International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Accuracy Invoked Course Recommender System using Collaborative Filtering

A.K.Mariappan, R.Ramana Surriyan, E.S.Shobana,

Abstract: It is always important that the student should choose the right course. The decision of taking the right course is very important as the student's future depends on the course they opt to study. Most of the students are not much aware of the courses that are available in their own field of study. Selecting wrong courses might be due to the mismatch between the student's aptitude, training and mental ability. Thus, the idea is to develop a system for helping the students to choose a course which would be best suited for him/her based on features like previous student selection, interest, languages known etc. Existing research has explored recommender system using content-based filtering but it can only do limited content analysis and the recommendation will not be precise at the end. An attempt is made to improve the performance of this system using Collaborative based filtering techniques which will recommend a ranked list of courses. Under Collaborative filtering techniques, user based collaborative filtering and item based collaborative filtering is used. Sample student datasets from Kaggle, has been used to test the performance of our system.

Keywords*: *Recommender System, Content-Based Filtering, Collaborative Filtering

I. INTRODUCTION

There have been various inquires about led in the field of information mining to decide the presentation of the understudies. The recommendation system will help student to make the right choice, which best suites his/her. We are trying to provide a solution to this problem using the latest technologies and researches on recommendation system. Recommendation system will improve the teaching and learning process of a student by providing good solution by analyzing the best courses. In the last few years, this system provided solutions by recommendation for wide range of choices by focusing on the logical relationship. So that people could make the right choice.

Thus by implementing the personalized recommendation system, the stage can give progressively insightful help, understanding the more inventive supplies of the instructive administrations. Here it presents a brief overview of the existing literature on this topic. ML-based approaches like regression, SVM, clustering and algorithms to predict the course a student should opt for are discussed. This describes the methodology used to implement the recommendation algorithms. This includes preprocessing few input data from the user and with the help of the algorithm which would combine all the features and results in output with ranked list of courses which he/she should proceed. An overview of the experiments conducted and applied on the matrix to find the most efficient method through the process of cross validation are also described. Here this presents the results of the experiments, datasets and evaluation metric. The performance of the system using various algorithms and methods are tabulated. Also describes the interpretation of the results, and gives a formal conclusion to the project. To develop a recommender system for helping the students to choose an online course which would be best suited for him/her based on features like previous performance, skill set, ability to learn, etc. • Existing research has explored recommender system using content-based filtering but it can only do limited content analysis and the recommendation will not be precise at the end. An attempt is made to improve the performance of this system by using Collaborative based filtering.

II. LITERATURE SURVEY

2.1 Recommendation system analysis

The prescient investigation is a report review [1] discusses the course proposal structure to find the courses which are skilled for an understudy looking for after admission to the school. Routinely, the conjecture relies upon the calling's target or the present business design. In this system proposed, the estimate is characterized subject to the assessments acquired by the understudy in twelfth standard; which is taken as a sign of the past insightful introduction and mental limit of the understudy. A model is generated from the bequest knowledge or data from the scholars UN agency have completed the course with success. This model [1] is employed for predicting the courses for brand new students. The concept behind this approach is that once a student with specific set of skills is triple-crown in an exceedingly course then another student with similar set of skills can have a better success chance within the same course. With the occasion of showing reorganization in resources and colleges, the applying of the framework is AN unavoidable pattern.

Manuscript published on November 30, 2019.
* Correspondence Author
A.K.Mariappan*, Information Technology,Easwari Engineering College,Chennai,
R.Ramana Surriyan, Information Technology,Easwari Engineering College,Chennai,
E.S.Shobana, Information Technology, Easwari Engineering College,

On the possibility of the framework, understudies will pick the courses unreservedly and severally in accordance with their instructive mastermind. Be that as it may, there are a few issues seeing the framework as following. Anyway, do the researchers pick the courses in the agreement of forte information? Anyway, do the understudies develop their own gifted quality in the agreement with social prerequisites? Anyway, do the understudies orchestrate their instructive way in the understanding of capacities preparing? Thus, the piece key [2] of the framework is to develop pertinent and great capacities from resources and colleges. This paper utilizes two popular algorithms: collaborative based recommendation using Pearson Correlation Coefficient and Alternating Least Square (ALS) [3], and compares their performance on a dataset of academic records of university students. The experimental results show that applying ALS in this domain is superior to collaborative based with 86 percent of accuracy.

2.2 Content based technique analysis

So also, it had been conjointly discovered that a gaggle recommender framework for on-line course study (GRS-OCS) is given and this GRS-OCS [4] utilizes the verifiable information with respect to course evaluations and previous students. Every agreeable separating and Content-based sifting is consolidated in GRS-OCS. A fundamental examination has been directed to bolster the significant data of MOOC. Experimental examination and client overviews show the efficiency= and adequacy of GRS-OCS. Additionally, the recommender frameworks are read for a long time with different promising models been anticipated. Among them, helpful Filtering (CF) models are ostensibly the principal independent one gratitude to its high precision in proposal and disposal of protection included individual meta-information from instructing. This paper expands the utilization of the CF-based model to the errand in truth proposal. we tend to imply numerous difficulties in applying the overall CF-models to make a course proposal motor, together with the shortage of rating and data, the lopsidedness in certainty enrollment conveyance, and in this manner the interest in actuality reliance displaying. we watch out for then propose numerous ideas to deal with these difficulties. In the long run, we will in general blend a two-arrange CF model [6] standard by course reliance with a diagram based recommender upheld course-change organize, to accomplish United Self-Defense Force of Colombia as high as zero.97 with a genuine world dataset. Presently a day E-learning is getting prevalent. E-learning is basically the PC and system empowered exchange of information and expertise. E-learning is additionally alluded to as electronic instruction and e-educating. A few instances of business frameworks are Blackboard, WebCT, and Top-Class while a few instances of free frameworks are Moodle, Ilias, and Claroline [7]. In this paper, we propose the engineering for Course Recommender System and how the information moves through this framework. This system predicts the best combination of subjects i.e. the subjects in which students are more interested. Here we use the learning management system such as Moodle to collect the data from student regarding their course choices.With the appearance of web-based e-learning systems, an enormous quantity of instructional information is obtaining generated. This large information gave rise to massive data in instructional sectors. Currently, massive information analytics techniques are being employed to investigate these instructional data and generate completely different predictions and suggestions for college kids, teachers, and colleges [11]. Recommendation systems are already terribly useful in ecommerce, industry and social networking sites. Recently recommendation systems are verified to be economical for the education sector furthermore. During this work, we have a tendency to are employing a recommendation system for large information in education This work uses collaborative-filtering based mostly suggestion techniques to recommend offered courses to students, relying upon their grade points obtained in alternative subjects. We have a tendency to are mistreatment item-based recommendation of driver Machine learning library [11] on high of Hadoop to get set of recommendations. Similarity Log-likelihood is employed to find patterns among grades and subjects. Root Mean sq. Error between actual grade and counselled grade is employed to check the advice system. The output of this study may be utilized by faculties, faculties or universities to recommend different offered courses to students.

2.3 User Based technique analyses

Picking the correct course in youth is an unbelievably fundamental call as his future relies upon this one choice. Understudy without anyone else's input isn't full-grown enough to require the correct bring in his youth. Picking incorrectly courses implies that pair between understudy power, capacity, and private intrigue. School or people have neither the ideal data nor skill. Since there's the same dependable stockpile ordinarily offered that may manage the researcher towards the premier proper bearing, the recommender framework has been advanced to supply him guiding in picking the correct course. Recommender framework might be a pc programmed prepared with the help of experts any place the important part of the foundation of the researchers and their aptitudes help finding a course for his future investigation. This paper proposes potential forecasts for understudy's course decision bolstered their imprints and choice of employment intrigue. Pack system [8] is utilized to search out structures and connections at interims the data. This paper additionally uncovers the investigation and thusly given the inconstancy in understudy learning it's changing into logically important to tailor courses similarly as course successions to understudy needs. blessing a logical philosophy for giving modified course succession suggestions to understudies. Initial, a forward-search in reverse enlistment decide [14] is built up that may ideally pick course successions to diminish the time required for an understudy to graduate. The standard records for prerequisite necessities (ordinarily blessing in more significant level training) and course availability. Second, exploitation the tools of multi-armed bandits, an algorithmic rule is developed that may optimally suggest a course sequence that each reduces the time to graduate whereas additionally increasing the general standard of the coed.

The algorithmic rule dynamically learns however students with totally different discourse backgrounds perform for given course sequences so recommends associate optimum course sequence for brand new students. Exploitation real-world student knowledge from the UCLA Mechanical and part engineering department, we tend to illustrate however the projected algorithms crush different strategies that don't embrace student discourse data once creating course sequence recommendations.

III. METHODOLOGY

The Recommendation framework we have anticipated takes contribution from the researchers identifying with their inclinations. The framework moreover takes the information with respect to the past grades, stipulations and examination criteria of courses from the data set gave to the framework. The framework attempts to anticipate the evaluations of researchers inside the courses offered, abuse fluctuated procedures like client-based helpful filtering; thing based collective filtering, piece based edge instrument, and Content-based methodology. At that point, it looks at the outcomes and utilizations the system giving the littlest sum blunder. The yield of the framework might be a hierarchic rundown of courses in accordance with the info given by the client along the edge of the foreseen evaluations by above determined method.

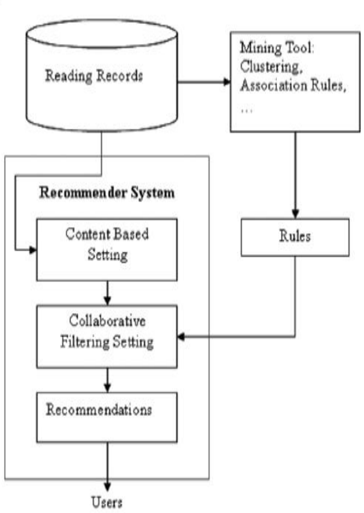


Figure 1. Overall Flow of Methodology

3.1 Content based filtering method

A content-based recommendation system recommends a course to a user based upon a description of the course and a profile of the user's interests. A substance based proposal calculation attempts to prescribe things that are like a specific client's past inclinations. The key innovation in include displaying for things and client profile building. The procedure of suggestion is to coordinate the thing highlights with the client's inclinations, discover the things generally like the client's inclinations and produce a proposal list.

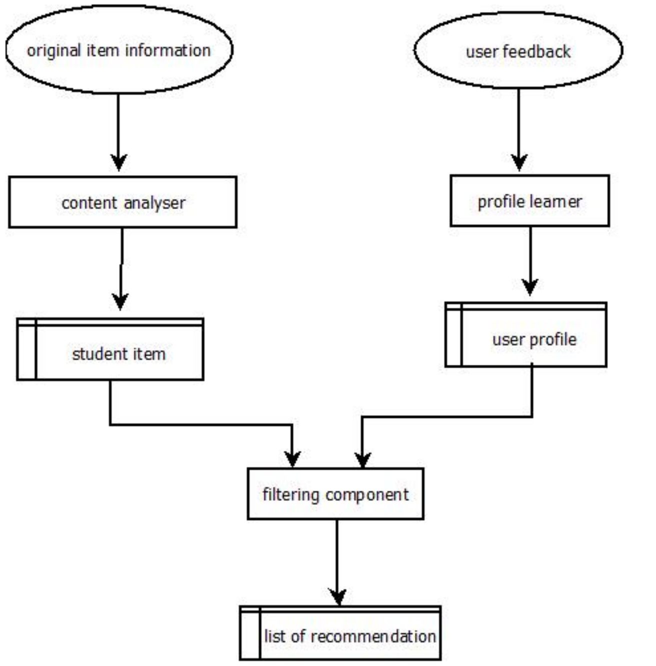


Figure 2. Content based filtering

3.2 User based collaborative filtering method

Client Based Collaborative Filtering calculation is a great customized suggestion calculation, it's generally utilized in numerous business recommender frameworks. This is a calculation dependent on the accompanying three 2010 Third International Conference on Knowledge Discover suppositions thought: People have comparable inclinations and interests. Their inclinations and interests are steady. We can foresee their decision as per their past inclinations. The accompanying equation is utilized to foresee the client based CF for the objective understudies

$$sim_{u,v} = \frac{\sum_{i\in I}(r_{u,i}-\bar{r}_u)(r_{v,i}-\bar{r}_v)}{\sqrt{\sum_{i\in I}(r_{u,i}-\bar{r}_u)^2}\sqrt{\sum_{i\in I}(r_{v,i}-\bar{r}_v)^2}} \quad (1)$$

Here $r_{u,i}$ denotes the rating of user u for course i, and $\bar{r}_u$ is the average rating given by user u calculated over all courses rated by u. Similarly, r_v denotes the rating of user v for course i, and $\bar{r}_v$ is the average rating given by user v calculated over all courses rated by v. For this step there are many algorithms can be used but Pearson -r correlation coefficient performs the best. Using this, sim (u, v) is calculated i.e. similar value of user u with v.

3.3 Item based collaborative filtering

The fundamental difference between thing based CF and client based CF is that thing put together CF creates expectations based with respect to a model of thing likeness as opposed to client closeness. In thing based cooperative filtering, first, likenesses between the different courses are registered. At that point from the arrangement of courses recently taken score by the objective client, k courses most like the objective course are chosen. For figuring the forecast for the objective course, weighted normal is taken of the objective understudy's scores on the k comparative courses prior chose.

$$sim_{i,j} = \frac{\sum_{u\in U}(r_{u,i} - \bar{r}_u)(r_{u,j} - \bar{r}_u)}{\sqrt{\sum_{u\in U}(r_{u,i} - \bar{r}_u)^2}\sqrt{\sum_{u\in U}(r_{v,i} - \bar{r}_v)^2}} \quad (2)$$

Here $r_{u,i}$ denotes the scores of student u for course i, and $\bar{r}_u$ is the average scores scored by user u calculated over all courses previously taken by u. Similarly, $r_{u,j}$ denotes the rating of student u for course j.

3.4 Prediction Methodology

Score Prediction strategy: The technique predicts understudy score in online course from imprints acquired in requirements schedule and normal execution of understudy in all courses taken. The informational index comprises of four properties for example Understudy, Domain, Subject (which establish the courses) and weight (appointed to courses of areas). The acquired stamps in essential prospectus will be utilized to discover the score of understudies in individual course.

3.4.1 Student Course and its score

To find the score of future courses and the following formula will be used

$$\text{Course Score} = \sum_{i=0}^{n} S_i * W \quad (3)$$

Where i indicate course number s_i indicate obtained score grade in the i^{th} course W indicate weight of course in curriculum. After multiplying subject percentage score with respective weight value of each subject which constitute the overall curriculum, sum all of them. For any of the procedure to be done, there must be a successive progression of the framework which unmistakably clarifies the significant working and the progression of the framework. The accompanying chart speaks to the general framework design of the course proposal framework which our paper pursues throughtout for getting the determained yield.

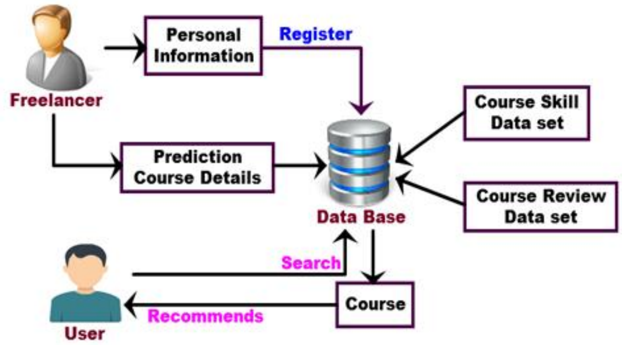


Figure 3. System Architecture

The system takes the personal information from the new user and creates a profile for the user and the user provides the criteria for searching. System obtains user information and predicts the similarity between the current user preference and the existing user preferences. System generates the list of courses recommended for the current user. Course skill data set and course review data set are used to generate the similarity course for prediction.

IV. RESULTS

4.1 Evaluation Metric

The evaluation involves completion of the matrix which has user's choices in particular. The above methods have been applied on the matrix to find the most efficient method through the process of cross validation.

Table1.Evaluvation Metric

TN/True Negative	Courses recommended were not Relevant
TP/True Positive	Courses recommended were Relevant
FN/False Negative	Courses which are not Recommended are Relevant
FP/False Positive	Courses which are not Recommended are not Relevant

4.2 Precision

Out of all the recommended courses, how many did the user actually like? It is given by:

$$\text{Precision}= \frac{TP}{FP+TP} \quad (4)$$

Here tp represents the number of courses recommended to a user that he/she likes and fp+tp represents the total courses recommended to a user. If 5 courses were recommended to the user out of which he liked 4, then precision will be 0.8. Larger the precision, better the recommendations.

4.3 Recall

What proportion of courses that a user likes was actually recommended? It is given by:

$$\text{Recall} = \frac{TP}{FN+TP} \quad (5)$$

Here tp represents the number of courses recommended to a user that he/she likes and fn+tp represents the total courses that a user likes. Larger the recall, better are the recommendation. The following table shows the deviation between item based and user based collaborative filtering methods with varying number of user's

Table 2. Deviation Results

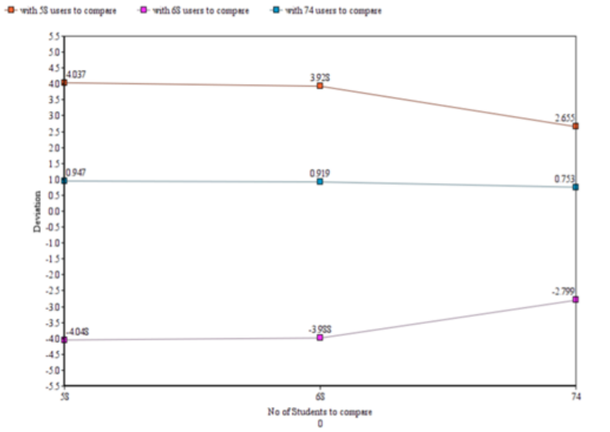
Methods	Positive Deviation	Negative Deviation	Average Deviation
Item-based collaborative filtering	4.1	-2.9	1.137
User based collaborative filtering			
With min users	-4.04	4.03	0.9
With avg users	-3.98	3.92	0.9
With max users	-2.79	2.65	0.7

For content-based approach prediction is based only on student's own preferences and course credit structure, there is no need of hiding records of other student it will not affect the predicted rank list of courses. If 5 courses were recommended to the user out of which he liked 4, then precision will be 0.8. If a user likes 5 courses and the recommendation engine decided to show 3 of them, then the recall will be 0.6.

4.4 Summary of Results

Content Based Algorithm - Since here comparability and preferences of one specific understudy in different courses is determined, at that point in positioned request the rundown of courses is prescribed to the understudy however it brings about less precision when contrasted with different strategies. Item Based Algorithm - Since Data set for thing is less.

Utilizing this course expectation of scores for understudy veered off greatest to 4.1 positive deviations and - 2.9 negative deviations for couple of understudies, normal deviation is 1.137. Subsequently, it demonstrates that Item Based Collaborative methodology performing poor in given dataset, as a result of less no of k neighbors (comparable courses). User Based Algorithm - Since here we are computing comparability between understudies which are huge in numbers than courses, it is giving better precision. Chart given underneath incorporates most extreme positive deviation, greatest negative deviation and normal deviation with diferent no of k-neighbors.



It demonstrates that when information or number of k-neighbors (comparative clients) lessens deviation from real worth increments. Those Courses which don't have essential have a most extreme deviation of +3 to - 3 of genuine evaluation acquired; this is on the grounds that evaluation anticipated is normal score of understudies till now which may change much for the understudy. For different courses, the most extreme deviation of - 2 to +2 of real evaluation acquired and normal deviation is - 1 to +1.

V. CONCLUSION

In this way, suggestion framework can be utilized to pick the best courses for understudies. Separating strategies are utilized for suggestion and Pearson relationship coefficient is utilized for comparability computation. From the outcomes it is seen that User Based Collaborative Filtering performs superior to anything Item Based Collaborative Filtering and Content based Filtering for course determination for accessible measure of information. The efficiency of the framework can be additionally improved by expanding the informational index to show signs of improvement results and can likewise by utilizing profound neural systems to give better outcomes. Later on we might want to create multi relapse models for the same number of as wide scope of courses accessible online with a superior precision by including number of informational collections. New advances like profound neural system framework and different ideas can be incorporated to improve this task further.

REFERENCES

1. Deepthi Upendrana, Shiffon Chatterjeea, Sindhumol Sa, Kamal Bijlania "Application of Predictive Analytics in Intelligent Course Recommendation",6th International Conference On Advances In Computing & Communications, ICACC 2016, 6-8 September 2016, Cochin, India.
2. Jun Tao, Gui Wu "An Application of Undergraduate Academic Growth Path on the Credit System Based on Data Mining" ,12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) ,2015.
3. Kiratijuta Bhumichitr, Songsak Channarukul, Nattachai Saejiem, Rachsuda Jiamthapthaksin, Kwankamol Nongpong "Recommender Systems for University Elective Course Recommendation", Computer Science Department, Assumption University, Bangkok, Thailand, 2017.
4. Ding Yanhui, Wang Dequan, Zhang Yongxin, Li Lin ," A Group Recommender System for Online Course Study" , 7th International Conference on Information Technology in Medicine and Education, 2015.
5. Bushra Alhijawi, Nadim Obeid, Arafat Awajan, Sara Tedmori, "Improving Collaborative Filtering Recommender Systems Using Semantic Information", 9th International Conference on Information and Communication Systems (ICICS), 2018.
6. Eric L. Lee , Tsung-Ting Kuo, Shou-De Lin, "A Collaborative Filtering-Based Two Stage Model with Item Dependency for Course Recommendation ", International Conference on Data Science and Advanced Analytics, 2017.
7. Sunita B Aher1, Lobo L.M.R.J.2, "Course Reccomender System In E-Learning", Vol. 3, No. 1, January-June 2012, pp. 159-164.
8. 8. Grewal DS1 and Kaur K2, "Developing an Intelligent Recommendation System for Course Selection by Students for Graduate Courses", Vol. 7:2, Bus Eco J 2016.
9. Zhe Yang, Bing Wu, Kan Zheng, "A Survey of Collaborative Filtering-Based Recommender Systems for Mobile Internet Applications", Vol. 4,May 2016.
10. Puneet Garg, Rinkle Rani and Sumit Miglani , "Mining Professional's Data from LinkedIn", Fifth International Conference on Advances in Computing and Communications, 2015.
11. Surabhi Dwivedi and Dr Kumari Roshni VS, "Recommender System for Big Data in Education ", IEEE, 2017 .
12. Guibing Guo, Jie Zhang, and Neil Yorke-Smith, "A Novel Recommendation Model Regularized with User Trust and Item Ratings",IEEE, 2015.
13. Chao Wu, Yaoxue Zhang, Jia Jia and Wenwu Zhu, "Mobile Contextual Recommender System for Online Social Media", Vol.14, August 2015
14. Jie Xu, Tianwei Xing, and Mihaela van der Schaar, "Personalized Course Sequence Recommendations", IEEE, 2016.
15. Xiwang Yang, Chao Liang, Miao Zhao and Hongwei Wang, Hao Ding, Yong Liu, "Collaborative Filtering-Based Recommendation of Online Social Voting", IEEE, 2017.

AUTHOR PROFILE

Mariappan A.K received the Bachelor of Engineering degree in Computer Science and Engineering from Madurai Kamaraj University , Madurai, India in 1984,the M.E degree in Computer Science and Engineering from Jadavpur University, Calcutta, India and Ph.D degree in Computer Science and Engineering from Jawaharlal Nehru Technological University , Hyderabad , India. He is currently working as a Professor in Easwari Engineering College, Chennai, India. He is senior member in Computer Society of India and member in ISTE. His research interest includes data mining, Search Engines, and NLP. He has published more than 30 research papers in peer reviewed Journals and Conferences.

Ramana Surriyan completed his Bachelor degree in Information Technology from Anna University , Chennai, India, in 2018. His area of research includes webmining, recomender system and machine learning.

E.S.Shobana completed her Bachelor degree in Information Technology from Anna University , Chennai, India, in 2018. Her area of research includes software Engineering, recomender system and machine learning.