the outcomes of functionality verification, performance benchmarking, and user feedback analysis

A. Functional Output

Each module of *Course Crafter* was individually tested to ensure its intended functionality and integration within the system. The testing process followed a module-wise validation approach, where each core component (Login, Course Generation, Video Fetching, Summarization, and Progress Tracking) was independently verified.

Module	*Functionality*	*Status*
Login	*User authentication and session tracking*	Implemented
Course Generator	*Gemini API outline creation*	Implemented
Video Fetcher	*YouTube API integration*	Implemented
Summary + Quiz	*Gemini API content generation*	Implemented
Progress Storage	*Local database tracking*	Implemented

Table 1 Modules Implemented

The above results show that all core functionalities were successfully implemented and tested under real-use conditions. The APIs responded consistently to varied topic inputs, generating accurate course outlines and fetching relevant educational videos.

The generated quizzes were contextually aligned with video topics, demonstrating the Gemini API's strong ability to comprehend and summarize multimedia content. The *Progress Storage* module worked as intended, allowing users to save and resume learning sessions seamlessly.

B. Performmance Evaluation

To assess system performance, key metrics such as API response time, processing latency, and topic-video match accuracy were measured over multiple trials.:

1) *Average Gemini API response time: 2.8 seconds*.
2) Average YouTube API search latency: 1.4 seconds
3) End-to-end course creation time: Approximately 6 seconds.
4) Accuracy of topic-video match (manual evaluation): ~92%.

These results indicate that *Course Crafter* performs efficiently even under real-time use. The combined API processing time remains under seven seconds, which is acceptable for dynamic course generation.

The 92% accuracy rate for topic-to-video matching demonstrates that the heuristic-based filtering and keyword optimization process successfully eliminates most irrelevant results. Only a small fraction of videos were found to be slightly off-topic, primarily due to ambiguous search keywords or metadata inconsistencies.

Additionally, system stress testing confirmed that multiple users could create courses concurrently without performance degradation, validating the scalability of the Flask backend and database management.

C. User Feedback Analysis

To assess usability and learner satisfaction, a pilot test was conducted among 15 students from different academic backgrounds. Participants were asked to use *Course Crafter* to generate a course of their choice and evaluate their learning experience

1) Quantitative Results:

a. 87% of users agreed that the videos fetched were *"highly relevant and educational."*

b. 80% reported that AI-generated summaries were *concise and helpful* for quick review.

c. 73% preferred *AI-generated quizzes* over traditional manually created ones due to variety and instant feedback.

2) Qualitative Observations:

a. Users appreciated the time efficiency of the system compared to searching for tutorials manually.

b. The summaries helped reduce video-watching time while still understanding the main ideas.

c. The quiz feature motivated users to engage more actively with the learning material.

d. Some users suggested adding multi-language support and more visual quiz formats.

Overall, the pilot study demonstrated a positive user perception, confirming the effectiveness and practicality of Course Crafter in real-world educational contexts.

VII. Discussion

The integration of Google Gemini API and YouTube Data API has successfully automated the entire digital learning pipeline — from course creation to content curation, summarization, and assessment. This section discusses the system's significance, advantages, limitations, and comparative analysis with conventional learning methods.

A. Advantages of the System

The implementation of *Course Crafter* demonstrates several advantages over traditional and semi-automated e-learning systems:

1) Personalization and Adaptivity: Each learner receives a course specifically designed for their chosen topic. The outline, video recommendations, and quizzes are all dynamically generated, ensuring a unique and personalized experience.
2) Time Efficiency:Learners no longer need to browse through countless videos or tutorials. The system automatically retrieves and organizes the best educational content, saving substantial time.
3) AI-Powered Contextual Learning:The integration of Gemini ensures that learners not only access videos but also receive meaningful, summarized insights. This reduces cognitive load and helps in faster concept understanding.

4) Interactive and Self-Paced Learning:Through built-in quizzes and progress tracking, learners are encouraged to interact with the content, assess their knowledge, and resume later without losing context.
5) Continuous Knowledge Evaluation:Automatically generated quizzes allow for real-time performance assessment, reinforcing learning outcomes after each module.
6) Content Reliability:Since the system sources material directly from verified YouTube educational channels, learners gain access to up-to-date and authentic tutorials.
7) *Scalability and Reusability:*The modular architecture allows future integration of other AI models, voice-based input, and analytics, making the system easily extendable for institutional use.

B. Limitations

While *Course Crafter* demonstrates promising results, certain challenges and limitations were identified during testing:

1) API Rate Limits:Both Gemini and YouTube APIs impose usage restrictions, which limit the number of queries that can be processed per minute. Large-scale usage may require quota expansion or caching mechanisms.
2) Dependence on Video Metadata Quality:The accuracy of summaries and quiz generation depends largely on the quality of video descriptions and transcripts. Poorly annotated videos may result in less accurate AI outputs.
3) Language Constraints:At present, the system functions optimally in English. For non-English topics, summaries and quizzes may not maintain the same accuracy.
4) No Plagiarism or Fact-Verification Module:The system currently assumes that Gemini's generated content is factual. A verification layer could be added in future iterations.
5) Limited Multimedia Support:The platform primarily supports YouTube videos; integration of other platforms like Vimeo or Coursera could further enhance diversity**.**

Despite these limitations, the current implementation provides a strong and functional foundation for an intelligent course generation framework.

C. Comparative Analysis

To evaluate the distinctiveness of *Course Crafter*, it was compared against traditional e-learning platforms in terms of automation, interactivity, and adaptivity.

Feature	***Traditional E-Learning***	***Course Crafter***
Manual video search	✓*Yes*	✗ *No (Automated via YouTube API)*
AI summary generation	✗*No*	✓ *Yes (Gemini API)*
Quiz automation	*Limited or manual*	*Fully automated*
Progress save/resume	*Partial*	*Fully integrated*
Adaptivity & personalization	*Low*	*High*
Learning path creation	*Manual*	*AI-generated*
Multimedia integration	*Fixed content*	*Dynamic & topic-based*

Table 2 Comparative Analysis

This comparison highlights how *Course Crafter* provides a next-generation learning solution by automating all major educational tasks, significantly enhancing personalization and efficiency compared to existing platforms.

VIII. CONCLUSION

The successful implementation of *Course Crafter* demonstrates the immense potential of artificial intelligence in revolutionizing the field of e-learning. By leveraging the power of Google Gemini API for natural language processing and YouTube Data API for multimedia content retrieval, the system bridges the gap between unstructured online content and structured, personalized education.

The platform automates the entire course creation workflow — from generating topic outlines to summarizing lessons and creating quizzes — offering a seamless, intelligent, and adaptive learning experience.
Key achievements of the implementation include:

1) Automatic generation of personalized courses: AI dynamically builds complete syllabi tailored to user input.
2) Intelligent video curation: The system identifies high-quality educational videos from YouTube, ensuring relevant and up-to-date learning materials.
3) Real-time summaries and quizzes: Gemini AI produces topic summaries and quizzes that help learners review and test their understanding instantly.
4) Progress persistence and resume functionality: The integrated database tracks user activity, allowing learning continuity without manual bookmarking.

By combining automation, personalization, and interactive assessment, *Course Crafter* offers a compelling demonstration of how AI can enhance accessibility, efficiency, and engagement in digital learning environments.

In essence, this research contributes to the growing movement toward AI-assisted education, where technology not only delivers information but also structures, explains, and evaluates knowledge autonomously. As the system evolves with upcoming features like multilingual support, gamification, and mobile integration, it holds the potential to redefine the global online learning experience — making education more inclusive, dynamic, and intelligent.

REFERENCES

[1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota.

[2] YinhanLiu, MyleOtt, NamanGoyal, JingfeiDu, MandarJoshi, DanqiChen, OmerLevy, MikeLewis, Luke Zettlemoyer, Veselin Stoyanov, "RoBERTa: ARobustly Optimized BERT Pretraining Approach," *Computation and Language,July 2019.*

[3] Nils Reimers and Iryna Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3982–3992, Hong Kong, China.M.

[4] Hiroshi Noji, Hiroya Takamura, "An Analysis of the Utility of Explicit Negative Examples to Improve the Syntactic Abilities of Neural Language Models," *In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 3375–3385, Online. Association for Computational Linguistics,2020.*

[5] M. Shubha, K. Kapoor, M. Shrutiya and H. R. Mamatha, "Searching a Video Database using Natural Language Queries," *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, 2021, pp. 190-196

[6] A. Ekin, A. M. Tekalp and R. Mehrotra, "Integrated semantic-syntactic video modeling for search and browsing," in *IEEE Transactions on Multimedia*, vol. 6, no. 6, pp. 839-851, Dec. 2004

[7] Raghad Obeidat, Rehab Duwairi, and Ahmad Al-Aiad, "A Collaborative Recommendation System for Online Course Selection Using Data Mining Techniques," *International Conference on Deep Learning and Machine Learning in Emerging Applications (DLML), IEEE, 2024.*

[8] Zijin Luo, Xu Wang, Yiquan Wang, Haotian Zhang, and Zhuangzhuang Li, "A Personalized MOOC Learning Group and Course Recommendation Method Based on Graph Neural Network and Social Network Analysis," *Preprint, 2024*.

[9] Mohd Suffian Sulaiman, Amylia Ahamad Tamizi, Mohd Razif Shamsudin, and Azri Azmi, "Course Recommendation System Using Fuzzy Logic Approach," *Indonesian Journal of Electrical Engineering and Computer Science, 2020.*

[10] Paul Covington, Jay Adams, and Emre Sargin, "Deep Neural Networks for YouTube Recommendations," *In Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16). Association for Computing Machinery, New York, NY, USA,2016*.

[11] Bal´azsHidasi, Alexandros Karatzoglou, LinasBaltrunas, DomonkosTikk, "SESSION-BASED RECOMMENDATIONS WITH RECURRENT NEURAL NETWORKS," *Published as a conference paper at ICLR 2016*.

[12] XiangnanHe, LiziLiao, Hanwang Zhang,Liqiang Nie,XiaHu, Tat-SengChua, "Neural Collaborative Filtering," *Information Retrieval, 2017*.

[13] Han, P., Zhou, S., Yu, J. *et al.* "Personalized Re-ranking for Recommendation with Mask Pretraining.," *Data Sci. Eng.* **8**, 357–367 (2023).

[14] Xiaocong Chen, Lina Yao, Julian McAuley, Guanglin Zhou, and Xianzhi Wang, "A Survey of Deep Reinforcement Learning in Recommender Systems: A Systematic Review and Future Directions," *Journal of the ACM (J. ACM), 2021*.

[15] Yifan Liu, Chunhua Shen, Changqian Yu, Jingdong Wang, "Efficient Video Segmentation Models with Per-frame Inference," *Computer Vision and Pattern Recognition. Feb 2022.*

[16] Mohammad Hajizadeh Saffar, Mohsen Fayyaz, Mohammad Sabokrou, Mahmood Fathy, "Semantic Video Segmentation: A Review on Recent Approache," *Computer Vision and Pattern Recognition,2018.*

[17] Imran, M., Almusharraf, N, "Google Gemini as a next generation AI educational tool: a review of emerging educational technology," *Smart Learn. Environ.* 11, 22 (2024).

[18] Z. Ji, K. Xiong, Y. Pang and X. Li, "Video Summarization With Attention-Based Encoder–Decoder Networks," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 6, pp. 1709-1717, June 2020

[19] Kaiyang Zhou, Yu Qiao, and Tao Xiang, "Deep reinforcement learning for unsupervised video summarization with diversity-representativeness reward," *In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence* (AAAI'18/IAAI'18/EAAI'18). AAAI Press, Article 929, 7582–7589, 2018.

[20] W. Xie and P. Pu, "A Course Recommendation System based on User Intents and Resource Alignment," *2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*, Nanjing, China, 2024, pp. 545-548

[21] W. Xie and P. Pu, "A Course Recommendation System based on User Intents and Resource Alignment," *2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*, Nanjing, China, 2024, pp. 545-548

[22] Zhong, Mingxia & Ding, Rongtao, "Design of a Personalized Recommendation System for Learning Resources based on Collaborative Filtering," *International Journal of Circuits, Systems and Signal Processing. 16. 122-131. 10.46300/9106.2022.16.16.2020*.

[23] S. Asadi , S. M. B. Jafari and Z. Shokrollahi, "Developing a Course Recommender by Combining Clustering and Fuzzy Association Rules," *Journal of AI and Data Mining, 7 2 (2019): 249-262*.

[24] Shaikh, M. S., Nair, R., Shaikh, M. Z., Shaikh, A., & Ravale, U, "Machine Learning based MOOC Course Recommendation System," *International Journal of Science and Social Science Research, 1(3), 260–264.2020.*

[25] VikramSingh and Balwinder Saini, "AN EFFECTIVE TOKENIZATION ALGORITHM FOR INFORMATION RETRIEVAL SYSTEMS," *Department of Computer Engineering, National Institute of Technology, Kurukshetra, Haryana, India, 2014.*

[26] ChaitaliDiwan,, Srinath Srinivasa, GandharvSuri, Saksham Agarwal, PrasadRam, "AI-based learning content generation and learning pathway augmentation to increase learner engagement," *Computers and Education. Artificial Intelligence,Volume 4,2023.*

[27] QiZhang,Zijian Yang,Yilun Huang,Jiarong He,Lixiang Wang, "AnEffective Way for Cross-Market Recommendation with Hybrid Pre-Ranking and Ranking Models," *In Proceedings of Proceedings of the ACM Interna tional WSDM Conference (WSDM'22). ACM, New York, NY, USA,2022.*

[28] D. G. M., R. H. Goudar, A. A. Kulkarni, V. N. Rathod and G. S. Hukkeri, "A Digital Recommendation System for Personalized Learning to Enhance Online Education: A Review," *in IEEE Access, vol. 12, pp. 34019-34041*.