International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

SmartElect: An AI-Based Personalized Elective Course Recommendation System Using TF-IDF and Cosine Similarity

Jafrin Sultana[1], Sneha Popat Mahadik[2], Vishakha Hanmant Sutar[3], Prof.Madhura M. Raste[4]

[1,2,3]Students, Department of Computer Science Engineering - Artificial Intelligence & Data Science, Padmabhooshan Vasantraodada Patil Institute of Technology, Sangli, Maharashtra, India
[4]Assistant Professor, Department of Computer Science Engineering - Artificial Intelligence & Data Science, Padmabhooshan Vasantraodada Patil Institute of Technology, Sangli, Maharashtra, India

ABSTRACT

The rapid growth of online learning platforms and digital education systems has significantly expanded access to elective courses across diverse academic and professional domains. Platforms such as Coursera, Udemy, and edX offer a wide range of courses, enabling learners to enhance their knowledge beyond traditional classroom boundaries. However, the large volume of available options creates a challenge for students in selecting courses that best align with their skills, interests, and long- term career goals. Conventional methods, such as manual browsing and generic recommendation approaches, often lack personalization and adaptability, resulting in inefficient decision-making.

To address these challenges, this paper proposes SmartElect, an AI-powered elective course recommendation system designed to provide personalized and relevant course suggestions using machine learning techniques. The system follows a structured pipeline including dataset collection, data cleaning, text preprocessing, and exploratory data analysis (EDA). TF-IDF (Term Frequency– Inverse Document Frequency) vectorization is applied to convert textual data into numerical representations.

The recommendation engine uses a content-based filtering approach, where cosine similarity measures the relevance between user input and course descriptions. The model is evaluated using Precision, Recall, and Normalized Discounted Cumulative Gain (NDCG), demonstrating effective and accurate recommendations. The system is implemented using a React frontend and Flask backend, along with an AI-based chatbot and career roadmap feature to enhance user experience.

The system effectively reduces decision-making complexity and improves course selection accuracy for students.

KEYWORDS: Machine Learning, Term Frequency–Inverse Document Frequency, Cosine Similarity, Content-Based Filtering, Course Recommendation System

INTRODUCTION

The emergence of digital learning platforms has revolutionized the way education is accessed and delivered. Modern platforms provide a wide variety of courses across multiple domains, allowing

learners to develop skills at their own pace. Despite these advantages, the availability of numerous options often makes it difficult for students to identify courses that best suit their individual profiles. As a result, many learners rely on random selection methods, peer influence, or incomplete information, which can lead to ineffective learning decisions.

The integration of Artificial Intelligence (AI) and Machine Learning (ML) in education has opened new possibilities for addressing such challenges. In particular, recommendation systems have shown great potential in delivering personalized suggestions by analyzing user preferences and content features. Applying these techniques to course selection can significantly improve the relevance and efficiency of recommendations.

In this work, SmartElect is introduced as an intelligent elective course recommendation system designed to assist students in making informed choices. The system utilizes a structured data processing pipeline, including data preparation, text cleaning, exploratory analysis, and feature extraction. Course descriptions are converted into vector representations using TF-IDF, enabling effective comparison with user inputs. Similarity between user profiles and course data is computed using cosine similarity, forming the basis for generating ranked recommendations.

To validate the model, the dataset is divided using a train-test split (test size = 0.2), and performance is evaluated through metrics such as Precision, Recall, and Normalized Discounted Cumulative Gain (NDCG). The system is implemented using a Flask backend and a React frontend, supported by a MySQL database for efficient data handling. Additionally, an intent-based chatbot is incorporated to provide focused assistance for course-related queries.

The proposed approach aims to simplify the elective selection process by offering accurate and personalized recommendations, thereby enhancing user satisfaction and supporting better academic and career decision.

LITERATURE REVIEW-

The rapid growth of online learning platforms has increased the demand for intelligent systems that can recommend relevant courses to learners based on their interests and goals. Artificial Intelligence (AI) and Machine Learning (ML) techniques have played a significant role in enhancing personalized recommendation systems in the education domain.

Ch. Suneetha et al. [1] proposed a machine learning-based course recommendation system that utilizes user preferences and historical data to suggest relevant courses. Their approach improves recommendation accuracy using classification techniques; however, it lacks advanced Natural Language Processing (NLP) methods for analyzing course content semantically.

Anushifa et al. [2] developed an AI-powered personalized course recommender system using TF-IDF, cosine similarity, and K-Nearest Neighbors (KNN). This approach effectively matches user interests with course descriptions and provides accurate recommendations. Despite its effectiveness, the system does not fully address dynamic user behavior or real-time recommendation updates.

Adarsh et al. [3] introduced a hybrid recommendation system that integrates Generative AI with rule-based filtering. The use of GPT-based techniques enhances personalization and helps overcome common issues such as the cold-start problem. However, the system requires significant computational resources and involves complex implementation.

Rithica and Sai Srija [4] presented an AI-driven personalized learning recommendation system aimed at improving student academic performance. Their system uses data-driven techniques for recommending

courses but relies heavily on structured datasets and lacks advanced semantic analysis capabilities.
Oubalahcen et al. [5] conducted a comprehensive survey on AI-based e-learning recommender systems, discussing various techniques such as collaborative filtering, content-based filtering, and hybrid models. The study highlights the importance of AI in improving recommendation systems while identifying challenges such as scalability, sparsity of data, and adaptability.

PROBLEM STATEMENT-

Selecting an appropriate elective course is a critical factor influencing a student's academic progression and future career opportunities. However, this decision-making process remains challenging due to several limitations in the current educational ecosystem. The rapid growth of online learning platforms has resulted in an extensive range of available courses, making it difficult for learners to identify options that are most relevant to their individual needs.

A major issue is the lack of structured and personalized guidance during course selection. Students often depend on informal factors such as peer recommendations, popular trends, or basic course ratings instead of evaluating their own skills, interests, and learning objectives. This unstructured approach frequently leads to suboptimal choices that do not align with their capabilities.

Moreover, existing recommendation systems typically provide generalized suggestions based on popularity or overall ratings, without adequately considering individual learner profiles. Such approaches fail to address important factors like prior knowledge, skill gaps, and personal preferences, resulting in recommendations that may be unsuitable in terms of difficulty level or relevance.

Another significant concern is the disconnect between selected courses and long-term career goals. Without proper analytical support, students may struggle to understand the practical value of a course in relation to their future aspirations. This mismatch can reduce engagement, lower satisfaction levels, and lead to inefficient utilization of time and learning resources.

Therefore, there is a need for an intelligent and personalized recommendation system that can effectively analyze user profiles and provide relevant elective course suggestions, thereby improving decision-making and enhancing learning outcomes.

PROPOSED SYSTEM-

The proposed system is an AI powered elective course recommendation platform designed to assist students in selecting suitable courses and career paths based on their skills and interests. By capturing user input, the system generates personalized recommendations that align with current industry requirements and learning trends.

The core of the system is a content-based recommendation engine that utilizes TF-IDF vectorization and cosine similarity to analyze course data and match it with user preferences. This approach enables the system to identify semantic relationships between user input and course content, ensuring accurate and relevant suggestions.

In addition to course recommendations, the system incorporates a roadmap feature that provides structured, step-by-step guidance for achieving specific career goals. This helps users understand the progression of skills from beginner to advanced levels, enhancing their learning experience. Furthermore, an intent-based AI chatbot is integrated to handle course-related queries, improving user interaction and providing instant support.

The system follows a modular architecture comprising a React.js-based frontend for interactive user

experience, a Flask backend for handling application logic and API communication, and a MySQL database for efficient data storage and retrieval. Python libraries are used to implement machine learning and text processing functionalities.

To evaluate system performance, metrics such as Precision, Recall, and NDCG are used, along with user feedback to measure usability and effectiveness. The overall objective of the system is to simplify the course selection process, reduce decision-making complexity, and provide personalized, data- driven guidance for academic and career planning.

METHODOLOGY

The SmartElect system follows a structured methodology to generate personalized course recommendations based on user skills and interests. It uses a client-server architecture where the React frontend collects user inputs and displays results, while the Flask backend processes data and performs recommendation logic. A MySQL database is used to store course and user data.

The dataset consists of course information collected from online learning platforms such as Coursera and Udemy, including course titles, descriptions, and skills covered. The dataset was preprocessed and used for training and evaluation.

The dataset is collected from online sources and preprocessed using techniques such as lowercasing, stop-word removal, and tokenization. TF-IDF vectorization converts textual data into numerical form. Cosine similarity is then applied to compute similarity scores between user input and course descriptions, and the top-ranked courses are recommended.

The dataset is split into training and testing sets (80:20) to evaluate performance. The system also includes a roadmap feature and chatbot for better interaction. Performance is evaluated using Precision, Recall, and NDCG metrics.

IMPLEMENTATION-

The system is implemented using a combination of web technologies and machine learning libraries to ensure efficiency and scalability. The frontend is developed using React, HTML, and CSS, providing an interactive interface with modules for authentication, dashboard, recommendations, roadmap visualization, and chatbot interaction. Communication with the backend is handled through API requests.

The backend is built using Flask, which manages request processing, integrates the recommendation model, and handles communication with the database. The course dataset, sourced from platforms like Coursera, is stored in CSV format and processed using the Pandas library.

Data preprocessing involves text cleaning, tokenization, and normalization to prepare the dataset for analysis. TF-IDF vectorization is applied to convert textual data into numerical features, ensuring consistent representation of both course data and user inputs. The recommendation engine uses cosine similarity to rank courses based on relevance and returns the top-N results.

Additional features include a roadmap module for structured learning paths and an intent-based chatbot for course-related assistance. MySQL is used for secure storage of user data and course information, ensuring reliable data management.

SYSTEM ARCHITECTURE-

The system follows a layered architecture to ensure efficient data flow, scalability, and modular design. It consists of three main layers: the user interface layer, the processing layer, and the data management

layer. The overall system architecture is shown in Figure 1.

The user interface layer is developed using React, HTML, and CSS, providing an interactive platform for users to input their skills, interests, and preferences. This layer captures user input and sends it to the backend through API requests.

The processing layer, implemented using Flask, handles core functionalities of the system. It performs data preprocessing, including text cleaning and normalization, followed by feature extraction using TF-IDF vectorization. The recommendation engine then applies cosine similarity to compare user input with course data and generates ranked course recommendations.

The data management layer uses MySQL to store and manage course information and user data. It supports efficient data retrieval required for generating recommendations.

Communication between layers is achieved through RESTful APIs, ensuring smooth interaction between frontend, backend, and database. The architecture is scalable and allows integration of additional features such as chatbot assistance and roadmap generation.

Overall, the system architecture provides a structured and efficient framework for delivering personalized course recommendations.

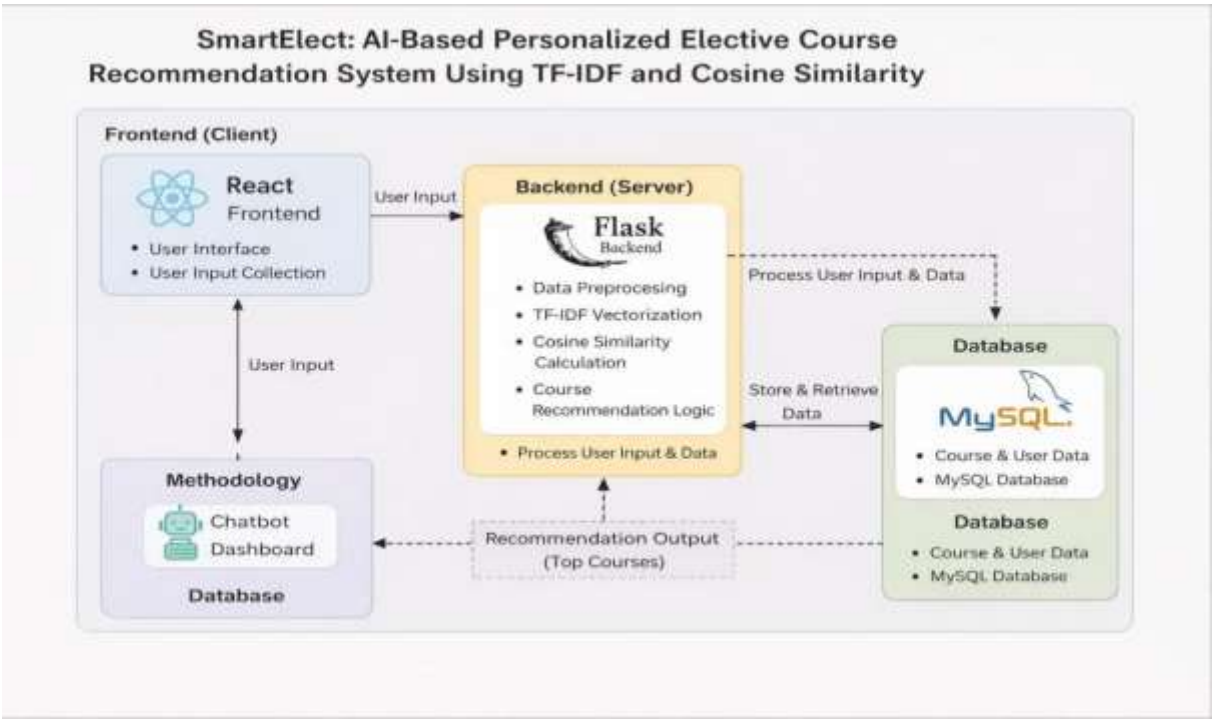

Figure 1: System Architecture of SmartElect

RESULTS AND DISCUSSION

The SmartElect system effectively generates relevant course recommendations using TF-IDF and cosine similarity. It successfully maps user-provided skills and interests to suitable courses, demonstrating good semantic understanding.

The system achieved:

- Precision: 0.1766
- Recall: 0.6863
- NDCG: 0.6939

These results indicate strong recall and effective ranking performance. The roadmap and chatbot features

enhance usability and interaction. However, the system may recommend similar courses due to the nature of content-based filtering.

Method	Precision	Recall	NDCG	Remark
Traditional(ManualSelection)	0.10	0.40	0.45	Low personalization
Basic-Recommendation System	0.14	0.55	0.60	Moderate performance
Proposed SmartElect System	0.1766	0.6863	0.6939	relevance and ranking

Table 1: Comparison Table

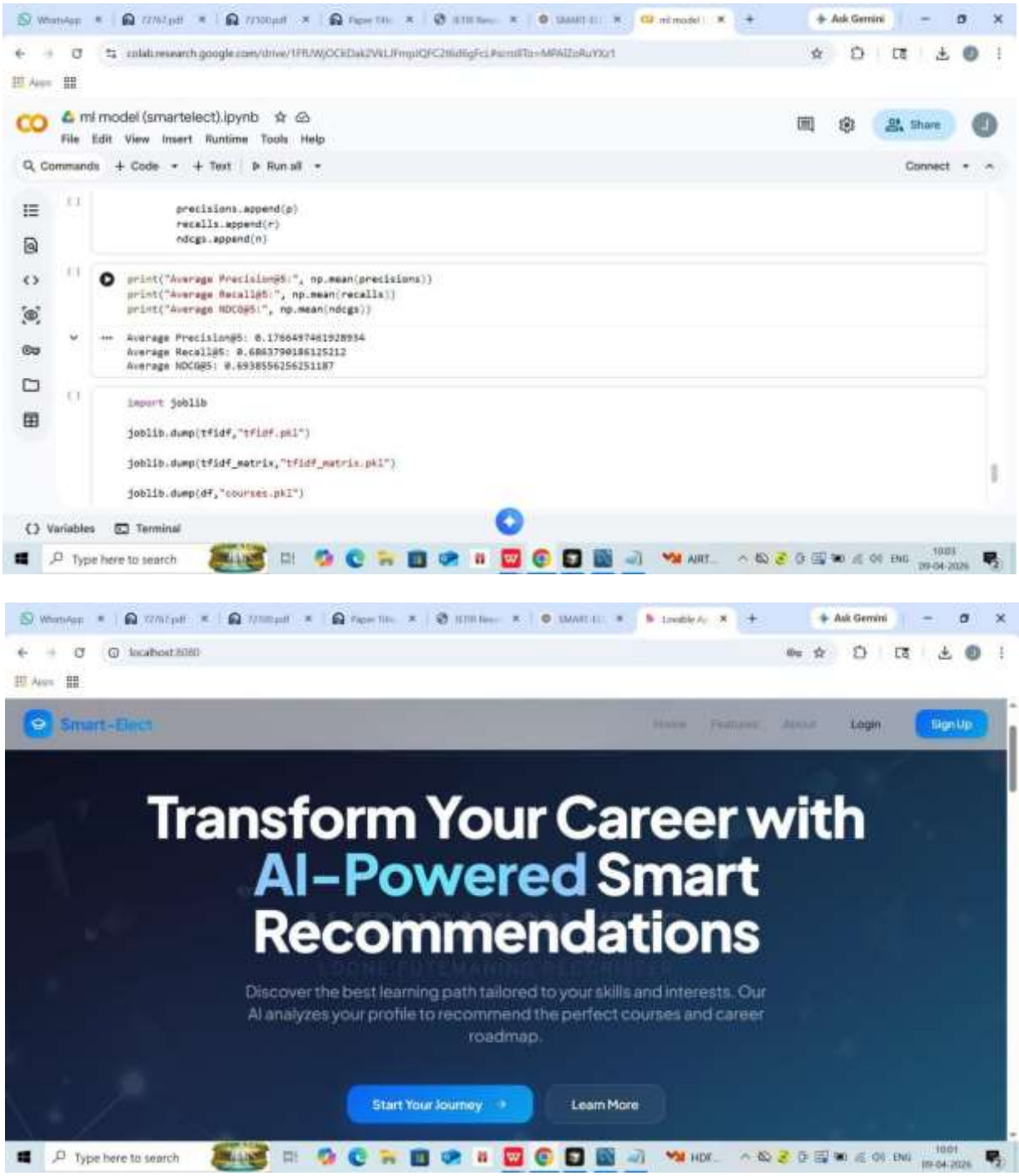

Figure 2: Result of SmartElect

ADVANTAGES

The SmartElect system provides highly personalized course recommendations based on user skills and interests, ensuring better relevance compared to generic systems. It effectively handles the cold-start problem, as it does not require prior user data.

The use of TF-IDF and cosine similarity ensures efficient processing and scalability. The roadmap feature provides structured learning paths, while the chatbot enhances user interaction, making the system practical and user-friendly.

FUTURE WORK

The SmartElect system can be improved by integrating hybrid recommendation techniques to increase accuracy. Advanced NLP models such as BERT can enhance understanding of user input and course descriptions.

Real-time data integration from learning platforms can keep recommendations updated. The chatbot can be enhanced with advanced conversational AI. A feedback system can be added to improve recommendations over time. Cloud deployment and mobile application development can improve scalability and accessibility.

CONCLUSION

The SmartElect system successfully addresses the challenges of elective course selection by providing personalized and data-driven recommendations. Using TF-IDF and cosine similarity, the system effectively matches user input with relevant courses.

The system overcomes the cold-start problem and includes additional features such as a roadmap and chatbot to improve usability. Although it depends on dataset quality, the overall performance is effective and scalable.

SmartElect provides a reliable and intelligent solution for course recommendation, helping students make informed academic and career decisions.

The novelty of the proposed system lies in the integration of content-based recommendation with a structured career roadmap and an AI-based chatbot. Unlike traditional systems, it provides real-time personalized suggestions without relying on prior user history, effectively addressing the cold-start problem.

REFERENCES

1. Ch. Suneetha, M. Jahnavi, M. Yashwanth, and N. K. Goutam, “Course Recommendation System Using Machine Learning,” *International Journal of Scientific Development and Research (IJSDR)*, vol. 8, no. 6, June 2023, ISSN: 2455-2631.
2. Anushifa J., Akshaya Lakshmi R., and Ashvini Uma A., “AI-Powered Personalized Course Recommender,” *International Journal of Innovative Research in Technology (IJIRT)*, vol. 11, no. 11, April 2025, ISSN: 2349-6002.
3. Adarsh S., Yeshas H. Rao, Rishikesh R., Tejo Sai Yaswanth Kemisetti, and Harshit Ranjan, “Elective Recommendation System Using Generative AI: A Hybrid Approach for Personalized Course Selection,” *TIJER – International Research Journal*, vol. 12, no. 1, January 2025, ISSN: 2349- 9249.
4. Rithica B. and Sai Srija R., “An AI-Driven Personalized Learning Recommendation System for

Enhancing Student Academic Performance," *International Journal of Scientific Research & Engineering Trends*, vol. 11, no. 6, Nov–Dec 2025, ISSN: 2395-566X.

5. Houda Oubalahcen, Lahcen Tamym, and Moulay Driss El Ouadghiri, "The Use of AI in E- Learning Recommender Systems: A Comprehensive Survey," *Proceedings of the 10th International Symposium on Emerging Inter-networks, Communication and Mobility*, August 2023.