A Strategic Approach in Course Recommendation System Using Similarity Models

Angel Susan Vino[1], DiyaAlDin[2], Parvathy Menon[3], Dr. S.Sandosh[4]

{angelsusan.vino2020@vitstudent.ac.in[1], diyaal.din2020@vitstudent.ac.in[2], parvathy.menon2020@vitstudent.ac.in[3], sandosh.s@vit.ac.in[4]}

Vellore Institute of Technology Chennai, India

Abstract. In response to the changing dynamics of the modern world, where information and education play key roles, recommendation systems have developed as critical components across a wide range of digital platforms. This study digs into the development of an online course selection system geared to fulfill the educational needs of a diverse audience, including working professionals, students, and ardent learners committed to lifelong learning. Against the backdrop of an ever-changing ecosystem, this research project intends to harness recommendation algorithms to greatly improve learning outcomes by providing individualized and adaptable learning opportunities. The prevalence of recommendation systems throughout social media, applications, websites, and numerous technologies emphasizes their importance, driving our initiative to con-tribute to the educational area through the development of a specific online course recommendation system.

Keywords: Recommendation , E-learning , Content-Based , NLP , Similarity etc.

1 Introduction

In today's digital age, where access to knowledge and educational resources has radically increased, navigating the vast landscape of available courses and learning resources has become significantly more difficult. Modern students have a variety of requirements and preferences, and traditional educational methods no longer meet their needs and passions. The profusion of instructional information brought about by the rise of online education platforms, e-learning environments, and Massive Open Online Courses (MOOCs) has made the task of choosing appropriate courses a challenging one(Fig.1.).

Course recommendation systems have emerged as a critical innovation for resolving this complex problem and improving student learning outcomes in both traditional and digital educational environments. These systems utilize cutting edge algorithms, data analytics, and user behavior analysis to provide learners

ICAMDMS 2024, February 23-24, Coimbatore, India
Copyright © 2024 EAI
DOI 10.4108/eai.23-2-2024.2346986

with individualized course recommendations that are tailored to their specific goals, interests, and skills. The extensive literature reviews and research citations that form the basis of this course recommendation system initiative emphasize the significance and necessity of providing students with individualized course recommendations. This project paper intends to contribute to the ongoing development of course recommendation systems by analyzing cutting-edge research, including works on MOOC recommendation algorithms, educational recommender systems, and hybrid recommendation approaches [6],[7].

This paper's objective is to design a course- recommendation system that transcends the limitations of traditional educational paradigms. The paper aims to empower students with a tool that not only simplifies the course selection process but also enhances the overall educational experience by combining in-sights from existing research with cutting-edge methods. It is anticipated that the project's completion will ushering a transformative shift in education by guiding each student along a unique, enriching, and empowering educational path.

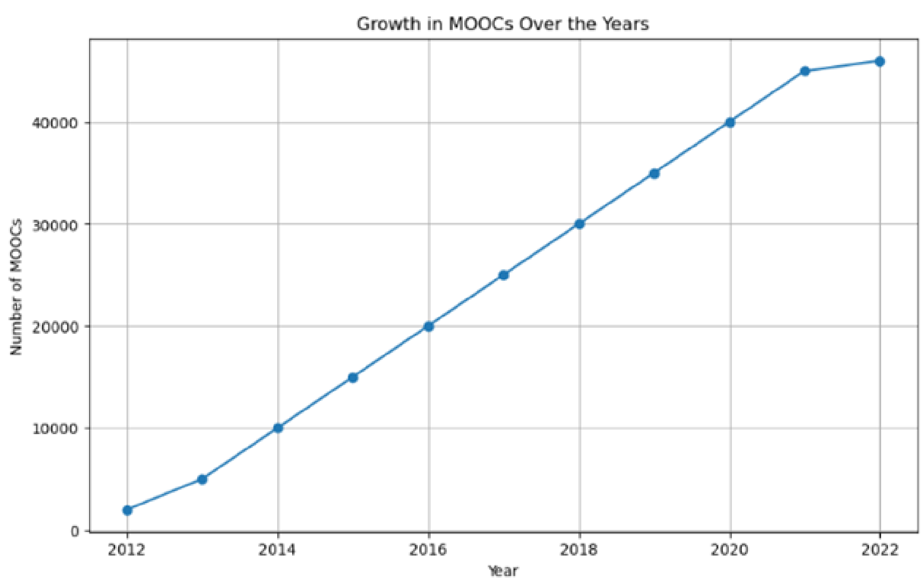

Fig.1. Growth of MOOCs over the years (2012-2022)

2 Literature Review

2.1 Encouraging User Participation in a Course Recommender System: An Impact on User Behavior-Rosta Farzan, Peter Brusilovsky

Course Agentisa community-based course recommendation system, and Rosta Farzan and Peter Brusilovsky explore the implications of increasing user participation in this study. The authors address the critical issue of user participation in collaborative and social recommender systems, underlining the significance of user ratings (both in terms of quantity and quality) to the success of the system. Even if their contributions aren't immediately beneficial to them, they devise an incentive scheme that aims to make user input useful to them personally. Students' course evaluations are used alongside a career progression tool to provide them feedback on their progress. The authors conducted two user tests to investigate the efficacy of this incentive mechanism, and the results showed both positive and negative effects on user behavior. The report also examines the potential drawbacks of incentive systems, such as reduced internal drive and gaming, and highlights the concept of "positive rating bias." The article concludes by discussing the pros and cons of such procedures, expanding our knowledge of the potential impact of incentives on user behavior in community-based recommender systems[1],[8],[9],[10].

2.2 Recommender Systems for University Elective Course Recommendation– Kiratijuta Bhumichitr,Songsak Channarukul, Nattachai Saejiem, Rachsuda Jiamthapthaksin, Kwankamol Nongpong

The scholarly article authored by Kiratijuta Bhumichitr et al. explores the pertinent domain of recommender systems, specifically focusing on the provision of recommendations for optional courses within a university setting. Despite the significant study and utilization of recommender systems across other domains, the task of recommending university courses remains a challenging area that has attracted less research attention. This research holds great significance as it caters to the needs of undergraduate students seeking course recommendations and streamlines the course selection process during pre-registration. The study employs Alternating Least Squares (ALS) and collaborative-based recommendation using the Pearson Correlation Coefficient, which are both widely recognized recommendation techniques. The evaluation of their achievement is conducted by utilizing authentic academic records obtained from university students. Based on the research data, it has been determined that ALS outperforms collaborative-based recommendations, achieving a notable accuracy rate of 86 percent. Furthermore, this research has made a substantial contribution to the field by providing valuable insights and presenting avenues for further investigation and practical applications within academic settings. Additionally, this conclusion underscores the efficacy of ALS in the realm of recommending elective courses at the university level. [2]

2.3 A Survey Paper on E-Learning Recommender System – Reema Sikka , Amita Dhankhar, Chaavi Rana

Sikka, Dhankhar, and Rana (Year) propose a revolutionary vision for e-learning through a recommender system in their ground-breaking work. The paper sup-ports for individualized learning experiences utilizing real-time web mining tools, based on the success of similar systems in e-commerce. The suggested system combines a "learning" module that learns from previous access patterns and "advising" module that dynamically tailors recommendations using methods such as clustering, association rule mining, and collaborative filtering. This novel technique addresses the lack of automated recommendation systems in education, highlighting the potential to transform learning resources. The study presents recommender systems as a promising frontier in education, ready to bring in a more adaptive and individualized approach to learning by addressing the diversity of learners' interests and skills.[3]

2.4 Recommender Systems for Learning: Building User and Expert Models through Long-Term Observation of Application Use-Frank Linton, Hans-Peter Schaefer

OWL, a robust recommender system focusing on information technology(IT), is introduced by Linton and Schaefer (Year) to improve learning experiences. Unlike traditional systems, OWL goes beyond simply recommending URLs or programming classes, instead focuses on IT abilities as measured by user interactions. By recognizing departures from predetermined patterns, it analyses user behavior, determines expected values, and detects learning opportunities. Examining Microsoft Word commands shows a Zip of distribution, which serves as the foundation for OWL's recommendation structure? OWL excels in providing personalized learning recommendations based on individual user behavior and preferences for successful skill development. The research highlights the observability of IT tasks, claiming that monitoring and evaluating them improves learning outcomes. OWL's paradigm change in recommender systems extends beyond desktop applications, providing tailored learning opportunities in a variety of IT disciplines, employing user behaviour data for individualized recommendations, and enhancing IT skill acquisition.[4]

2.5 A Hybrid Approach for Supporting Adaptivity in E-learning Environments-Mal -Omari, JCarter, FChiclana

Using agent technology and the Event-Condition-Action (ECA) model, the study provides a hybrid framework for facilitating flexibility in e-learning settings. Intelligent Tutoring Systems (ITSs) and Adaptive Hyper Media Systems (AHMSs), which provide adaptive display and navigation, have impacted the suggested framework. To provide LMSs with real-time adaptation, the framework integrates agent technology and the ECA paradigm. As the foundation for adaptation, the Felder-Silverman Learning Style Model (FSLSM) is chosen. The ECA model is used to identify events in the e-learning environment, and the hybrid strategy of agent technology with the ECA model enables the LMS to dynamically change learning materials and experiences. The system can allow dynamic and individualized adaptive processes and is relevant to different e-learning environments such as MOOCs. The authors concluded with a glimpse of the future research directions, including a review of the framework using a variety of real-world case studies.[5]

3 Proposed Model

3.1 Data Collection

To determine the target audience's interests and ability levels, data must be gathered as the initial stage in the evolution of the course recommendation system. There are two main parts to this data collection process:

3.1.1 Survey Administration Data on people's skill levels and areas of interest for more research are gathered via a questionnaire. In order to ensure representation from students, working professionals, and learning enthusiasts, the survey is distributed to a wide sample of respondents. The dataset was built around 200 surveys and the questionnaire was focused on the area of interest of the individuals. It mainly consisted of some beginner to advanced level questions from that particular area.

3.1.2 Pre-existing Course Dataset Concurrently, a pre-existing dataset of courses is assembled, which includes course titles and corresponding levels of difficulty.The dataset had more than 3500 courses. The basis for the course suggestions is this dataset.

3.2 Data Preprocessing

To guarantee correctness and consistency, the gathered data is pre-processed before being combined and analyzed. The following are the steps in data preprocessing:

3.2.1 Cleaning of Survey Data In order to provide a clean and trustworthy dataset, any missing or inconsistent survey responses are resolved.

3.2.2Course Dataset Refinement Duplicate entries are removed and course difficulty level classifications are improved by reviewing the course dataset.

For both the above datasets the preprocessing was done using natural language processing techniques including text vectorization and other needed mechanisms.

3.3 Evaluation of Skill Level

The survey data is used to evaluate respondents' skill levels in their areas of interest. Through the quantitative examination of survey responses, this is accomplished. Based on the statistics, skill levels can be divided into three categories: beginner, intermediate, and advanced.

3.4 Recommendation Algorithm

The algorithm that matches a person's interests and ability level with relevant courses from the database already in place is the foundation of the recommendation system. In order to improve the recommendation system's accuracy, thepaperusesacontent-basedrecommendationstrategy.Thecontent-basedapproach makes use of aspects related to things and user preferences in order to capitalize on the intrinsic qualities of both persons and items. In particular, we employ feature vectors, or collections of features, to represent products and user profiles.

3.4.1 Cosine Similarity The paper use cosine similarity for feature vectors, which calculates the cosine of the angle between two vectors. The calculation of the cosine similarity (cosinesim) between vectors P and Q is as follows:

$$\text{cosinesim}(\mathbf{P},\mathbf{Q}) = \frac{\mathbf{P}\cdot\mathbf{Q}}{\|\mathbf{P}\| \cdot \|\mathbf{Q}\|} \tag{1}$$

Here P .Q represents the dot product of vectors P and Q, while ||P|| and ||Q||denote the Euclidean norms of vectors P and Q respectively.

3.4.2 Jaccard Similarity The paper uses the Jaccard similarity coefficient to calculate how similar two sets of features are to one another. The following is the definition of the Jaccard similarity between two sets, P and Q:

$$J(P, Q) = \frac{|P \cap Q|}{|P \cup Q|} \tag{2}$$

Within the framework of a course recommendation system, P and Q stand for collections of attributes related to products or user preferences. The number of features that both sets have in common is represented by the numerator,

$|P \cap Q|$, while the total number of unique features in both sets is represented by the denominator, $|P \cup Q|$.

3.5 Iterative Improvement

Based on user feedback and system performance, the process is continuously re-fined. To guarantee that the recommendation system continuously provides top-notch course recommendations, upgrades and improvements are implemented on a regular basis. The methodology focuses on enhancing personalized recommendations by aligning user preferences with item features, providing a foundation for a more accurate and tailored recommendation system.

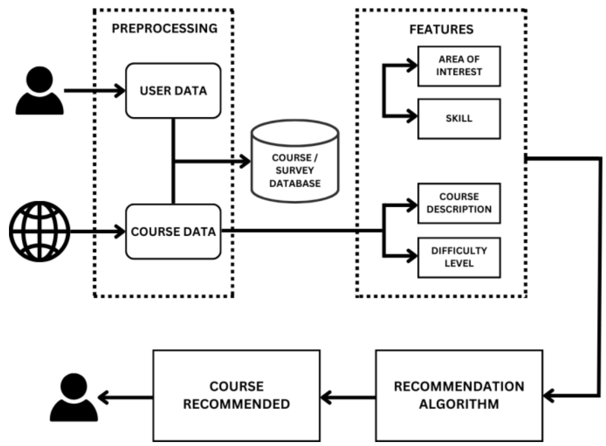

Fig.2. Architecture of the Proposed System

So, in the proposed system the two datasets were collected, preprocessed and was implemented on the recommendation algorithm and the nutshell of the entire methodology is being described using Fig. 2. The methodology focuses on enhancing personalized recommendations by aligning user preferences with item features, providing a foundation for a more accurate and tailored recommendation system.

4 Results and Discussion

The estimated similarity scores are used by the recommendation system algorithm developed in this paper to find products that closely match user preferences. The paper computes the similarity between a user's preferences and the features of the things that are offered for them. The user is then recommended the top-N items with the highest similarity ratings. The paper is majorly build based on user preferences and a content-based recommendation system is used in here.

Table1. Evaluation Metrics

Model	Precision	Recall	F1Score
Cosine Similarity	0.75	0.50	0.60
Jaccard Similarity	0.63	0.47	0.54

The Table1 is an evaluation metrics output based on the two model's cosine similarity and Jaccard similarity that comes under content-based recommendation system. Cosine similarity works better than Jaccard according to these values. The explanation is depicted using theTable2.

Table2. Comparison Study between two Methods

Metrics	Cosine	Jaccard
Precision	Out of all items recommended by the model, 75% are relevant (Precision=0.75)	Out of all items recommended by the model, 63% are relevant (Precision=0.63)
Recall	Captures 50% of all relevant items (Recall=0.50)	Successfully recommends 47% of all relevant items (Recall=0.47)
F1Score	Achieves a balance with a score of 0.60 (F1Score)	Demonstrates balanced performance with a score of 0.54 (F1Score)

Furthermore, when a harmonic balance of precision and recall is prioritized, with a keen regard for avoiding false positives and false negatives, the Cosine Similarity model emerges as a marginally preferable choice based on the metrics offered. The better precision (0.75) and F1 Score (0.60) compared to Jaccard Similarity demonstrate its efficacy in delivering a more well-rounded performance in the context of this research's specific aims. The paper is focused on providing all are the main improvements observed during the approach done here compared to the already done ones. Since the dataset used in the suggested recommendation system has been optimized for performance, it might not be possible to directly compare it with other recommendation systems. As such, the focus of this discussion will be on classic recommendation system problems that could occur in systems proposed by other researchers, as shown inTable3.

Table 3. Existing vs. Proposed

Existing System	Proposed System
There is lack of clarity in the proposed recommendation system developed by Farzan et al. since rating can be artificially boosted based on the company's requirement. There is no direct role-play available to the user.	This proposed recommendation system was developed based on the data that was directly collected from the users. Required data has been collected through questionnaires and the ratings of the course is not influencing the recommendation as of now.
Requires user histories and similar user data for effective recommendations this also known to us a scold start problem.	The paper sorts out as this recommendation system are helping a completely clueless user from avoiding confusions and by simply accessing their level in their area of interest and then suggesting the courses based on their skill level.

5 Conclusion

Concluding the research work has provided an extensive investigation of a content- based course recommendation system employing both cosine and Jaccard similarity metrics. The proposed system aims to improve individualized learning experiences by matching user preferences ascertained by means of a skill and interest assessment questionnaire with an already-existing course dataset.

Cosine similarity performed better than Jaccard similarity in the recommendation system evaluation, indicating its superiority in capturing the semantic links between user profiles and course descriptions. Together, the precision, re- call, and F1 scores attested to the cosine similarity model's ability to produce recommendations that are more pertinent and accurate. A properly designed questionnaire is used to gather the system's base of user skills, which adds a layer of personalization that closely matches user preferences. By matching users with courses that complement their skill set and correspond with their chosen areas of interest, this strategy improves recommendation accuracy.

All things considered; the study adds to the expanding body of knowledge in personalized recommendation systems by clarifying the usefulness of cosine similarity in content-based models for course suggestions. By utilizing a questionnaire to gather user capabilities, the system is better able to customize recommendations, which in turn makes learning more interesting and helpful for users.

5.1 Future Work

This study on course recommendation systems, employing cosine and Jaccard similarity metrics, paves the way for future advancements. Future work includes enhancing skill analysis through NLP techniques, enabling dynamic user profile updates, integrating contextual learning paths, exploring hybrid recommendation models, implementing user feedback mechanisms, ensuring transparency, extending to mobile platforms, conducting large-scale deployments, addressing ethical considerations, and devising user engagement strategies. These efforts aim to transform the recommendation system into a more sophisticated, user- centric platform, providing an enriched and adaptive learning experience.

References

[1] Bhumichitr K, Channarukul S, Saejiem N, Jiamthapthaksin R, and Nongpong K.: Recommender Systems for university elective course recommendation. International Joint Conference on Computer Science and Software Engineering (JCSSE) Vol. 14 (2017).

[2] Sikka R, Dhankhar A, and Rana C.: A Survey Paper on E-Learning Recommender System. International Journal of Computer Applications. Vol. 47 (2012).

[3] Linton F, and Schaefer H-P.: Recommender Systems for Learning: Building User and Expert Models through Long-Term Observation of Application Use. User Modeling and User-Adapted Interaction Vol. 10 (2000).

[4] Al-Omari M, Carter J, and Chiclana F.: A Hybrid Approach for Supporting Adaptivity in E-learning Environments. International Journal of Information and Learning Technology Vol. 33 (2016).

[5] Shani G, and Gunawardana A.: Evaluating Recommendation Systems. Recommender Systems Handbook. (2010).

[6] Algarni S, and Sheldon F.: Systematic Review of Recommendation Systems for Course Selection. Machine Learning and Knowledge Extraction. (2023).

[7] Barnes T.: Q-matrix Method: Mining Student Response Data for Knowledge. Technical report. (2005).

[8] Bose RJC, Deshmukh O, and Ravindra B.: Discovering Concept Maps from Textual Sources. In Proceedings of the 8th International Conference on Educational Data Mining. (2015).

[9] Brusilovsky P, Eklund J, and Schwarz E.: Web-based education for all: a tool for development adaptive courseware. Computer Networks and ISDN Systems, Vol. 30, pp. 291–300 (1998).