A course recommendation system based on grades

Conference Paper · March 2020

DOI: 10.1109/ICCSEA49143.2020.9132845

4 authors, including:

Bhaskar Mondal
National Institute of Technology Patna

Priyadarsan Patra
DIT University

A course recommendation system based on grades

Bhaskar Mondal*, Om Patra†, Sanket Mishra‡ and Priyadarshan Patra§
Xavier School of Computer Science and Engineering
Xavier University Bhubaneswar
Bhubaneswar, India
Email: *bhaskarmondal.cs@gmail.com, †ompatra10@gmail.com, ‡ucse17009@stu.xub.edu.in, §ppatra@xub.edu.in

***Abstract*—The online courses are playing a crucial role in developing new skills in learners and in the education system. Now a days a massive number of online courses and certifications are available over the internet from universities as open learning platforms. As there is no in-person consultation with any expert, the learners may opt for irrelevant courses inadvertently and may not be able to analyze their own suitability and adaptability of the courses which will west learners time and resources. This paper proposes a machine learning approach to recommend suitable courses to learners based on their learning history and past performance. The framework first classifies a new learner based on their past performance using the k-means clustering algorithm. Collaborative filtering will be applied in the cluster to recommend a few suitable courses. Further, based on an online test the adaptability of the learner will be tested to the customized recommended courses according to learners needs. The framework will provide a personalized environment of study to each learner.**

***Index Terms*—Collaborative Filtering, K-nearest neighbour algorithm, Non-negative Matrix Factorization, Cosine similarity.**

I. Introduction

The main problem for today's learners is that they need tailored access to factsstatistics based on preferences and requirements. To address this issue, the recommender system is used to analyse data automatically according to the user preferences and the most suitable one is presented in a plethora of different alternatives [1]. To personalize data, the recommender system is used either to acknowledge a comparable user or to identify specific items of user's concern. RS prioritize information, linked to items, and provided users with significant suggestions to their interest.

In India, a few online learning platforms or massive open online course (MOOC) like national program on technology enhanced learning (NPTEL), active learning platforms for young aspiring minds (SWAYAM) are developed by the government of India. These MOOC can lead the learners to potential carrier failure without proper guidance and irrelevant course chosen by the learners. A CRS may help the student to choose correct courses and the personalized environment will able to engage the learner to the framework.

A recommender system is an intelligent system that recommends a personalized set of information extracted from a dynamically generated huge volume of data. The recommender systems mainly filter the information based on user rating or opinion on some item called collaborative filtering or previous users' choice known as content-based filtering. Content-based filtering recommends the choice opted by some similar person in the past [2] [3]. Clustering is the machine learning technique to classify a data set into a finite number of groups based on similarity among data. These methods are unsupervised learning technique and mainly used for classifying unlabeled historical data. In the classified groups, members in each group are similar to each other and dissimilar to the members of other groups [4]. Frequent pattern mining is a data mining technique to extract the set of items chosen together frequently in the past from the historical transactional data sets. The course opting by the student is a transactional database. A recommender system may be proved extremely beneficial for the online open university environment. It can filter out and recommend a suitable set of courses to the Learner which will save time and keep the learner's aspiration high. A personalized environment will help the learners to adopt the concepts more easily. Therefore, the framework will help the learner to achieve the course outcomes and skills. A typical course recommender system is presented in Fig. 1. The proposed framework follows a machine learning (ML)

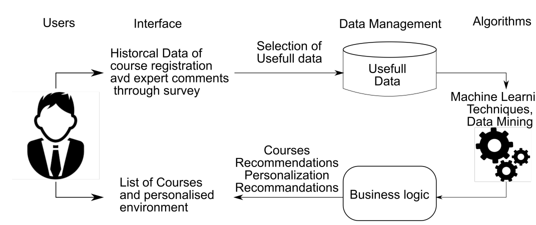

Fig. 1. A typical course recommender system

approach to recommend suitable courses to learners based on their learning history and past performance. The framework first classifies a new learner based on their past performance using the k-means clustering algorithm. A few courses will be recommended using Apriori frequent pattern mining algorithm on the historical student data. Further, based on an online test, the adaptability of the learner will be tested, and customized courses are then recommended based on learner's needs. The framework will provide a personalized environment of study for each learner, which includes addition or deletion of some preliminary or advanced topics, frequency of tests, level of test questions' hardness and recommendation of a suitable course after completion of the current one.

Most of the universities and online learning platforms are widely offering Internet-based learning facility to the students.

In this situation, the most important concern is the adaptability and understandability of the courses to the learners. There is a tremendous chance that a learner may opt for an irrelevant course and lose the egger of study in the environment. Therefore, a learner-centric dynamically personalize framework for the Lerner is the need of the hour. A personalized framework will help the learners to opt for most suitable courses and by personalization, it will keep the Lerner engage to study and help to adopt skills by completing a course faster.

II. Data Collection and Preparation

A. Collection

The data collection is the initial step of research in which the data is collected from the universities database. The dataset had 300 student records and 22 variables or attributes. Of the 48 variables, 12 key parameters are considered for better result.

B. Data Pre-processing

At first, our dataset was composed of 300 student details. Then we recognized that there were variables that didn't have all the data and they were all in their raw, so the unavailability of features was the main reason for removing some student details from our dataset. Most of the student didn't have subject data available. Lower Case Conversion

Since many individuals can jolt the same thing down differently, it can be difficult to process character data. The significant criterion for selecting a feature is matching the string We have converted our text into a lower case with abbreviations to prevent any ambiguity.

- Removing Punctuation: Redundant commas, question marks, and other special symbols are omitted from the data set to clean up the data.
- Striping White spaces: In this step all the text data is preprocessed and cleaned off with all the unnecessary white spaces, tab from the text.

III. Related Works

In recent years a couple of research have been published, most of the system among them are considering a university course registration. The previous frameworks are totally silent on the personalization of the learning environment for each course. The recommender systems are mainly using content-based filtering, collaborative filtering [5] and pattern mining or knowledge mining [6]. Badarenah et. al. [7] (2017) have proposed an elective course recommender system. The system recommends elective courses to the university students based on similar historical data. The collaborative recommender system uses rough set theory for frequent pattern mining and k means clustering algorithm for grouping the students' data. Finally, the courses are recommended using association rules. Ng et. al. [8] (2017) proposed a course recommendation system with a professor for college students in regular mode. The proposed CRS mainly depends on topic, sentiment and tag analysis for which the system needs good and reliable survey data. Collecting reliable survey data for an open university framework is hard and may lead to an inefficient system. Bakhshinategh et. al. (2017) have proposed a course recommendation system based on graduate attributes. Grewal et. al. [9] [9] (2016) proposed a CRS based on learner's previous subject knowledge and future interest. The CRS is designed for prospective students to the graduate programs. The system works based on clustering of historical data. Bhumichitr et. al. [10] (2017) developed a framework for elective course recommendation to the university system in regular mode. The CRS works based on Correlation Coefficient among the student to measure similarity and Alternating Least Square (ALS). The authors claim that 86% accuracy of their system. Jinjiao et. al. [11] (2018) have proposed an intelligent CRS for Chinese universities. The intelligent CRS works based on the sparse linear method and requires previous course enrollment data. Gulzar et. al. [12](2018) have proposed a CRS for the Ph.D. research students. The CRS is based on Ontology and N-gram query classification which means the system can read the learners requirements and suggest courses. The author's claims more than 95% of accuracy. Cheng et. al. [13] (2018) have proposed another ontology-based hybrid CRS. The system collects information from multiple sources collaborative filtering with the filtering of content for a recommendation. [14] (2016) have proposed a new custom algorithm for course sequence recommendation as the sequence of study of courses have an important role in adaptability to the topics. The system uses a forward-search backward-induction algorithm which needs a previously stored course with prerequisites. Bodily et. al. [15] (2017) have presented a brief review of online learning recommendations and dashboards. The authors have tried to highlight the usability of recommender systems on student's behavior, achievements and skill development. However, our proposed framework follows a machine-learning approach to recommend adequate courses to learners depending on their learning history and past performance (Garde Point average) in academics and is more effective. The method first classifies a new learner based on their past results using the k-means cluster algorithm. A few courses will be recommended using Apriori's frequent pattern mining algorithm for historical student data.

IV. Preliminaries

In this section, the preliminaries of the required techniques and algorithms are discussed in short.

A. Collaborative Filtering (CF)

CF techniques analyse huge amounts of data of the preferences and train the system to predict a new users preference based on historical data of similar users. These systems can recommend a complex preferable item without any knowledge of the item recommended. The performance of the CF-systems depends heavily on the similarity used in; similarity majors like Pearson Cosine Similarity, Correlation Coefficient, and Euclidean Distance are widely used to estimate similarity between users. A similarity major is chosen based on the data available in the repository.

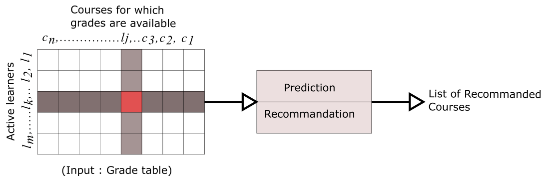

Fig. 2. A typical process of Collaborative Filtering

The CF algorithm can recommend one or a set of new learners' courses based on previous performance in different courses and background. In a typical CF scheme, there is a list of m learners $L = l_1, l_2, \ldots, l_m$ and a list of N courses $C = \{c_1, c_2, \ldots, c_n\}$. Every learner l_i already pursued a list of courses C_{li}, and the grade of the corresponding courses.

The grades are given directly by the trainer or instructor, typically within a fixed numerical scale, or implicitly can be found from the background data of the learner. We must consider that $c_l i \in C$ and C_{li} can be a null set. There exists a distinguished active learner $l_a \in L$ for whom the likeliness of some course to be determined by the CF algorithm and it may be in two ways as follows. Prediction gives a numerical value, $P_{(l,j)}$ represents the likeliness of course $c_l i \in C_{(l,i)}$ for a learner l_i. A predicted value is a value within a range (e.g. from 1 to 10) provided that the teacher gives the grades. On the other hand, the recommender system gives a list of N courses, $C_i \in C$, that the learner will like the most. The list of courses recommended must be already opted by the active learner, i.e., $C_i \cup C_{ji} = \phi$. Due to this reason CF algorithms is also called as Top-N recommender system. There are mainly two types of CF algorithms, namely Collaborative Filtering Algorithms based on Memory and Collaborative Filtering Algorithms based on Model.

Memory-based CF algorithms utilize the full user-item database to predict. These algorithms are using statistical methods to find a user set, called as neighbours, who have a past-record of similar behaviour with the target user. After neighbourhood formation, these techniques use different algorithms to combine neighbourhood preferences to produce a prediction or top-N recommendation for the active user. Such methods are therefore also regarded as user-based collaborative filtering or nearest-neighbour. There are two types of memory-based CF algorithms; user-user CF and item-item CF. Model-based CF algorithms propose a list of items based on a user expectations model, using a probabilistic approach and imagining the CF process to determine the expected preferential value of each item, given the user preference on other items. For constructing these models, machine learning algorithms such as clustering, the Bayesian network, and rule-based techniques are used. The clustering techniques tackle CF as a classification problem and create clusters of similar users and estimate the probability that a new user would be in a specific class C and based on that it estimates the conditional probability of preference. A probabilistic model for CF problems is generated by the Bayesian network techniques. The rule-based techniques use association rule mining algorithms to find the most common sets and then recommend a confidence-based set of frequently used items.

B. Content-based Recommendations:

A recommender system based on content makes use of the data provided by the user either explicitly (grade, preferences, rating) or implicitly (keywords, tags). A user profile is crated based on that data, which is used to make recommendations to the user. Term Frequency (TF) and Inverse Document Frequency (IDF) are widely used to determine a particular article's / document's/ course's/ movie's/ news item's etc. relative importance.

TF is the occurrence of a word in a document and IDF is the logarithmic inverse of the document frequency among the whole corpus of documents. For example, if we search for "the use of recommender system" on a search engine. It is expected that 'the' will have higher frequency that 'recommender' in the corpus. But, 'recommender' is the most important part for the search query. In such situation the TF-IDF ignores the effect of unimportant high frequency words. In TF-IDF, log is used to reduce the effect of high frequency words. For example: TF = 5 vs TF = 3 is massively diverse from TF = 1000 vs TF = 100. So, the relevance of a word in a document cannot be determined by raw count. For an article a with total words $|a|$ and k is a key word, the TF is defined as Eq. 1

$$TF_{k,a} = \frac{frequency(k)}{|a|} \tag{1}$$

Therefore, it's represented by the Eq. 2 with a weighted term frequency.

$$w_{k,a} = \begin{cases} 1 + \log_{10} kf_{ka} & \text{if } tf_{lc} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The weights are calculated by Eq. 3

$$w_{k,a} = \frac{kf_{ka} \log \frac{N}{af_k}}{\sqrt{\sum((kf_{ka})^2 (\log \frac{N}{af_k})^2)}} \tag{3}$$

Assume a set C of x courses with m binary attributes and a vector $R_{u \times v}$ where r_i indicates the grade for course C_i. For each attribute c_j the IDF is calculated. To determine the preferability of a learner in an attribute, user profile for each attribute is computed.

Learner profile P is a vector comprising a weighted value for each attribute for a specific learner. The value of an element p_i of the vector P represents the learner's grade or preference for the attribute k_i based on the background of the learner. The learner profile for an attribute is the dot product of the attribute value for all the items in C with the learner's grade vector V which ia given by Eq. 4.

$$l_j = \sum_{x=1}^{j} u_x (kaf)_x C_{kx} \tag{4}$$

For course c_j in n. l_j is a numeric value representing preference for each course in the set of course C. A higher value of l_j indicated higher preferability for the course.

V. Proposed Method

The framework will work based on historical and survey data. The first step to this is collecting data. Then the collected data need to go through a clean and selection process. Only useful data sets will be chosen for the use of the framework. The selected data will then need to normalize which includes the integration of data from heterogeneous sources followed by scaling. The final data set is denoted by D. A clustering algorithm will be applied to the data set D to create a group of similar learners. Once the clusters of learners are created a frequent pattern mining algorithm will be applied to each of the clusters. The system classifies the students based on historical data by finding out what was the background of those students who scored higher grade in each course. Every time a new student will enter in the system will be classified using the clusters and a set of courses will be recommended to the learner based on frequent pattern mining. On the other hand, an online test will be conducted for each learner to understand their current expertise based on which the courses will be personalized for a new learner. The personalization will use the previously collected subject expert's comments and association rule mining. The framework for course recommendation and personalized online learning is presented in Fig. 3.

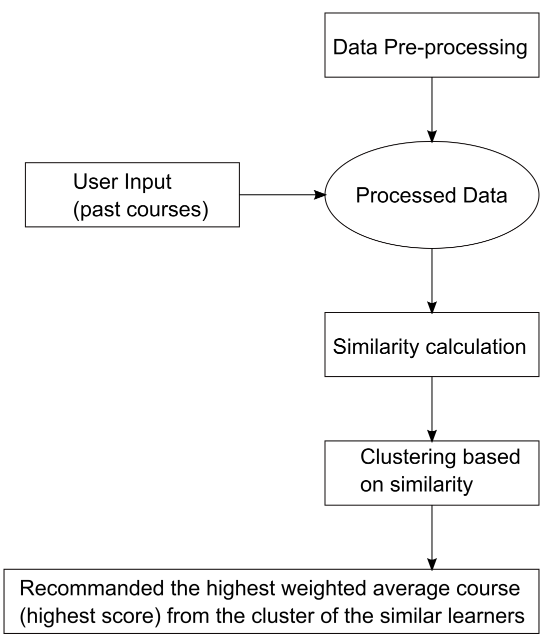

Fig. 3. A typical process of Collaborative Filtering

VI. Results and Discussions

The efficiency of the proposed system is discussed in this section. The heat map in 4 depicts the difference of predicted score and actual score of some random learners

A. Root means square error (RMSE)

Two criteria to evaluate predictive accuracy are used, the root-mean-square error (RMSE) and the mean absolute error.The RMSE is being used to calculate model performance

Fig. 4. difference of predicted score and true score of some random learners

in research studies as a standardized statistical metric as well as another useful measure widely reviewed for model assessments is the mean absolute error (MAE). The RMSE and MAE are defined as Eq. 5 and Eq. 6 respectively.

$$RSME = \sqrt{\sum_{i=1}^{n} \frac{(\widehat{y_i} - y_i)^2}{n}} = \sqrt{\sum_{i=1}^{n} e_i^2} \tag{5}$$

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n} \tag{6}$$

where e_i denotes $(y_i - x_i)^2$, $(y_i - x_i)$ where $y_1, y_2, \ldots, y_n$ are predicted values $x_1, x_2 \ldots, x_n$ are the observed values and n is the number of observations. The calculated values of MSE is 3.6092737, MAE is 1.133 and RMSE is 1.8998089.

B. Precision and recall

These are two quality and accuracy measure of a recommender system. The precision is the ratio between correctly recommended items to first n recommended items. Recall is the ratio between the number of correctly recommended items to all m recommended items. The precision vs recall is ploted in Fig. 5.

The Precision-Recall curve is a graph with Precision values for the y-axis and the Recall values for the x-axis. The high region under the curve reflects both high recall and high precision, in which high precision is associated with a low false-positive rate, and high recall is associated with a low false-negative rate. High scores in both indicate that the learning algorithm returns correct results as well as preserving the majority of all positive outcomes A high-recall but low-precision system yields several results, and most of its predicted labels are inaccurate especially in comparison to training labels. A machine with high precision but the low recall is indeed the contrary, producing few results, and most of its labels are correct when compared to training labels. A best possible situation with high precision and high recall produce multiple results, with all results labeled correctly.

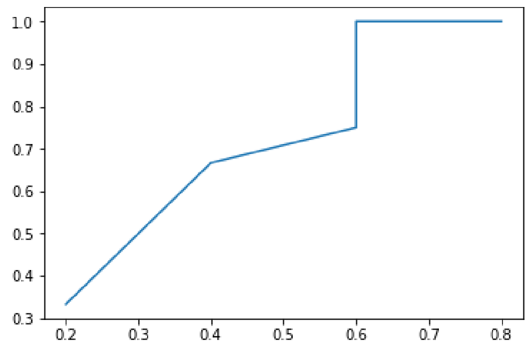

Fig. 5. Plot of Precision vs recall

VII. Conclusion

It is proposed to take into consideration the recommendation system by overcoming the limitations of the current approach to the individual recommendation system. We have compared three techniques through experiments and found that the proposed system is more suitable, more effective and more beneficial to the learners. There are some common reasons to implement a recommendation system including user satisfaction and another reason to increase the platform's fiscal success. Added to generate recommendations from knowledge base in future neighbourhood generation. The aim to use neighbourhood formation is to identify other related learners based on their area of interest and target learner needs.

References

[1] B. Mondal, *Artificial Intelligence: State of the Art*. Cham: Springer International Publishing, 2020, pp. 389–425. [Online]. Available: https://doi.org/10.1007/978-3-030-32644-9-32

[2] F. Isinkaye, Y. Folajimi, and B. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egyptian Informatics Journal*, vol. 16, no. 3, pp. 261–273, 2015.

[3] N. Pritam, M. Khari, L. Hoang Son, R. Kumar, S. Jha, I. Priyadarshini, M. Abdel-Basset, and H. Viet Long, "Assessment of code smell for predicting class change proneness using machine learning," *IEEE Access*, vol. 7, pp. 37 414–37 425, 2019.

[4] B. Mondal and J. P. Choudhury, "A comparative study on k means and pam algorithm using physical characters of different varieties of mango in india," *International Journal of Computer Applications*, vol. 78, no. 5, 2013.

[5] M. R. McLaughlin and J. L. Herlocker, "A collaborative filtering algorithm and evaluation metric that accurately model the user experience," in *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2004, pp. 329–336.

[6] L. Esmaeili, M. Nasiri, and B. Minaei-Bidgoli, "Personalizing group recommendation to social network users," in *International Conference on Web Information Systems and Mining*. Springer, 2011, pp. 124–133.

[7] A. Al-Badarenah and J. Alsakran, "An automated recommender system for course selection."

[8] Y. Ng and J. Linn, "Crsrecs: A personalized course recommendation system for college students," in *2017 8th International Conference on Information, Intelligence, Systems Applications (IISA)*, Aug 2017, pp. 1–6.

[9] G. DS and K. K, "Developing an intelligent recommendation system for course selection by students for graduate courses," *Business and Economics Journal*, vol. 7, no. 2, pp. 1–9, 2016. [Online]. Available: https://www.omicsonline.org/open-access/developing-an-intelligent-recommendation-system-for-course-selectionby-students-for-graduate-courses-2151-6219-1000209.php?aid=73333

[10] K. Bhumichitr, S. Channarukul, N. Saejiem, R. Jiamthapthaksin, and K. Nongpong, "Recommender systems for university elective course recommendation," in *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, July 2017, pp. 1–5.

[11] J. Lin, H. Pu, Y. Li, and J. Lian, "Intelligent recommendation system for course selection in smart education," *Procedia Computer Science*, vol. 129, pp. 449 – 453, 2018, 2017 INTERNATIONAL CONFERENCE ON IDENTIFICATION,INFORMATION AND KNOWLEDGEIN THE INTERNET OF THINGS. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050918302424

[12] Z. Gulzar, A. A. Leema, and G. Deepak, "Pcrs: Personalized course recommender system based on hybrid approach," *Procedia Computer Science*, vol. 125, pp. 518 – 524, 2018, the 6th International Conference on Smart Computing and Communications. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050917328314

[13] T. Su, L. Cheng, Y. Wang, X. Liang, J. Zheng, and H. Zhang, "Time-optimal trajectory planning for delta robot based on quintic pythagorean-hodograph curves," *IEEE Access*, vol. 6, pp. 28 530–28 539, 2018.

[14] J. Xu, T. Xing, and M. van der Schaar, "Personalized course sequence recommendations," *IEEE Transactions on Signal Processing*, vol. 64, no. 20, pp. 5340–5352, Oct 2016.

[15] R. Bodily and K. Verbert, "Review of research on student-facing learning analytics dashboards and educational recommender systems," *IEEE Transactions on Learning Technologies*, vol. 10, no. 4, pp. 405–418, Oct 2017.