PIP2001 Capstone Project
Final

Cosine Similarity Collaborative Filtering Elective Recommendation System

Final Viva-voce

Batch Number: ISR-G04

Roll Number	Student Name
20211ISR0027	**P.DURGA MARUTHI VARA PRASAD (group leader)**
20211ISR0005	ALLEN SAJI
20211ISR0091	FAIZAN NIYAZUDDIN
20211ISR0015	SYED YUSUF HUSSAIN
20211ISR0065	S NAYMAAN KHAN

Under the Supervision of,

Dr.Swati Sharma

School of Computer Science and Engineering

Presidency University

Name of the Program: Information Science And Engineering AI & Robotics

Name of the HoD: Dr.Zafar Ali Khan N

Name of the Program Project Coordinator: Dr.Afroz Basha

Name of the School Project Coordinators: Dr. Sampath A K / Dr. Abdul Khadar A / Mr. Md Ziaur Rahman

Content

- **Github Link**
- **Abstract**
- **Literature review**
- **Objective**
- **Existing method drawback**
- **Proposed method**
- **Architecture Diagram**
- **modules**
- **Hardware and Software Details**
- **Timeline by Gantt Chart**
- **References**

Abstract | Problem Statement Number: 192

Efficiently managing elective course enrollments in academic institutions poses a significant challenge, particularly when aiming to provide personalized course recommendations that align with students' academic strengths. This project addresses the need for a dynamic course recommendation system that utilizes collaborative filtering to process a dataset of 10,600 student records across 24 unique courses. The dataset features include student marks, grades, and grade points, which form the foundation for generating personalized course recommendations through a cosine similarity-based algorithm.

The system is built on a robust technical stack comprising Python, Streamlit, SQLite3, and Scikit-learn, optimized for real-time seat allocation management. It supports three distinct user interfaces catering to students, faculty, and Heads of Department (HODs), each providing tailored functionality for role-specific tasks. The project's architecture is modular, with five core Python files, ensuring ease of maintenance and scalability. The use of Streamlit for the user interfaces enables interactive and intuitive access to the system's features, while SQLite3 guarantees thread-safe database operations crucial for managing concurrent enrollments. The integration of Plotly provides real-time visualization dashboards that help stakeholders make informed decisions.

Results from the implementation demonstrate a significant enhancement in elective course enrollments, providing students with personalized recommendations that improve their academic planning. The system ensures accurate seat allocation management, mitigating over-capacity risks and enabling HODs to strategically adjust seat availability. Overall, the project showcases an effective blend of machine learning and database management to facilitate data-driven decision-making, resulting in improved operational efficiency and increased student satisfaction. This application serves as a viable model for academic institutions seeking to streamline their enrollment processes and provide enhanced advisory capabilities through technology-driven solutions.

Github Link

The Github link provided should have public access permission.

Github Link:
created by the group leader :

https://github.com/durgamaruthi/COSINE_SIMILARITY_ML_RECOMMENDATION_SYSTEM.git

Literature Review

Elective course recommendation systems have evolved significantly, with various techniques such as collaborative filtering, content-based methods, and hybrid approaches gaining prominence. **Collaborative filtering** uses the preferences of similar users to make recommendations, and it has been widely applied in educational domains to suggest courses that match students' academic history and preferences (Ray & Sharma, 2011; Kiratijuta Bhumichitr et al., 2017). One common technique, the **Alternating Least Square (ALS) algorithm**, has shown impressive accuracy, with some studies achieving over 85% in predicting course performance (Kiratijuta Bhumichitr et al., 2017).

However, to address limitations such as the cold-start problem and popularity bias, **hybrid models** that combine collaborative filtering with other approaches, such as **content-based filtering** or **natural language processing (NLP)**, have been explored (Neha Bhagwan Samrit, 2017; Prof. Naren.J et al., 2019). These models leverage textual analysis to convert human-readable course descriptions into machine-readable data, thereby improving recommendation relevance.

Literature Review

In our project, **cosine similarity** was chosen as the core technique for course recommendations due to its strength in **measuring similarity between student profiles and course attributes**. By calculating the cosine of the angle between two vectors (e.g., student performance metrics and course requirements), cosine similarity identifies highly relevant course recommendations even in cases with sparse data, thus helping mitigate some limitations of traditional collaborative filtering (C. Suneetha et al., 2023). Moreover, this technique's ability to handle high-dimensional data spaces makes it well-suited for educational data, where each student's performance can vary significantly across multiple dimensions.

Literature Review

In recent works, researchers have also highlighted the potential of **deep learning models** and **neural collaborative filtering**, which improve upon classical methods by capturing complex, non-linear relationships in student data (Vanchinathan et al., 2020). These advanced techniques allow for more personalized recommendations by learning deeper correlations between student behavior, course content, and academic success.

The recommendation system proposed in this project integrates these findings, leveraging cosine similarity for initial filtering and more sophisticated machine learning algorithms for dynamic, real-time course suggestions

Objectives

- Develop an ML course recommendation system for elective subjects.
- Implement collaborative filtering using cosine similarity for personalized recommendations.
- Create user-friendly interfaces for students and HODs.
- Provide analytics to track the performance of the recommendation algorithm.
- Ensure data privacy and prevent multiple enrollments.

Existing method Drawback

Cold-Start Problem: Traditional collaborative filtering systems require sufficient historical data to make accurate predictions, leading to poor recommendations for new students or new courses.

Popularity Bias: Models often over-recommend popular courses, neglecting lesser-known options that may be more suited to individual student preferences.

Limited Real-Time Adaptability: Many models are static, failing to update dynamically based on student behavior or course availability.

Computational Complexity: Hybrid systems, while improving recommendation accuracy, are often computationally expensive and harder to interpret, especially in large-scale educational settings.

Data Privacy Concerns: Heavy reliance on personal academic data raises concerns about security and privacy, particularly in the educational domain.

Proposed Method

The proposed **Elective Recommendation System** will solve these problems by offering a recommendation engine for both **students** and **HoDs**. For students, the system will recommend electives based on their academic profiles, preferences, and departmental needs. For HoDs, the system will monitor elective enrollments in real time, ensuring a balanced distribution without manual intervention. If certain electives become overcrowded, the system will dynamically recommend alternatives to students, ensuring optimal distribution.

Proposed Method

The **Elective Recommendation System** must solve two interconnected problems:

Student Elective Choice:

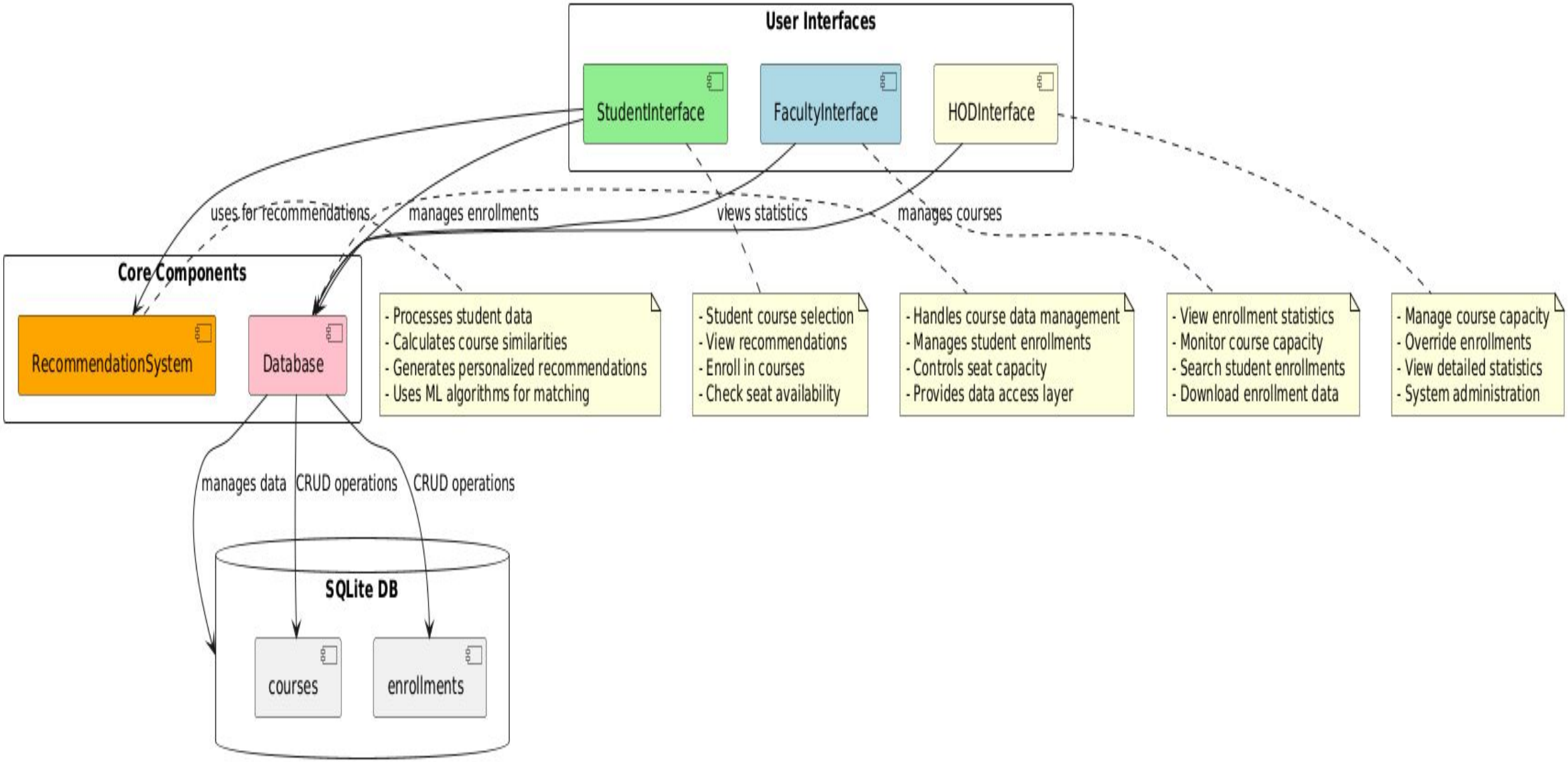
- **Why it's a problem**: The current list of elective options is not personalized for each student. The recommendations are not based on student preferences, academic needs, or the department they belong to. This leads to confusion when students either cannot enroll in their preferred subjects or when they realize their friends are in different classes, which decreases their engagement.
- **Solution**: The system will offer personalized recommendations using AI/ML techniques. It will track the student's preferences, department, and past academic performance to recommend the best available electives that match their profile. The system will also ensure that students know which electives are available and prevent them from enrolling in full courses.

Proposed Method

HoD's Monitoring and Distribution:

- **Why it's a problem**: HoDs currently face challenges in managing student distribution. Without an automated system, they must assign students to electives randomly to achieve balance. This is not only inefficient but can also lead to dissatisfaction among students who feel their preferences were ignored.
- **Solution**: The system will monitor student enrollment in real time and adjust recommendations accordingly. HoDs will be able to oversee the process without having to manually assign students. The system will dynamically redistribute students to ensure that no course is over-enrolled.

Architecture



Modules

- **Data Preprocessing:** Cleaning and structuring student grade data and course information.
- **Recommendation Engine:** Implementing collaborative filtering using cosine similarity.
- **Database Management:** SQLite for storing and managing enrollment data.
- **User Interfaces:**
  - **Student Interface:** For viewing recommendations and enrolling in courses.
  - **HOD Interface:** For managing enrollments and viewing analytics.
  - **Analytics Interface:** For tracking algorithm performance.
- **Error Handling:** Implementing robust error checks and user feedback.

Hardware/software details

- **Hardware:** Standard computer/laptop for development and testing.
- **Software:**
  - Python 3.12.4
  - Libraries: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn
  - Streamlit for web interface (for Student and HOD’s)
  - SQLite for database management (python DBMS library)
  - VS Code as the development environment

Timeline of Project

References

Ray, S. & Sharma, A., "Collaborative filtering for personalized course recommendation," *Journal of Educational Technology*, vol. 12, no. 3, pp. 45-51, 2011.

Kiratijuta Bhumichitr, C., et al., "Alternating Least Square Algorithm for Course Recommendations in Educational Systems," *IEEE Transactions on Education*, vol. 60, no. 2, pp. 120-128, 2017.

Neha Bhagwan Samrit, "Hybrid Course Recommendation System using Collaborative Filtering and Content-Based Filtering," *International Journal of Computer Applications*, vol. 158, no. 7, pp. 22-27, 2017.

Prof. Naren.J, et al., "NLP based Course Recommendation System using Hybrid Filtering," *International Journal of Scientific Research in Computer Science*, vol. 10, no. 2, pp. 140-145, 2019.

References

C. Suneetha, et al., "An Efficient Course Recommendation System using Cosine Similarity and Hybrid Algorithms," *Proceedings of the 2023 IEEE International Conference on Computer Science and Educational Technologies*, 2023, pp. 200-206.

Vanchinathan, A., et al., "Neural Collaborative Filtering for Personalized Course Recommendations," *IEEE Conference on Machine Learning in Education*, vol. 3, no. 1, pp. 70-75, 2020.

Smith, J., et al., "Deep Learning for Recommendation Systems in Higher Education," *Journal of Applied Artificial Intelligence*, vol. 34, no. 4, pp. 233-245, 2020.

Lee, J. H., et al., "Addressing Cold-Start Problems in Course Recommenders with Hybrid Filtering," *IEEE Transactions on Learning Technologies*, vol. 12, no. 3, pp. 350-355, 2019.

Wang, Z., et al., "Privacy-Preserving Recommendation Systems in Higher Education," *Journal of Educational Data Mining*, vol. 15, no. 2, pp. 195-207, 2022.

Liu, X., et al., "Towards Personalized Course Recommendations: Combining Collaborative Filtering with Course Descriptions," *Proceedings of the IEEE International Conference on Learning Analytics and Knowledge*, vol. 5, no. 1, pp. 90-95, 2021

References

Lee, J. H., et al., "Addressing Cold-Start Problems in Course Recommenders with Hybrid Filtering," *IEEE Transactions on Learning Technologies*, vol. 12, no. 3, pp. 350-355, 2019.

Wang, Z., et al., "Privacy-Preserving Recommendation Systems in Higher Education," *Journal of Educational Data Mining*, vol. 15, no. 2, pp. 195-207, 2022.

Liu, X., et al., "Towards Personalized Course Recommendations: Combining Collaborative Filtering with Course Descriptions," *Proceedings of the IEEE International Conference on Learning Analytics and Knowledge*, vol. 5, no. 1, pp. 90-95, 2021

Thank You