PATNA WOMEN'S COLLEGE AUTONOMOUS 1940

Explore – Journal of Research
Peer Reviewed Journal
ISSN 2278–0297 (Print)
ISSN 2278–6414 (Online)
Vol. XIV No. 2, 2022

https://patnawomenscollege.in/explore-journal-of-research/

A Literature Review on School Student Interest using Datamining Algorithms

• **Khushbu Agrawal** • **Bhavana Narain**

Received : April 2022
Accepted : May 2022
Corresponding Author : Khushbu Agrawal

Abstract: The educational performance plays a vital role in the classification of the student in school education. The student performance effect various factor like learning process, personal and social. This paper demonstrates the impact of student positive or negative performance of student success. With the help of data mining algorithm we predict the student dropout rate, which is helpful for academic progress. For this we have reviewed various papers in three section algorithms, tools and dataset. These literatures help us to frame our technical work which will be continued in our future work.

Keywords: *student performance, prediction, classification algorithm, data mining, academic performance*

Khushbu Agrawal
Research Scholar
MATS School of Information Technology
MATS University, Raipur (c.g.), India
Email:khushi9r@gmail.com

Bhavana Narain
Professor
MATS School of Information Technology
MATS University, Raipur, Chhattisgarh, India
Email:narainbhawna@gmail.com

Introduction:

Educational data mining is a scientific research area, it use the multiple algorithm to improve academic result and procedure for further decision making. Predicting student performance in academic data is an important issue in e-learning environments. Student performance is based on various factors such as personal, social, psychological and other issues. Data mining techniques is a promising tool to attain these objectives, data mining techniques are use to bring hidden information, patterns and relationship among the large dataset, which help us in categorization of data into knowledgeable facts. To identify the prediction of risk students with a large no. of student data set, it is very difficult and time consuming to using traditional data mining research methods such as questionnaires. Using traditional method in data mining has some limitations like it cannot properly handle the missing values, requires detailed information about the data, and cannot deal with uncertainty or vagueness in any information domain. Various tools and techniques required for achieving the best result from data mining like data cleansing, AI, association rule mining, clustering, regression, machine learning and classification. So the classification is one of the most useful predictive data mining techniques to solve this problem, and customized traditional method by applying various classification techniques. Student performance

predictions with high advantageous is beneficial for identifying the students with low academic achievements.

Methodology:

In this paper we have done literature review based on dataset, tools and algorithm. This review will help us to find out best algorithm for large dataset, pre-processing of dataset and tool which are applied for preprocessing of dataset and tool which are applied for preprocessing of dataset and finding the results in data.

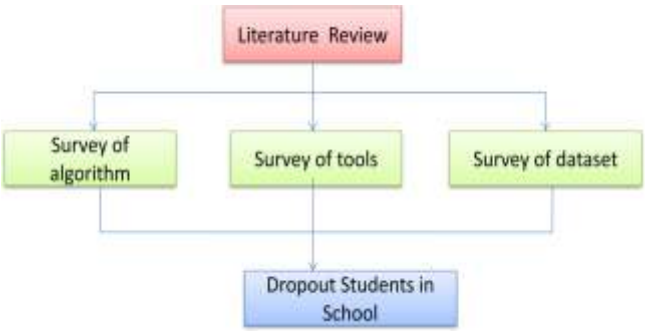

Fig.1. Hierarchy of Literature review

Literature Review:

Several empirical studies have been conducted the data mining in academic data. In several research papers, most of the research work is done in small scale and structured dataset that are discussed below:-

Literature review based on Algorithm-

Ali et al., (2019) Investigate k-nearest neighbor classification (k-NN) performance on heterogeneous data sets which include the combination of both numerical and binary data. Earlier traditional k-NN works on numerical data only.

Algarnih (2016) has concluded that Educational data mining (EDM) use to improve teaching and learning process for data mining. Data mining is a process to finding knowledgeable information.

Amra and Maghari (2017) study focuses on KNN and Naïve Bayesian Algorithms Model for prediction of student performance. KNN and Naïve Bayesian Algorithm were applied on the educational data set of secondary school.

Bhavsar and Panchal (2012) reviewed about the Support Vector Machine (SVM), SVM helps in solving statistical learning problems such as spam filtering, text classification, handwriting analysis, face and object recognition, and countless others.

Fairos et al., (2019) data mining algorithm was applied to predict student performance either excellent or non-excellent. The student performance was conducted on selected university in Malaysia.

Romero et al., (2021) had study about early prediction of student academic performance using j48 algorithm. We have shown that how early the prediction can be done varies based on the type of educational system.

Namdeo et al., (2021) had study about student prediction and survey on prediction system for education area. They use Naïve bayes and SVM fine tuning done automatically. Classification takes less execution time and gives better accuracy.

Literature review based on Dataset-

Anandavally et al., (2015) have studied about clustering method in educational data mining, which play important role in data mining techniques and partitioning student dataset for the purpose of student prediction after their final examinations, student learning.

Abu saa (2016) has concluded that data mining methodology to expose multiple factor related to student performance prediction on academic data.

Bharadwaj and Pal (2011) applied a decision tree model to predicted student division on previous student database, studied the Purvanchal University, Janunpu from session 2007 to 2010.

Pal et al., (2022) had study about student behavior or prediction and throws light on different – different situation and real time problem in teaching platform. They also discussed about teacher role and all development of student.

Zhang et al., (2021) demonstrate the importance of SVM algorithm in context of big data in data mining classification algorithm and analyze various improved method of SVM.

Literature review based on Tools-

Gupta et al., (2020) has concluded that predict student performance by traditional classification techniques in data mining. They presented 7 different classification techniques namely- REPTree, PART, Random Forest, J48, 1-Bk, Naïve Bayes and BayesNet. Finally J48 algorithm of decision tree gave the best result compared to other data mining techniques

Gowri et al., (2017) has found the importance of educational data mining application, through this method they have predicted the performance of the students using Weka apparatus. The objective of the research is to find out that the student is prone to violence or not.

Hashim et al. (2003) presents a case study to student's academic records in faculty of mathematical science and statistics their study was done on the data set that has 124 graduate students and study is implemented using WEKA. Viswanathan and Vengatesh Kumar [16] have concluded that student performance prediction and define methodology for implementing student prediction. They present ensemble support vector machine (ESVM)

Conclusion:

Educational data mining play an important role in school education system, the use of rising technology need to largest dataset. With the help of literature review from three domain such as algorithm, tools and dataset we got overall type of big data as related to school information like students, faculty members etc., which is use in data mining concept to predict student performance. By this study provide improvement in public and private field to improve academic performance. It can be concluded that various research works have been done to find out drop ratio of students in school education. Regress work is still required.

References :

1. Ali N.,Neagu D. and Trundle P.(2019November 6).Evaluation of nearest neighbour classifier performanceonhetrogeniusdataset,SNa p p l i e d science1:1559.https://doi.org/10.1007/s42452-019-1356-9
2. Algarni A.(2016).Data mining in Education, (IJACSA) International Journal of Advanced Computer Science and Applications,Saudi Aribia Vol. 7, No. 6.www.ijacsa.thesai.org
3. Abu Amra I.A., Maghari A.Y.A.(2017 May) "Students Performance Prediction Using KNN and Naïve Bayesian", International Conference on Information Technology (ICIT), https://www.researchgate.net/publication/320672596
4. Bhavsar H., Panchal M.H.(2012 December 10). A Review on Support Vector Machine for Data Classification: International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 10, ISSN: 2278-1323.
5. Yaacob W.F.W., Nasir S.A.M., Yaacob W .F. W., Sobri N.M.(2019 December 3).Supervised data mining approach for predicting student performance:Indonesian Journal of Electrical Engineering and Computer Science Vol. 16, pp. 1584~1592 ISSN: 2502-4752, DOI: 10.11591/ijeecs.v16.i3.pp1584-1592.
6. Zambrano J.L.,Lara Torralbo J.A., and Romero C.(2021).Early Prediction of Student Learning Performance Through Data Mining: A Systematic Review: ISSN 0214 - 9915 CODEN PSOTEG , Vol. 33, No. 3, 456-465 doi: 10.7334 /psicothema 2021.62
7. Ingale N.V., Dr. Sivakkumar M., Dr. Namdeo V.(2021 june 4).Survey on Prediction System for Student Academic Performance using Educational Data Mining:Turkish Journal of Computer and Mathematics Education Vol.12 No.13 (2021), 363-369, published online.
8. C.Anuradha, T.Velmurugan, R. Anandavally. (2015) Clustering algorithm in educational data mining: A review", International Journal of Power Control and Computation (IJPCSC) ISSN: 0976-268X, Vol 7. No.1 – 2015 Pp.47-52.www.ijcns.com
9. Saa A.A.(2016).Educational data mining and student performance prediction: (IJACSA) International Journal of Advanced Computer

Science and Applications, Vol. 7, No. 5.www.ijacsa.thesai.org

10. Bharadwaj B.K., and Pal S.(2011).Mining Educational Data to Analyze Student's Performance: International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6,.www.ijacsa.thesai.org
11. Pal S., Sarkar Dr. P. ,Bhattacharya S. (2022 January 1). Scope for Applications Management in India of Soft Computing in effective Classroom.International Journal of Instructional Technology and Educational Studies (IJITES),ISSN (Print): 2682-3918 - ISSN (online): 2682-3926.https://ijites.journals.ekb.eg/article_204403_80de6500f5bd1a1cd02906806cf3dbed.pdf
12. Gaye B., Zhang D. and Wulamu A.(2021)Improvement of Support Vector Machine Algorithm in Big Data Background:Hindawi, Mathematical Problems in Engineering, Article ID 5594899. https://doi.org/10.1155/2021/5594899
13. Kaur B., Gupta A., Singla R.K.(2020 march 6).Applicability of Traditional Classification Techniques on Educational Data:International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8.Published By: Blue Eyes Intelligence Engineering & Sciences Publication.
14. Gowri G.S., Thulasiram R., and Baburao M.A.(2017).Educational data mining application for estimating student performance in WEKA environment: IOP Conf. Series: Materials Science and Engineering 263 (2017) 032002, 14th ICSET, IOP Publishing.
15. Viswanathan S. , Vengatesh S.K. (2021 April 5).Study Of Students' Performance Prediction Models Using Machine Learning:Turkish Journal of Computer and Mathematics Education,Vol.12 No.2 , 3085 – 3091
16. More SS, Narain B, Jadahv BT (2017) A comparative analysis of unimodal and multimodal biometric systems. In: International conference on innovative trends in engineering science and management (ITESM-2017).
17. Narain B, Zadgaonkar AS, Kumar S (2013) Impact of digital image processing on research and education. Natl Semin Work.
18. More S.S., Narain B., Jadhav B.T. (2021) Advanced Encryption Standard Algorithm in Multimodal Biometric Image. In: Rizvanov A.A., Singh B.K., Ganasala P. (eds) Advances in Biomedical Engineering and Technology. Lecture Notes in Bioengineering. Springer, Singapore. https://doi.org/10.1007/978-981-15-6329-4_7
19. SS More and B Narain,BT Jadhav.(2018)Data encryption standard algorithm in multimodal biometric image, Int. J. Comp. Sci. Eng.
20. U Sharma, B Narain, V Nohria - SPAST Abstracts.(2021)Hybrid Support Vector Machine and Distance Classifier in Breast Tumor Detection
21. P Singh, B Narain - NOVYI MIR.(2021).Student Satisfaction in Educational Organization using Machine Learning.
22. B Narain, AS Zadgaonkar, S Kumar Natl Semin Work.(2013).Impact of digital image processing on research and education.
23. M Nayak, B Narain - SPAST Abstracts(2021).The Online Retail Market Analysis for Social Development with Machine Learning.
24. On-Line Big Data Analysis using K-Mean and Modified K-Mean Algorithm with Machine Learning Techniques* B Narain 2021.
25. *M Nayak, B Narain - Computational Intelligence in Data Mining.(2020).Big Data Mining Algorithms for Predicting Dynamic Product Price by Online Analysis.