Hybrid Course Recommendation System

Chirag Jain[1], Ayush Attawar[2], Parth Narechania[3] and Dr. (Mrs.) Vinaya Sawant[4]
Student[1,2,3], Assistant Professor[4]
Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
chiragjain55551@gmail.com[1], *ayushattawar@gmail.com*[2], *parthnarechaniacool@gmail.com*[3], *vinaya.sawant@djsce.ac.in*[4]

ABSTRACT

E-learning has gained enormous momentum and is becoming a prominent way of learning among the student community. A great amount of useful and varied educational resources are available over the internet through Massive Online Open Courses (MOOCs) websites such as Coursera, Udacity, EdX, etc. Though students have many opportunities to explore, it becomes increasingly challenging and time-consuming to search and examine the vast number of courses for suitable content. To overcome this problem, recommendation systems can be used. This paper proposes a course recommendation algorithm based on user's profile and their similar characteristics to other users. The proposed algorithm combines content-based filtering with collaborative filtering to provide more accurate and targeted results. For subjective testing, a web-based system is developed with 8 courses and 15 users.

1. Introduction

Online learning has gained immense significance in the last decade due to advancements in electronic technologies and media channels. It provides a great opportunity for people to gain new skills using the powerful combination of the internet and education in a cost-effective manner. Hence, it increases flexibility and efficiency. There are various online learning platforms such as Udemy, Coursera, etc., that serve millions of people. The boom of online learning can be recognized as Research and Markets forecasts the online education market as a few hundred Billion by 2025. So, there's a huge demand from people to learn online. Due to this demand and rapid growth of the market, a massive amount of online learning content has been created over the internet through numerous platforms. Hence, the biggest concern for learners is to figure out the best content to gain proper skills so that it proves beneficial and saves time and money as well.

Online learning is being adopted steadily by the student community. Students can easily learn and sharpen their skills using the internet. But, with a wide variety of resources and content available on the internet, it is difficult for students to find specific related content, and many times they enroll in courses without proper research and may get diverted from their actual point of focus [1]. This results in losing a lot of time and money and not getting clarity of the topic [2]. To avoid such problems, proper guidance in terms of course selection, identifying authentic content, decision-making, etc. is very important in every student's life [3].

With enormous data availability, proper technologies and systems need to be developed to handle the data properly and make the most out of it. For huge online learning content, Recommendations systems can be used for handling and suggesting online learning courses to users. Recommendation Systems are Machine Learning (ML) based algorithms. Different courses are recommended to users based on their interests. A recommendation system is a subclass of Information filtering Systems that seeks to predict the rating or the preference a user might give to a specific item. The recommendation system is mainly divided into two broad categories: content-based filtering and collaborative-based filtering.

A content-based recommendation system works with data that is provided by the user, either implicitly (clicking on a link) or explicitly (rating). Based on this data, a user profile is generated, which is then used to make suggestions to the user. Content-based filtering relies mainly on content knowledge sources and descriptions of the items. Collaborative filtering is one of the Recommendation System techniques which is based on gathering and analyzing large data sets including user's activities, opinions, or preferences. It is a "user to user" correlation method, thus predicting items to the target user by finding similarities with other users and their preferences [4].

Sometimes Content-based filtering is used to support collaborative filtering, which is called hybrid-based filtering [5] to get better results and overcome limitations. The hybrid technique was created to achieve better performance and overcome the disadvantages of traditional recommendation techniques. This technique combines the better part of two or more techniques. The hybrid technique will be the most accurate since it will be the intersection of the Content-based and Collaborative filtering technique results.

2. Related Work

This section briefly describes different recommendation systems that use content-based filtering, collaborative filtering, and hybrid filtering for recommending courses to students [6]. Several projects use either content-based or collaborative filtering for the recommendation [7]. Identifying people's preferences and making decisions becomes a big challenge to build recommender systems that cope with information overload. Extracting users' data can be done implicitly by identifying users' actions or explicitly by collecting users' ratings.

A lot of research has been conducted on building a course recommendation system that will help users to find a course of their choice without much effort. In one of the first recommender systems, Tapestry [8] used collaborative filtering (CF) to calculate the relationship between the user and help target users filter retrieve similar items. Konstan [9] et al. proposed a mathematical model that by using the user project scoring matrix calculated the similarity between users, searching for the nearest neighbor from the user set of the target users, thus obtaining the recommendation result. Breese et al. confirmed that popular items have an important impact on recommendation results and integrated them with many improvement ideas.

According to Ping-Lung Hsin, colleges should offer more courses in career development and workplace competency building to improve students' capacity for workplace adaptation and form a sound understanding of career development [10]. A personalized recommendation system for online pages was created by Wang and Shao. The hierarchical bisecting medoids algorithm is used to cluster time-framed navigation experiences, and association mining is then used to analyze those navigation sessions in order to generate recommendations for the future. When recommending courses, Polyzou et al. employed the Markov chain-based collaborative filtering technique. Association rules were utilized by Ma et al. to suggest college electives. A genetic algorithm-based multicriteria hybrid recommendation approach was proposed by Esteban et al.

The collaborative recommender approach put out by AlBadarenah and Alsakran [11] suggests elective courses based on how similar students are to one another. They utilized the K-means algorithm to group comparable students, the K nearest neighbor method to choose the cluster that most closely resembled the target student, then association rule mining for each cluster [12]. The hybrid recommendation system that Wen-Shung Tai et al. devised combines SOM with data mining methods. They utilized a dataset from a course with 1000 students, clustered the students using SOM neural networks based on similarity, and then generated association rules for each cluster using the Apriori method [13]. A hybrid based personalized course recommendation system with text content and an N-gram model was proposed by Gulzar et al. [14]. Based on the knowledge gap, Khosravi suggested a peer learning recommendation system. Four alternative course recommendation methods that are all connected to the course order were proposed by Wang et al.

To recommend courses in E-Learning Systems based on historical data, Aher and Lobo combined machine learning algorithms such as the Simple K-means algorithm for clustering and the association rule using the Apriori algorithm [15]. A course recommendation system based on students' academic achievement and cognitive ability was proposed by Upendrana et al. [16]. Their system was created to assist students in selecting the courses that would benefit their grades the most. It recommends several courses with a high probability of success in terms of grades.

Research says that the course recommendation system suggests to the students the best combination of courses in which they are interested and the similarity between them which may, in turn, help them to identify individuals sharing a common taste of domains/courses and collaborate with them for making projects/doing research work on related topics. The existing recommendation systems recommend courses to users based on various similarities between them. This paper presents a recommendation system that recommends courses to students based on the similarity between courses taken by the different students and their close relatedness to each other.

3. Proposed System

An integrated E-learning system is developed for implementing a course recommendation system. Users can either enroll as a teacher or a student. Students can register and select preferable technologies. Teachers, after registration, can create courses for different domains and subdomains which are then used for recommendations. Once registered, courses are recommended to the students according to their preference using content-based filtering.

Although these recommendations may not be accurate, students can still have a good experience since the subdomains of the recommended courses are of interest to students. Once the students start to enroll in a few courses and increase their engagement with our system, collaborative based filtering can be used to recommend courses based on other students' enrollment histories. The combined effect of content-based and collaborative-based to recommend courses using the hybrid technology will be more accurate and targeted since courses that are enrolled by similar students and are best suited according to students' profiles are recommended. The hybrid technique will help students in finding the best courses in an easy, effective and efficient way. Students can easily focus on learning and developing new skills rather than wasting time on finding the courses. Due to this, the interest and engagement of students are intact.

4. Algorithms

Recommender systems play a vital role in increasing user engagement and benefiting the business. Various recommendation techniques like trust-based, context awareness-based, social network-based, fuzzy set-based, knowledge-based, etc., are available using which we can find patterns in behavioral data of the customers [1]. For our purpose, we will be focusing on the traditional techniques: Content-Based Filtering, Collaborative-Based Filtering, and Hybrid Filtering.

A. Content-Based Filtering

Content-Based filtering recommends products based on user preferences. It compares the features of two or more objects to find the similarities between them. It is used as a standalone technique when we have less information about the user. For our work, Content-Based filtering compares the sub- technologies in which a particular user is interested with categories of the courses. So, by using the traditional method of data extraction like cosine similarity, recommendations can be generated. The probability of recommending a course to a user can be calculated as the categories of course and sub-technologies of the user in common divided by the total number of categories of the course as represented by equation 1.

$$P\,(U, C) = \frac{(a \cap b)}{\sum a} \tag{1}$$

- P (U, C) = Probability of recommending a course C to user U
- a = Set of all the categories of the course C
- b = Set of all sub-technologies preferred by the user U

The probability value ranges from 0 to 1. The courses with higher values of probability are recommended to the users.

In Figure 1, two courses with associated sub-domains are displayed along with the user and his preferable sub- technologies. Out of the two available courses, the course which has a higher intersection of its sub-domains with the sub-technologies of user is recommended to the user.

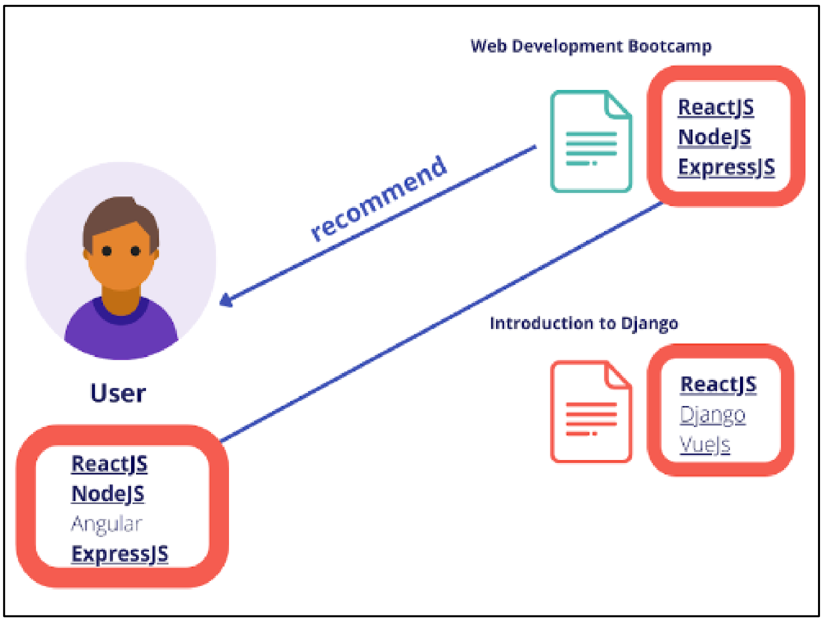

Figure 1. Content Based Filtering

B. Collaborative-Based Filtering

Collaborative-Based filtering considers the opinion of other users who share similar purchase histories for recommending products. Collaborative filtering can be achieved by considering the similarities between different users or items.

- In a user-based approach, products that similar users purchase are recommended.
- In an item-based approach, products are recommended based on similar items purchased by users.

In the proposed system, we have implemented a user-based approach. It is a two-step process in which the first step is to find the similarities between the objects followed by finding out the probabilities for recommendations. Since we are using the user-based approach for our work, we have to calculate the similarities between different users of our system to provide recommendations based on similar users' purchase histories. To calculate the similarity of users, various techniques like Pearson correlation-based similarity, constrained Pearson correlation-based-similarity, cosine-based similarity, or even adjusted cosine-based measures can be used [1].

1. Calculate the similarity value

The similarity value between the users can be calculated using various parameters. For our work, we consider two parameters: preferable sub technologies and courses enrolled by the user. The similarity value is calculated by considering the weighted average of two parameters as represented by equation 4.

$$x = \frac{(c1 \cap c2)}{(c1 \cup c2)} \tag{2}$$

- c1 = Courses enrolled by user U1
- c2 = Courses enrolled by user U2

The numerator of equation 2 is the intersection of the courses enrolled by two users. And denominator corresponds to the union of all the courses enrolled by either of the users.

$$y = \frac{(s1 \cap s2)}{(s1 \cup s2)} \tag{3}$$

- s1 = Sub technologies preferred by user U1
- s2 = Sub technologies preferred by user U2

The numerator of equation 3 is the intersection of the sub-technologies preferred by two users. And denominator corresponds to the union of all the sub technologies preferred by either of the users.

$$S\,(U1, U2) = \frac{(0.5 \times x) + (0.5 \times y)}{(0.5 + 0.5)} \tag{4}$$

- S (U1, U2) = Similarity value between two users U1 and U2

2. Probability of Recommending a Course to a User

In this step, we use the similarity value of the previous step to find the probability of recommending the courses to the users. For a course C that is not yet enrolled by user U, the probability of recommending it is given by:

$$P\,(U, C) = \frac{z}{M} \tag{5}$$

- P (U, C) = Probability to recommend course C to user U
- Z = Sum of similarities of User U with all the users who have enrolled for course C
- M = Number of users who have enrolled for course C

The numerator of equation 5 is the sum of similarities of all the users with user U who have already enrolled in course C. The denominator is the total number of users who have enrolled in course C. The higher the similarity between users, the higher the probability of recommending courses of similar taste.

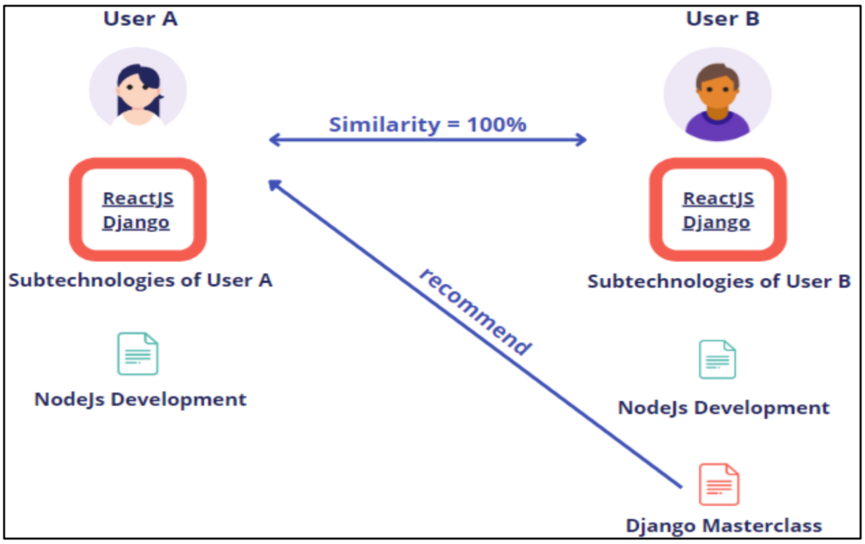

Figure 2. Collaborative Based Filtering

In figure 2, User A and User B have enrolled in the same course "NodeJS Development". Also, both users have similar preferred technologies. So, the similarity value is 100% between both the users. Now, when User B has enrolled in the course "Django Masterclass", it is recommended to User A based on collaborative-based filtering.

C. Hybrid-Based Filtering

Hybrid-based filtering is used to make more accurate recommendations by amalgamating the advantages of two or more techniques. To achieve better performance, a hybrid recommendation system is implemented by combining the Collaborative filtering technique with other techniques, to avoid cold-start, sparseness, and problems of scalability [1]. For our paper, the results of collaborative and content-based filtering are combined to provide the recommendations which are best suited for users according to their profiles.

$$P\ (U, C) = \text{Content Filtering} \cap \text{Collaborative Filtering} \qquad (6)$$

Hybrid-Based Filtering is intersection of the results of content and collaborative filtering to provide more accurate results. It is represented as equation 6.

5. Experimental Results

For testing the implemented algorithm, we have developed a system with a user interface for students to view recommended courses, enroll in them, and learn new skills. The backend of the system encompasses the implemented Hybrid course recommendation algorithm. For the experiment, we have considered eight courses of the Web Development domain. Each course is associated with a few sub-domains. Fifteen students took part in the experiment. They were asked to register in the system and select domains and subdomains according to their personal preferences. The results of the experiments are provided ahead.

Table 1. Details of Courses

Course ID	Course Name	Course Sub-Domains
Course 1	Introduction to Web Development	React, Node JS, Express JS
Course 2	Backend Bootcamp	Node JS, Vue JS, Express JS
Course 3	Backend Development	PHP, Laravel, Angular
Course 4	Web Development Bootcamp	Angular, Node JS
Course 5	Web Development Tutorial	Vue JS, Node JS, Laravel
Course 6	Front End Tutorial	Flask, React
Course 7	Full Stack Web Development	Vue JS, Angular, PHP, Spring
Course 8	Key to Web Development	React, PHP, Express JS

Table 1 displays the details of all the courses available on our system.

Table 2. Domains and Sub-domains selected by users

User ID	Domain	Sub-Domains
User 1	Web Development	React, Node JS, Express
User 2	Web Development	React, Vue JS, PHP
User 3	Web Development	PHP, Laravel
User 4	Web Development	Node JS, Express JS, Angular
User 5	Web Development	Vue JS, Node JS, Laravel
User 6	Web Development	Flask, React
User 7	Web Development	PHP, Angular, Spring
User 8	Web Development	Angular, Node JS
User 9	Web Development	Express JS, React
User 10	Web Development	Vue JS, PHP
User 11	Web Development	Laravel, Angular
User 12	Web Development	React, PHP
User 13	Web Development	React, Angular, Spring
User 14	Web Development	Angular, Express JS, Laravel

User 15	Web Development	Flask, Node JS, Laravel

Table 2 displays the details of domains and subdomains of different users according to their personal preferences. For domain, we have restricted ourselves to just Web development.

Table 3. Courses selected by users

	Course 1	Course 2	Course 3		Course 8
User 1	✓		✓		✓
User 2		✓	✓		
User 3	✓				
User 4		✓			✓
User 5					✓
User 6		✓			
User 7	✓			....	
User 8				....	
User 9	✓				
User 10					
User 11			✓		
User 12		✓			✓
User 13	✓				
User 14	✓		✓		✓
User 15		✓	✓		

Table 3 displays the enrollment histories of different users of our system. For example, User 1 has enrolled in courses 1, 3, 6, and 8.

Table 4. Similarity value of User 1 with all other users

User	Common Courses	Total Courses	Similarity Value
User 2	2	5	0.4
User 3	2	5	0.4
User 4	1	5	0.4
User 5	2	4	0.5
User 6	0	7	0
User 7	1	4	0.25
User 8	1	5	0.2
User 9	2	6	0.33

User 10	0	5	0
User 11	1	5	0.2
User 12	2	5	0.4
User 13	1	6	0.16
User 14	2	5	0.4
User 15	3	5	0.6

Table 4 displays the similarity values for user 1 with all other users.

For example, users 1 and user 2 have enrolled in course 3 and course 6 in common. User 1 has also enrolled in course 1 and course 8, whereas user 2 has enrolled in course 2. So, two courses are enrolled by both and a total of five courses are in union. So, similarity is calculated as 2/5 which is 0.4. So, similarly, the similarity of user 1 with other users is calculated. Now user 1 has not yet enrolled in course 2, course 4, course 5, and course 7. So, recommendations can be made for these courses. Course 2 is already enrolled by user 2, user 4, user 6, and user 12. So, the sum of similarities of user 1 with all these users is 0.4+0.4+0+0.4 = 1.2. The probability of recommending course 2 to user 1 is 1.2/4 = 0.3. Similarly, for course 4, the probability is 0.218, for course 5, it is 0.12 and for course 7, it is 0.3325. So, course 7 and course 2 have the highest probabilities of user 1 enrolling in it. So, these courses were recommended to user 1 during experiment. Likewise, all 15 users were recommended courses.

6. Conclusion

Recommendation systems are very useful in this knowledge exploration age. With ever-increasing data, it becomes very difficult to find what we are looking for without any filters or recommendations. In e-learning websites, recommendation systems play a vital role in protecting the interest of learners in gaining knowledge. Course Recommendation system = can help in avoiding frustration and wastage of time by providing useful and interesting recommendations. In this paper, content-based recommendations along with collaborative filtering are used to provide accurate and targeted recommendations. The concluding recommendation results are based on the information provided by the user as well as the similarity with other users. Hence, users can easily find appropriate and best-suited courses according to their interests and focus on the learning process.

7. References

[1] T. Z. Islam et al., "College Life is Hard! - Shedding Light on Stress Prediction for Autistic College Students using Data-Driven Analysis," 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), 2021, pp. 428-437, doi: 10.1109/COMPSAC51774.2021.00066.

[2] B. Ma "Design of an Elective Course Recommendation System for University Environment" educational data mining 2019.

[3] J. A. Konstan B. N. Miller D. Maltz J. L. Herlocker L. R. Gordon and J. Riedl "GroupLens: applying collaborative filtering to Usenet news" Communications of the ACM vol. 40 no. 3 pp. 77-87 1997.

[4] V. Garg and R. Tiwari, "Hybrid massive open online course (MOOC) recommendation system using machine learning," International Conference on Recent Trends in Engineering, Science & Technology - (ICRTEST 2016), 2016, pp. 1-5, doi: 10.1049/cp.2016.1479.

[5] Z. Chen, X. Liu, and L. Shang, "Improved course recommendation algorithm based on collaborative filtering," 2020 International Conference on Big Data and Informatization Education (ICBDIE), 2020, pp. 466-469, doi: 10.1109/ICBDIE50010.2020.00115.

[6] R. Obeidat, R. Duwairi and A. Al-Aiad, "A Collaborative Recommendation System for Online Courses Recommendations," 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), 2019, pp. 49-54, doi: 10.1109/Deep-ML.2019.00018.

[7] J. L. Herlocker J. A. Konstan L. Terveen and J. Riedl "Evaluating collaborative filtering recommender systems" ACM Transactions on Information Systems vol. 22 no. 1 pp. 5-53 2004.

[8] Z. Gulzar A. A. Leema and G. Deepak "PCRS: Personalized Course Recommender System Based on Hybrid Approach" Procedia Computer Science pp. 518-524 2018.

[9] H. Thanh-Nhan H. Nguyen and N. Thai-Nghe "Methods for building course recommendation systems" 2016 Eighth International Conference on Knowledge and Systems Engineering (KSE) pp. 163-168 2016.

[10] Chung-Yi Huang, Rung-Ching Chen and Long-Sheng Chen, "Course-recommendation system based on ontology," 2013 International Conference on Machine Learning and Cybernetics, 2013, pp. 1168-1173, doi: 10.1109/ICMLC.2013.6890767.

[11] A. Al-Badarenah and J. Alsakran "An Automated Recommender System for Course Selection" International Journal of Advanced Computer Science and Applications vol. 7 no. 3 2016.

[12] S. Parvatikar and B. Joshi "Online book recommendation system by using collaborative filtering and association mining" 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC) pp. 1-4 2015.

[13] Z. Chen, X. Liu and L. Shang, "Improved course recommendation algorithm based on collaborative filtering," 2020 International Conference on Big Data and Informatization Education (ICBDIE), 2020, pp. 466-469, doi: 10.1109/ICBDIE50010.2020.00115.

[14] Y. H. Wu and E. H. Wu, "AI-based College Course Selection Recommendation System: Performance Prediction and Curriculum Suggestion," 2020 International Symposium on Computer, Consumer and Control (IS3C), 2020, pp. 79-82, doi: 10.1109/IS3C50286.2020.00028.

[15] S. Aher and L. Lobo, "Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data", Knowledge-Based Systems, vol. 51, pp. 1-14, 2013.

[16] D. Upendran, S. Chatterjee, S. Sindhumol and K. Bijlani, "Application of Predictive Analytics in Intelligent Course Recommendation", Procedia Computer Science, vol. 93, pp. 917-923, 2016.