A Study on Learning Analytics with Recommended System

N.UttejKumar[a] , S.Ranjeeth[b], A.Sravya[c], Dr.RatnaRajuM[d]

[a,b]Assistant Professor, Computer Science and Engineering, Vignan's Foundation For Science, Technology and Research, Andhra Pradesh, India
[c]Assistant Professor, Computer Science and Engineering, Vignan Lara Institute of Technology and Science, Andhra Pradesh, India.
[d]Associate Professor, Computer Science and Engineering, St.Ann's College of Engineering and Technology.

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 23 May 2021

Abstract: Education data mining is becoming popularfrom the education data we can derives the insights using data mining techniques. Insights which are used for students, faculty, and education organizations and one of the famous applications in data mining is recommender system. In the field of learning analytics recommender system have become common in recent years. In education sector student performance can increase by using different tools, which are used in recommender system application implementation. The style and expectations of the user should be correctly constructed to include the most appropriate suggestions. Researchers have proposed various types of recommender systems in learning analytics (RSL).The authors of this paper reviewed several current state of recommendation system models and presented their selection criteria in learning analytics. Authors investigated various significant preference factors and classified them based on their similarity in the RS (Recommender System).This article presents the future directions of RSL (Recommender Method in Learning Analytics) models and compiles a detailed.

Keywords: Recommender system, learning analytics, collaborative filtering, education data mining

I. INTRODUCTION

Data Mining (DM) is a technique for extracting information from a company's various databases and repurposing it for purposes other than the databases that were originally intended for data mining implementation, depending on the type of data and the organization. Data mining is the process of extracting information from data. which specifies juice or mining experience from massive data so to understand from scratch in this article we explained about recommender system and its working, its techniques (machine learning methods) [19, 20] and in next sectionexplained about learning analytics and its techniques finally recommender system in learning analytics.

II. RECOMMENDER SYSTEM

Recommender Systems are pieces of software that suggest things you might be interested in if you haven't tried them yet. If you bought a book on machine learning, for example, it suggested a list of books on artificial intelligence (AI), pattern recognition (PR), data mining (DM), and so on. Software tools Rs (Recommender system) and techniques provide recommendations such as items that are useful to a consumer.

Working Model of Recommendation system
On the internet, there are a lot of choices; the channel is necessary, important, and effective in conveying accurate information to reduce the issue of information overload, which has caused future confusion for many online customers. Recommender systems solve this problem by sifting through a large amount of artificially created data to provide clients with customized data and services.
The following are the steps of working of recommendation system

Step 1: Collection of Data
In order to construct a proposition motor, it is necessary to collect data. It depends on the type of data whether it is understood or expressed. Client feedback on items, such as ratings and comments, is treated as express data. Solicitation and return history, truck events, page views, navigate, and discover logs are all treated as understood data. This information will be collected from each client when they visit the website.
Step 2: Storing
If the data set is large, algorithms can make good recommendations. The database tool that is used to create the recommendations is chosen based on the type of data. Database tools include NoSQL, standard SQL, and some sort of repository item..

Step 3: Analyzing
What are the chances of us noticing things that have similar client commitment data? In order to attempt to do, in this way, we watch out for clean the data by using absolutely variation investigation ways. In the event

that you wish to create rise proposals to the customer as they're seeing the thing, at that point you may need an extra snappy considering kind investigation. We can separate the data in a variety of ways, including:

a) Real-time frameworks: Made information can measure by it. Preparing and dissecting floods of occasions apparatuses normally include in an ongoing framework. This framework would be expected to give in the piece suggestions.

b) Batch investigation: Handling the information occasionally is an interest for clump examination. This worry recommends sufficient data should be made to frame the investigation intently, for example, every day deals limit. This framework may function admirably to send an email on the following date.

c) Near-continuous investigation: if you collect data quickly, you can restart the investigation at regular intervals or seconds. Giving recommendations during the clone (same) browsing period is best when working in a near-real-time system.

Algorithm

For each item in the product catalog, I1
For every client C who purchased I1
For every item I2 purchased by client C
The record that a client purchased I1 and I2
For every item I2
Calculate the similarity between I1 and I2

Step 4: Filtering

The final step is information filtering, which is useful for bringing predictable data needed to make proposals to customers. We need to pick a technique that is better suited to the proposition engine (motor). The calculations will come next.

a) Content-based: A renowned, recommended item has similar highlights as what a purchaser likes or perspectives.

b)Cluster: Suggested items work well together; it is unconcerned about what different clients have done.

c) Collaborative: Different clients who like the same thing as another client's perspectives or preferences will also refer to a suggested item as a collaborative item.

In this article we are mainly focusing on collaborative filtering in learning analytics.

AAlgorithms of Recommenders system: In the proposal framework three principle channel methods, these strategies contain calculations. Methods like substance, community oriented, cross breed based strategies. Synergistic strategy separated into two classes: model-based and memory-based

a) Model-based techniques: A depiction model is created from a data set, and a functioning client receives forecasts from this model. The following are some examples of model structure measurements obtained through variation seeking after calculations.

i) Artificial Neural Network (ANN): An ANN is a system that consists of various related neurons that are organized in layers in a productive way. Loads are associated with neuronal relationships, depending on the proportion of effect one neuron has on another. We can use neural networks in a few unusual problem situations where we have a few preferences. Unnatural neural frameworks, for example, are extremely effective in managing the aggravation and jumbled up because they contain numerous neurons and also distribute a load to each affiliation. Collections of information.Assume nonlinear limits and getting difficult associations in data sets can also be done by fake neural organizations, which are capable of working even if some parts of the system fail. When geography is chosen, finding the ideal framework geography for a given issue is a significant detriment to ANN, and it acts as a lower set out toward the gathering botch.

ii) The Decision Tree Technique (DTT) is a construction that combines the root, branch, and leaf centers into a single structure. Every inside center depicts a quality test, each branch represents the result of a test, and each leaf center represents a class mark. The DT is created by separating a game plan for planning cases for which the class names are known, and it is based on the tree chart way of thinking. They are then linked to portraying inconspicuous events effectively. They can make astonishingly accurate forecasts [13] when based on high-quality data.Choice trees are more interpretable than SVM (support vector machine) and ANN (fake neural organization) classifiers because they join fundamental requests with respect to data in a justifiable manner. The decision tree is very flexible when dealing with products that have a mix of real-valued and categorical options, as well as when dealing with objects that have a few specific missing options.

ii) The path toward stirring up frameworks of interconnected things to explore examples and patterns is through link analysis [14]. It's used to boost the success of a website's appearance by demonstrating its incredible capabilities. HITS calculations and page rank are also present in connection examination, with a site page serving as an independent center point in the web chart [16], which is handled by standard interface investigation calculations.

iii) Matrix completion methods: We can predict unknown values within user item matrices using the matrix completion technique. One of the critical techniques used in community separating suggestion frameworks [15] is relationship-based K-nearest neighbor (KNN). They rely on customer rating data that has been collected over time. Many things that are addressed within the framework [13] are not given a rating by the client because the appraising matrix is frequently colossal and sparse. This problem consistently causes the system to be disappointed in its ability to provide customers with strong and precise recommendations.Various low-position models have been used in lattice consummation for training varieties, particularly in communitarian sifting toward application.

iv) Bayesian Classifiers: The probabilistic structure can help with order issues. It is based on what the Bayes hypothesis and restrictive likelihood mean. Bayesian classifiers treat class marks and each trait as arbitrary factors [12]. Innocent Bayes classifiers have two main advantages: they are resistant to disconnected upheaval focuses and unimportant properties, and they handle missing characteristics by ignoring the case during probability assessment calculations.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

P =Probability

H =Posterior Probability of a Hypothesis

X = training data

Fig. 1. Bayesian classifier probability formula

b) Memory-Based Techniques: The entire customer thing information base can be used to forecast age; we can find the neighbor using measurable procedures, which is also known as the closest neighbor. Memory-based CF can be achieved by relying on a client's abilities (collective sifting). Comparative customer based communitarian separating methods figure resemblance among customers by looking at customer assessments on something, and it then determines the expected rating for a thing by the unique customer as a weighted average of the assessments of the thing by customers like the powerful customer, where loads are the similarities of these customers with the goal thing. Simplicity between things, not customers, is used to figure expectations in thing-based separating procedures. It compiles a model of thing resemblances by retrieving everything from the customer thing organization (framework) that has been evaluated by a working customer in order to determine the degree of similarity between them.

Thing/client numerous sorts of resemblance measurements are used. Pearson relationship coefficient and cosine are the 2 celebrated similarity measurements.

i) Pearson correlation coefficient: It will take a range of qualities (r) from +1 to -1 to calculate the Pearson correlation coefficient. With 0, there is no relationship between the two factors. If the no is greater than 0, it is referred to as positive affiliation; if it is less than 0, it is referred to as negative affiliation. When one variable's worth is increased in a certain affiliation, the value of another variable increases; however, when one variable's worth is reduced in a negative affiliation, the value of another variable decreases.

$$r = \frac{N(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[N(\sum X^2) - (\sum X)^2][N(\sum Y^2) - (\sum Y)^2]}}$$

N = number of pairs of scores

$\sum XY$ = sum of the products of paired scores

$\sum x$ = sum of X scores

$\sum Y$ = sum of Y scores

$\sum x^2$ = sum of squared X scores

$\sum Y^2$ = sum of squared Y scores

Fig.2. Pearson correlation coefficient formula

Compare this to a Pearson-based metric. In comparison to a factual methodology, cosine similitude is a vector-space show that relies on straight polynomial math based math. The point of comparability between two n-dimensional vectors is determined by it. The cosine-based measure is used in text mining (TM) to look at two content records and fields of data recovery. In this case, vectors of terms address the reports. The closeness measure of equivalence between two things is also referred to as likeness metric, and the methods used to calculate the scores that express how similar customers or things are to one another are also referred to as likeness metric. reiterating use scores as an establishment.Distance or connection measurements can also be insinuated as comparability measures depending on the state of use.
There are numerous areas utilizing proposal frameworks, for instance, the travel industry and travel, online business, film and music, wellbeing and Sensex, and so on This paper surveys about proposal framework in learning investigation.

III. LEARNING ANALYTICS (LA)

Learning analytics (LA) [11] is an emerging field in the education sector. It is used to evaluate important information on understudies and teachers at a small extension level that targets singular understudies and the courses are taken to comprehend understudy execution and advance understudy achievement [9]. It is primarily focused around learning measure [8], and it is used to evaluate important information on understudies and teachers at a small extension level that targets singular understudies and the courses are taken to comprehend understudy execution and advance understudy achievement. With advanced LA methodologies and examination instruments, understudy execution and analysis results can be improved through refocused help and intervention, advancing learning and instruction [10]. In order to improve the demonstration of learning and teaching for singular students, the research and further LA incorporate the development, use, and blend of new methodology and tools.

TABLE 1. LEARNING ANALYTICS AND ITS RELATED FIELDS

S.no	Field	Stake Holders	Objectives	Methods	Data
1	AA (Academic Analytics)	Institutions of Education	Admissions management, forecasting, marketing, and decision-making	Techniques of statistics	Data of education sector
2	EDM (Education Data Mining)	Faculty, Understudies	Improving the learning process by converting information into meaningful data	Techniques of DM (data mining)	Data of education sector
3	LA (Learning Analytics)	Faculty, Learners, Institutions of education	Suggestion, prediction, enrollment, syndication, domestication, personalization	Methods of quantitative and DM techniques	Data of education sector

A. Learning Analytics Techniques
Regression, classification, clustering, artificial neural network are the techniques used in learning analytics field. LA is a part of education data mining in previously, but now treating as different sectors like AA (academic analytics), LA (Learning analytics) and education data mining. To get clear knowledge about learning analytics [11]

IV. RECOMMENDER SYSTEM IN LEARNING ANALYTICS

TABLE-2 LITERATURE SURVEY OF RECOMMENDER SYSTEM IN LEARNING ANALYTICS

S.No	Author	Proposed Work and techniques used
1	Faisal M. Almutairi [1] 2017	In this article three approaches are developed under collaborative filtering framework, two approaches are built by using coupled matrix factorization with latent matrix factor third approach was build using tensor factorization to model grades. These three methods are used to incorporate additional information in the context of collaborative filtering (CF), slow learners' problem also handle when predicting for next semester, and author evaluated these three proposed models on grade data which is collected from university of Minnesota.
2	Nguyen Thai-Nghe [2] 2010	In this article, recommender system methods for learning analytics in education data mining are used to propose a novel approach for predicting student performance, common regression techniques such as logistic/linear regression are used to validate the proposed approach on different dataset (two datasets are taken from Knowledge discovery data mining (KDD challenge-2010 data sets))
3	Sanjog Ray [3] 2011	In this article, author proposed a CRS (Course Recommender System) to student which is used for improving their final grade in academic level. Based on the collaborative filtering technique CRS is developed. It is evaluated on real time data set.
4	S.JothiLakshmi [4] 2018	In this article, author proposed a recommender system called Integrated Recommender Educational Data mining (IRED) for higher educational institution target marketing, it is a unique design which provides solutions for target marketing in higher education institutions. Techniques used for building proposed work are collaborative filtering as filtering agent and C4.5 as pattern discovery model.
5	Phung Do [5] 2017	In this article, e-learning material recommendation system is proposed using collaborative filtering and knowledge based reasoning techniques, proposed method evaluated on three datasets which are taken from cognitive tour and compared model performance with other three different techniques called MF (Matrix factorization), RBR (Rules based reasoning) and CBR (case-based reasoning).
6	Alexandre L [6] 2018	In this article, student gets the recommendation based on the assessment taken by him. The assessments used in this article are calculus, hand-on, remote lab, simulations. After completion of assessment student will get analysis and suggestion.
7	Alexandre L. Gonçalves [7] 2018	In this article, on remote laboratories activities suggestions provided to students in order to scaffold their performance using learning analytics and recommender system techniques. It is also providing the performance analysis and possible errors done by student during lab experiment
8	David Adrian Sanders [18]	The system integrates with new client-based systems that filter Web pages and provide convenient, structured, focused, and controlled Internet access. The first system, known as iLessons, was built into Microsoft Internet Explorer 6 and gave teachers the ability to create lesson Web pages, define Internet zones that could be accessed during a lesson, and enforce these settings across a group of computers. A second system allowed students to use the Internet to research and collaborate. The system filtered Web pages based on their content relevance and aided students by inferring their learning style (active or reflective) and recommending pages discovered by other students based on their on page relevancy, student learning style,and state of mind measured by activity.

V. CONCLUSION

This paper focused on recommender system in learning analytics and its importance in the education field, giving recommendations to students, learners, teachers, and education organizations. It also discussed sample algorithm working flow of RS, types of recommendation algorithms, and this paper focused on recommender system in learning analytics and its importance in the education field, giving recommendations to students, learners, teachers, and education organizations. Based on demographic profiles, online courses taken, data from MOOCs, and other factors.

REFERENCES

1. Faisal M. Almutairi, Nicholas D. Sidiropoulos and George Karypis Context-Aware Recommendation-Based Learning Analytics Using Tensor and Coupled Matrix Factorization, IEEE journal of selected topics in signal processing, vol. 11, no. 5, august 2017
2. Nguyen Thai-Nghe∗ Lucas Drumond∗, Artus Krohn-Grimberghe∗, Lars Schmidt-Thieme∗, Recommender System for Predicting Student Performance, Procedia Computer Science 1 (2010) 2811–2819.

3. Sanjog Ray, Anuj Sharma, A Collaborative Filtering Based Approach for Recommending Elective Courses, International Conference on Information Intelligence, Systems, Technology and Management ICISTM 2011: Information Intelligence, Systems, Technology and Management pp 330-339
4. S. JothiLakshmi, Dr. M.Thangaraj, Design and Development of Recommender System for Target Marketing of Higher Education Institution Using EDM, International Journal of Applied Engineering Research ISSN 0973 -4562 Volume 13, Number 19 (2018) pp. 14431-14437.
5. Phung Do, Kha Nguyen, Thanh Nguyen Vu, Tran Nam Dung, and Tuan Dinh Le, Integrating Knowledge-Based Reasoning Algorithms and Collaborative Filtering into E-Learning Material Recommendation System, Springer International Publishing AG 2017 T.K. Dang et al. (Eds.): FDSE 2017, LNCS 10646, pp. 419–432, 2017. https://doi.org/10.1007/978-3-319-70004-5_30
6. A. L. Gonçalves, L. M. Carlos, J. B. da Silva and G. R. Alves, "Personalized Student Assessment based on Learning Analytics and Recommender Systems," 2018 3rd International Conference of the Portuguese Society for Engineering Education (CISPEE), Aveiro, 2018, pp. 1-7, doi: 10.1109/CISPEE.2018.8593493.
7. Alexandre L. Gonçalves, Gustavo R. Alves, Lucas M. Carlos, Juarez B. da Silva, João B. da M. Alves, Learning Analytics and Recommender Systems toward Remote Experimentation
8. G. Siemens, P Long. (2011) "Penetrating the fog: analytics in learning and education." Educause Rev. 46(5): 30–40.
9. Learning Analytics: Definitions, Processes, and Potential. http://learninganalytics.net/ Learning Analytics Definitions Processes Potential.
10. A van Barneveld K.E Arnold J.P Campbell (2012) "Analytics in higher education: establishing a common language". EDUCAUSE Learn. Initiative. 1:1–11.

11. S Ranjeeth, T.P. Latchoumi, P Victer Paul. "A Survey on Predictive Models of Learning Analytics", Procedia Computer Science, Volume 167, Pages 37-46, 2020 https://doi.org/10.1016/j.procs.2020.03.180.
12. Friedman, N., Geiger, D., Goldszmidt, M. Bayesian network classifiers. Mach Learn 1997; (Vol.29 (2-3), pp.131–63).
13. Caruana R, Niculescu-Mizil A. (2003) an empirical comparison of supervised learning algorithms. In: Cohen W. Moore AW, editors. Machine Learning, Proceedings of the twenty-third international conference, ACM, New York; (pp. 161–8)
14. Berry, MJA. Linoff, G. Data mining techniques: for marketing, sales, and customer support. New York: Wiley Computer Publishing; 1997.
15. Koren, Y., Bell, R., Kolinsky, C. Matrix factorization techniques for recommender systems. IEEE Comput 2009; (Vol.8, pp.30–7)
16. Deng, C., Xiaofe, H., Ji-Rong, W, Wei-Ying M. Block-level link analysis. In: Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval; 2004. p. 440–7
17. Ranjith, S., Victer Paul, P., "A Survey on Recent Recommendation Systems for the Tourism Industry." Accelerating Knowledge Sharing, Creativity, and Innovation Through Business Tourism, edited by Muhammad Waseem Bari, et al., IGI Global, 2020, pp. 205-237. http://doi:10.4018/978-1-7998-3142-6.ch012
18. D. A. Sanders and J. Bergasa-Suso, "Inferring Learning Style From the Way Students Interact With a Computer User Interface and the WWW," in IEEE Transactions on Education, vol. 53, no. 4, pp. 613-620, Nov. 2010, doi: 10.1109/TE.2009.2038611.
19. S Ranjeeth, T.P.Latchoumi and P Victer Paul, "Optimal Stochastic Gradient Descent with Multilayer Perceptron Based Student's Academic Performance Prediction Model", Recent Advances in Computer Science and Communications (2021) 14: 1730. https://doi.org/10.2174/2666255813666191116150319
20. Ranjeeth, S., Latchoumi, T.P. (2020). Predicting kids malnutrition using multilayer perceptron with stochastic gradient descent. Revue d'Intelligence Artificielle, Vol. 34, No. 5, pp. 631-636. https://doi.org/10.18280/ria.340514