

ELECTIVE RECOMMENDATION SYSTEM

Godavarthi Naga Manoj Balaji¹, Golla Meghana², Munagala Naga Sameera³,
Ayush Chhetri⁴

^{1,2,3,4}Presidency University Bangalore, India.

E-Mail : godaverthinagamanojbalaji@gmail.com, E-Mail : meghanagolla1@gmail.com ,

E-Mail: nagasameera@gmail.com ,E-Mail: itsmeayush2003@gmail.com

DOI: <https://www.doi.org/10.58257/IJPREMS38179>

ABSTRACT

The paper project discussed herein targets, in higher detail, the development of an elective recommendation system with the use of techniques related to machine learning and NLP. It will recommend courses based on textual course descriptions and user inputs relevant to the academic goal and preference of the learner. Recommendations would be done in consideration of a textual course description and the user input towards academic goal alignment and learners' preferences. To meet this objective, many strategies have been taken on board that entails: Content-Based Filtering, Collaborative Filtering, Semantic Analysis using SBERT. Results in detail would then be judged against different metrics, like Precision @K, Recall @K, and F1 Score. Of course, all the results undoubtedly show the superiority of SBERT; hence, it provides flexibility and scalability for any modern educational environment.

Keywords- Recommendation System, Cosine Similarity, SBERT, Collaborative Filtering, Personalised Education, Machine Learning in Education, NLP, Semantic Analysis

1. INTRODUCTION

The objective will be the revival of the discovery process of a course by using sophisticated machine learning better so as to recommend content in a manner closer to the interests or goals of a learner. For this, therefore, the system needs to process text-based course descriptions and user input with natural language processing and machine learning so as to capture patterns, similarity, and relevance. This approach takes into account subtlety and course topics along with its learning objectives or user preferences in generating highly personal and targeted recommendations.

Content-based filtering approach using the recommendation engine assesses courses and queries based on the content given to it by its users. It is more complex than simple keyword matching, as it really probes into understanding the semantics surrounding course descriptions in order to derive much more adequate and precise similarities to measure. It measures similarity scores for each course relative to the query on the part of the user. Therefore, it can give an optimized list of course recommendations, which can fulfill their immediate needs, but at the same time, courses that will expand their knowledge in other related areas.

Then, the ranking of courses will depend upon their relevance. In the meantime, it provides a comparison of courses that in one way or the other are related to an input. This will provide a complete view to users about the choices available and will discover courses that may help them learn some new things. For that matter, with regard to each course, its evaluation will be there, and therefore similarity scores will provide somewhat detailed material so that a user is able to make informed decisions. This recommendation engine will support flexible and adaptive learning since its suggestions are always improved with regard to the user feedback flowing in constantly. With further engagement of the users with the system and their additional input, the recommendations will be re-calibrated according to the new preferences of the users. The dynamic feature will ensure that suggestions remain relevant and up-to-date with changing needs.

It tries to create a sense of a well-crafted and easy journey through learning and saves time and efforts for the users. They shall decide just the right way in choosing up the course up, according to interests and preferences considering the option while choosing courses it will be much of a customized option as far as one needs to learn the thing, which will eventually assist one in the best possible outputs acquired from this training process either on academic or professionally. This digital learning system redefines course selection in a learning environment using its algorithms as well as design centered on a user.

2. LITERATURE REVIEW

The existing literature has elaborated on different approaches toward course recommendation systems in higher education. Rule-based systems are the simplest system but not scalable or personalized. Collaborative filtering is simple for recommendations of courses based on similar student preferences; however, sparse data lead the system to problems. Content-based filtering uses course description to recommend courses, but it is uncritical over the capture of complex relationships between students and courses.

Hence, Personalization is the key to good recommendations. Hybrid methods that merge either collaborative and content based filtering or exploit machine learning algorithms help to bridge over the above mentioned flaws as well as make it more accurate and relevant suggestions. Furthermore, heterogeneity in student interests, academic performance, career goals, as well as real-time feedbacks can further personalize and make it relevant over time.

Ethics would also come into studies. Data points here include aspects of privacy and no discrimination in recommendation. This aspect can create trust and equity in the system.

Although it can be helpful to students, it benefits teachers as well when learning is concerned. That's because analysis of students' data will make curriculum and methods used in teaching more vigorous. Also, they will know who the risky students are so that they can prevent them.

While such approaches as decision trees, genetic algorithms, and neural networks have now been applied in course recommendation, there are still challenges they face. Among the scalability issues of handling large scale data, there are also concerns on e-waste and energy consumption when running the system.

In short, there is immense scope that course recommended systems can be of much help in ubiquitous personalization in enriching course choices by the students, achieving quality improvements in learning processes, and hence becoming beneficial to both the students and teachers. Future research studies should thus mainly focus on scalable challenges and ethical considerations related to the exploration of deep learning techniques in this field.

3. METHODOLOGY

3.1 Dataset Description

The dataset for this research is course data that contains metadata and user reviews which are very rich to be used in recommending these courses to a user.

Course Name: Identifies the course by their ID or name/title; hence it has to be unique since two different courses may have the same title.

Course Description: Long description in text about what the course teaches and accomplishes. Pre-raw input for semantic and content-based analysis.

Level: Degree of difficulty, such as Beginner, Intermediate, Advanced.

Skills: List of the skills a learner will pick up at the end of a course serving to align a course with some user's career or learning goal.

Pre requisites: Presupposed previous knowledge or courses to be undertaken by students in respect of a certain course useful for the recommendation system in order to offer courses which precisely fit the level of skill.

Course Rating: A rating score provided by the users may also be utilized by the collaborative filtering models with a view to judging the quality and relevance of a course.

3.2 Data Preprocessing

The pre-processing of the data has been done to bring efficiency in the working of the entire system and deliver results accurately-albeit mandatorily.

Treatment of Missing Values: Records have been removed in case of missing description or pre-requisite of courses. If partial, i.e., rating of courses missing, the system fills it out with the mean or median taken from the other courses based on imputed records in similar course metadata.

Text Preprocessing: The course descriptions will be pre processed by removing excessive characters and bad formatting that might affect the working of the model. This will be done by the following: making everything small for text uniformity. Excluding punctuation and special characters because they are noisy. Removing common stop words like "the," "and," "in" since all these will not add meaning to the content.

Lemmatization: Words are then taken back to the original form through different algorithms, such as Porter's Stemmer.

Normalization: Since there will features with two different kinds, the numeric ones, in case of Course Ratings and CGPA Average, all of the numeric, are normalized using standard scaling, at comparable scale, techniques. Such a process addresses all the feature values equally nice so that this, maybe, recommendation model is not going into feature scale.

5.3 Model Implementation

It combines three recommendations that are: Cosine Similarity, Collaborative Filtering and Semantic Matching by SBERT. More in-depth how those were implemented on that model will be provided further.

3.3.1 Cosine Similarity:

Course descriptions are vectorized using TF-IDF, a technique which weighs the importance of each term with respect to the whole corpus of course descriptions. Now, compute cosine similarity between the course vectors in order to

recommend similar courses w.r.t their content. Run two variants - one taking the stemmed course description as an input, and one w/o stemming - describe both results w.r.t the performance impact of stemming.

3.3.2 Collaborative Filtering:

Description: Different collaborative filtering methods would be implemented to make recommendations based on preference or interaction by users with courses. This kind of model will, in general, depend on historic data comprising course ratings, course completion, or even the interaction of users with courses.

Implementation:

User-Item Matrix: A matrix will be created wherein a set of rows will denote users and columns will denote courses. Entries would show the rating or interaction made by the users with respect to courses.

Matrix Factorization: This matrix is then factored using some techniques such as Singular Value Decomposition into latent factors that capture the characteristics preferred by the users and courses.

Then, it predicts the missing ratings or course preferences by mapping user preferences to latent factors and provides personalized course recommendations.

3.3.3 Semantic Matching with SBERT (Sentence-BERT):

Course descriptions are taken for this purpose, such that this language model, SBERT, transforms these into a dense, high-dimensional embedding space representative of the meaningful elements of such descriptions.

Semantic Similarity Computation: This will allow the system to compute similarities of user queries or selected course descriptions against the whole course database by comparing their embeddings for cosine similarity. Indeed, the system will match courses for the user, not by keyword overlaps but in such a way that it will understand the context and intent behind the descriptions.

It's particularly good at catching relations between courses that are worded differently yet have the same learning objectives; SBERT is a godsend.

3.4 Evaluation Metrics

The performance of various recommendation models is evaluated by using the following metrics:

Recall @K: It gives the measure of recall, giving how many relevant courses out of all the relevant courses have been captured in top K recommendations. That is pretty important too, because usually, there are some courses which any system must recommend that might fail.

F1 Score: The F1 score is the harmonic average of Precision and Recall, hence striking a balance between the two, while yielding one value for model performance.

Mean Reciprocal Rank: MRR is a ranking quality measure in recommended courses, and it informs on how highly ranked relevant items are. A higher MRR means the system gives recommendations of higher quality towards the top of the list.

These will form the bases for objective comparisons among different approaches, including Cosine Similarity, Collaborative Filtering, and fine-tuning of SBERT in trying to optimize user satisfaction with models and model accuracy. The following extended methodology has comprehensively identified the design of the system, the steps of data processing, and strategies followed by the author in its evaluation. Please let me know if you have any additional information or even changes and/or modifications regarding this summary.

4. RESULTS

To evaluate the performance and quality of the recommendations provided by this Elective Recommendation System, several approaches were attempted; Cosine Similarity-with and without stemming-, Collaborative Filtering, and finally, SBERT. A summary of how good the recommendations produced by the model are goes like this:

4.1 Performance Metrics:

Name of metrics	Cosine similarity with steaming process	Collabrative filtering	Sentence bert	Cosine similarity without steaming process
Precision	0.0361	0.00	0.00	1.00
Recall	0.1818	1.00	1.00	0.50
F1 Score	0.0603	0.00	0.00	0.67
Mean Absolute Error(MAE)	919.5602	0.0189	0.189	0.0109

Hit Rate	0.0085	0.00	0.00	0.00
Mean Average Precision(MAP)	0.1818	0.00	0.00	0.00
Mean Squared Error(MSE)	1339437.0301	0.002	0.0021	0.0010
Root Mean Squared error (RMSE)	1157.3405	0.0457	0.0457	0.0312

4.2 Recommendation Accuracy:

It was doing great in comparison with other approaches since the SBERT could capture subtlety in semantics of text. The methods based on cosine similarity were performing well, but the limitation was with the keyword-based match. There is the cold start problem of new users and courses in the case of competitive collaborative filtering.

4.3 User Feedback:

Generally, very engaging and satisfactory responses toward the SBERT-based recommendations were retrieved from the user's satisfaction survey on the complicated queries.

4.4 Insights from Visualizations:

Important trends regarding preferred skills came out of the analysis through word cloud and the bar chart, and therefore, represented valid indicators in respect of student selection with respect to resource allocation and subsequent selection behaviour by any HEI.

4.4.1 Strengths of Methodologies:

Strengths of SBERT:

Although others may have been absolutely perfect, SBERT-based was obviously much more powerful in retrieving semantic meaning out of the input from the user and was hence absolutely suitable for intricate and varied questions.

Cosine Similarity:

This too was pretty fantastic, but still, methods which were cosine similarity-based suffered because their search based on lexical overlap remained highly challenging due to nuances found in questions asked.

Collaborative Filtering:

Collaborative filtering was performed very well with user interaction data, but it required loads of historical data to return good results and therefore was highly constraint whenever the dataset is sparse.

4.4.2 System Strengths:

The modularity of architecture enabled smooth interfacing of diversified methods with the system for greater flexibility. Cloud deployment enables scaling up large data sets and facilitates real time recommendations.

4.4.3 Challenge Faced:

Cold Start Problem:

This collaborative filtering cannot deliver course assignments to either new users or newly added ones. This is balanced out by the interfacing of methods in content-based methods.

Computation Cost:

The very effective pre-trained model, SBERT, utilized extremely vast amounts of computational resources. This had to be optimized using a scaling strategy that would accommodate growth.

4.4.4 Impact on Education:

This system can provide more course interaction, which is more relevant to the needs of a student about pursuing his or her academic and professional objectives. These learnings, in terms of trending courses and user preference, further can be capitalized at the institutional level through visualization tools for optimization.

5. FUTURE ENHANCEMENTS

There will be an integrated real-time data stream that will allow dynamic adaptation of model recommendations. Multimodal data, course video, user review, may further increase the dimension to scale up the system and result in recommending the richer versions. Better explainability to the users in order to build trust as to why those courses were recommended.

6. DISCUSSION

Strengths: SBERT semantic capability provides context-sensitive recommendations. Cosine-similarity and collaborative filtering get the basics right.

Weakness: Cold start challenge in CF, heavy computations involved in SBERT.

Implication on Education: The approach contributes to the realization of personalized learning for the benefit of students, and actionable insights with analytics by the institutional policies.

Future Work

1. Corporate Learning and Development:

This system would recommend courses, or certificates related to what they do now, their present performance, or their career goal. It would also recommend "Data Science with Python" to that wanna-be analyst.

2. Career Guidance Portal:

Suggest skill development courses that go along with the trend going on in that industry. For instance, full stack web development to a software engineer who needs to shift his roles to frontend.

3. EdTech Platforms:

Enhance search-ability of courses on such platforms: Coursera, Udemy, or Khan Academy

Example: Recommending courses from a user's history of enrollment or a pattern of searching.

4. K-12 Education:

Encourage the children to opt for optional activities, club or academic options which a child can participate based on his interest or performance.

5. Rural Class Skill Development:

Provide vocational courses which should be delivered for rural masses or deprived sections.

Example: Mobile Repair Technician Training for rural youth.

7. CONCLUSION

Elective selection in group recommendation systems is the challenge that helps students in the right direction toward courses that best fit their academic interests, career aspirations, and interdisciplinary goals. The Elective Recommendation System examines individual preferences by using advanced machine learning and natural language processing for personalized recommendations with precision. It would also make the process of decision making quite easier, since this system categorizes courses in a way that displays skills, academic relevance, and even future opportunities toward optimized enrollment. This would really encourage goal-directed searching and alleviate some of the barriers to learning. A tailored approach would ensure motivation, satisfaction, and active participation that shall, in return, enable these students to make informed choices toward the attainment of academic and professional goals with increased confidence and ease.

8. REFERENCE

- [1] Saurabh Amrutkar , Recommender Systems for University Elective Course Recommendation ,researchgate , April 2021
- [2] Anusha Naik , Ashrita B Kumar , Dhanush Virupaksha , H Priyanka , Elective Recommendation System, ieeexplore , 2024
- [3] Raihan Chaira , Ade Romadhony , Elective Course Recommendations System Using Recurrent Neural Network , ieeexplore , 2024
- [4] Shrooq Algarni , Frederick Sheldon , Systematic Review of Recommendation Systems for Course Selection , mdpi , 2023
- [5] Mfowabo Maphosa¹ , Wesley Doorsamy² , Babu Paul³ , A Review of Recommender Systems for Choosing Elective Courses (IJACSA) 2020
- [6] Mamata Garanayak , Sipra Sahoo , Sachi Mohanty , An Automated Recommender System for Educational Institute in India , researchgate , 2018
- [7] Sami Alghamdi , Osama Sheta , Mohmmmed Adrees , A Framework of Prompting Intelligent System for Academic Advising Using Recommendation System Based on Association Rules , ieeexplore , 2022
- [8] Erwin Hidayat , Recommendation System for Major University Determination Based on Student's Profile and Interest , academia , 2020
- [9] K. N. Lakshmaiah¹ , D. Anitha² , An Experimental Study on Partial Replacement of Rice Husk Ash (RHA) With Cement of Concrete , IJSR , 2017

-
- [10] Boxuan MA, Yuta Taniguchi, Shin'ichi Konomi, Course Recommendation for University Environments, Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020), 2020
 - [11] Nabila Kamal; Farhana Sarker; Arifur Rahman; Sazzad Hossain; Khondaker A. Mamun, Recommender System in Academic Choices of Higher Education: A Systematic Review, IEEE, 2024
 - [12] Mfowabo Maphosa¹, Wesley Doorsamy, Babu Paul, A Review of Recommender Systems for Choosing Elective Courses, (IJACSA) International Journal of Advanced Computer Science and Applications, 2020
 - [13] Pradeep Kumar Singh, Pijush Kanti Dutta Pramanik, Avick Kumar Dey, Prasenjit Choudhury, Recommender Systems: An Overview, Research Trends, and Future Directions, researchgate, 2021
 - [14] Mfowabo Maphosa, Wesley Doorsamy, Babu Paul, A review of recommender systems for choosing elective courses, International Journal of Advanced Computer Science and Applications, 2020
 - [15] M. Rekha Sundari, G. Shreya, T. Jawahar, Course Recommendation System, International Journal of Computer Applications, 2020