Report: Recommendation Algorithm for Personalized Study Material

1. Introduction

The goal of this recommendation system is to provide personalized study material recommendations for college students based on their course, interests, past performance, and material popularity. This platform provides study materials (articles, tutorials, and quizzes) to college students based on their course, interests, and past performance.

2. Approach and Methodology

The recommendation system was designed with several key factors in mind:

- **Student Profile**: Course, year, and areas of interest are integral to filtering relevant study materials.
- **Student Performance**: The student's past performance (average quiz scores) helps match materials.
- **Material Popularity**: Study materials with higher popularity (views and ratings) were considered more valuable .

3. Synthetic Data Generation

Given the necessity for robust testing, **synthetic data** was created to simulate various student profiles, materials, and engagement scenarios. This approach ensured evaluation of the recommendation system without being constrained by the limitations of real-world data availability. The dataset included diverse courses, performance levels, interests, and engagement metrics.

4. Data Visualization

To enhance understanding of relationships within the data, **heat maps** and **correlation matrices** were used . These visual tools helped in identifying patterns between student performance, material popularity, and engagement levels. For instance:

- The correlation map indicated how student quiz scores related to their engagement with specific study materials, allowing for fine-tuning of the recommendation criteria.

5. Scoring System

The scoring system is based on the following factors, each contributing a specific weight to the final score:

- **Interest Match (30%)**: A binary score (1 if the material's subject matches the student's interests, otherwise 0).
- **Performance Fit (30%)**: Alignment of material difficulty with the student's average quiz performance.
- **Popularity Score (20%)**: Normalized score based on how popular the material is (views and ratings).
- **Views Score (20%)**: Normalized score reflecting how frequently the material has been viewed.

The final recommendation score for each material was calculated as:

Final Score=(0.3×Interest Match)+(0.3×Performance Fit)+(0.2×Popularity Score)+(0.2×Views Score)\text{Final Score} = (0.3 \times \text{Interest Match}) + (0.3 \times \text{Performance Fit}) + (0.2 \times \text{Popularity Score}) + (0.2 \times \text{Views Score})Final Score=(0.3×Interest Match)+(0.3×Performance Fit)+(0.2×Popularity Score)+(0.2×Views Score)

6. Evaluation of Performance

To evaluate the effectiveness of the recommendation system, we used **Mean Average Precision (MAP)**, a metric that measures the precision of recommendations at various cutoffs:

- **MAP Calculation**: We generated a set of 5 recommendations for each student and compared them against the materials they had previously interacted with and rated. MAP was calculated as the mean precision over multiple recommendation levels (e.g., top 1, top 2, top 3, etc.).

A baseline MAP of **0.4305** was observed during initial testing. After refining the algorithm, the MAP was improved significantly (exact value subject to experimentation with different weights and models).

7. Deployment

The entire recommendation system was deployed using **Streamlit**, providing an interface for users to input their data and receive personalized study material recommendations.

Conclusion

The recommendation system successfully provides personalized study material suggestions by leveraging a hybrid approach. The system adapts well as more interaction data becomes available. Further refinements, such as exploring deep learning-based recommendation models, could improve the system's predictive accuracy.