International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 5.22 (SJIF-2017), e-ISSN: 2455-2585
Volume 4, Issue 11, November-2018

A Study on Prediction of Student's Learning Ability Using Data Mining Techniques

Mrs. S.Menaka[1], Dr. G.Kesavaraj[2],

[1]*PG & Research Department of Computer Science,*
Vivekanandha College of Arts and Sciences for Women (Autonomous),
Elayampalayam, Tiruchengode 637 205, Tamil Nadu, INDIA

[1]*PG & Research Department of Computer Science,*
Vivekanandha College of Arts and Sciences for Women (Autonomous),
Elayampalayam, Tiruchengode637 205, Tamil Nadu, INDIA

Abstract*—The most difficult tasks in the education sector in India to predict student's educational performance due to a large volume of student information. In the Indian circumstance, we don't have any existing system by which analyzing and monitoring can be done to check the progress and performance of the student mostly in Higher education system. Every association has their own criteria for analyzing the presentation of the students. The reason for this happening is due to the lack of study on existing prediction techniques and hence to find the best prediction methodology for predicting the student academics evolution and performance. Another important reason is the lack in investigating the suitable factors which affect the academic performance and achievement of the student in exacting classes. So extremely recognize the problem, a fact literature analysis on predicting student's performance using data mining techniques is planned. The main objective of this article is to provide a huge data and understanding of different data mining techniques which have been used to predict the student development and performance and hence how these prediction techniques help to find the most important student quality for prediction. By improve the performance of the student in learning by using best data mining techniques. These also provide some benefits for faculties, students, educators and management of the institution.

***Keywords*—Educational Data Mining, Prediction Techniques, Student attributes, Classification.**

I. INTRODUCTION

In Indian learning system checking student's performance is a very important in advanced education. But we don't have any fixed criteria to assess the student presentation. Some institution student performance can be experimental by using internal evaluation and co-curriculum. There are lots of definitions of student educational performance calculation should be given in the text. Different authors are using different student factors/attributes for analyze student presentation. Most of the writer used CGPA, Internal evaluation, External evaluation, Examination ultimate score and additional co-circular actions of the student as forecast criteria.

Most of the Indian institution and universities using ultimate examination grade of the student as the student academic presentation criteria. The ultimate grades of any student depend on dissimilar attributes like internal evaluation, external evaluation, laboratory case work and viva-voce, sectional test. The presentation of the student depends upon how lots of grades a student achieve in the ultimate examination. Norlida Buniyamin, Pauziaah Mohamad Arsad et al. (2013) declared that what are the significance of educational analytics for an learning institution and how they work for the development of education. They also planned an intellectual suggestion involvement system to improve the student's performance and success in education.

This system uses two difficult student quality to measure the success and that is student mark and student data [2]. Zaidah Ibrahim and Daliela Rusli et al. (2007) stated that predicting student's performance is very serious for any educational institution because it is essential for the arrangement of new regulation and principles for the improvement of the education and standing. They used CGPA and demographic attributes of the first year student to expect their effect in the first year of learning in engineering [3].

Data mining techniques which are used in frequently education are known as Educational data mining. There are lots of data mining techniques are obtainable to predict the student presentation. Education data mining help to find the hidden information from a huge database of education setting, because at present lots of data are generate in educational institution associated to student [4]. Further, this secreted data can be used for presentation, failure and final effect calculation of the student. It also helps the instructor, organization and faculties to occupation according to the learning principles of the students. In fact data mining facilitate in the different field of education division [5]. So to suitably

understand the actual meaning of the data mining in education we need to do a efficient literature assessment on different occupation done by the different investigator.

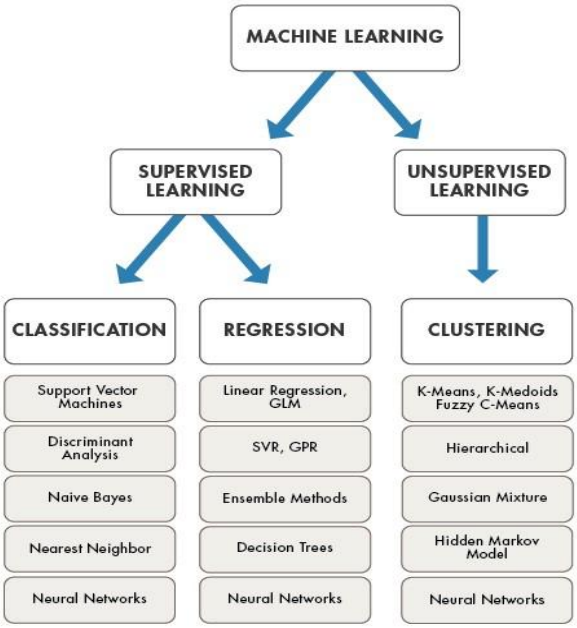

Fig 1.1 Shows the Classifications of Algorithms

Our main objectives to this future effort are:

i. To understand, analyze and then discover the difference between different forecast techniques of data mining in education.

ii. To classify and understand different student attribute which are mainly used for predict the student presentation.

iii. To recognize and identify with the different prediction techniques which are mainly used for predicting the student presentation

The above points are the most important focus of our learning. In segment 2, the major focus will be given on the attitude adopted for the configuration of research questions for this document and literature review. In segment 3-4, the major focus of the study is to find or recognize the significant factors on Predicting Student's presentation and prediction methods used for student presentation. In segment 5-6, the main focus is one the largely argument on the effect of the study and in the final conclusion and future work scale is given.

II. LITERATURE REVIEW

The main purpose of literature review is to find out latest techniques to work on the previous information set and then find out some latest data figure that. To do some relational study, the copy of more than 10 years should be taken into concern and then find out some awareness gaps between works done by the investigator. It helps to defend your study questions and gave some way for future study.

Formulation of Research: Research pattern is one of the important responsibilities when going for write any research paper. Before the development of any research try to understand and following the Kitchen hams steps. B. Kitchen hams, R Pretorius et al. (2010), declared that PIOC (People, interference, Outcome, Circumstance) are the most dangerous research for your research paper [6]. Factors which are measured when going to structure

TABLE 2.1. RESEARCH DEVELOPMENT CRITERIA

Criteria	Detail of targeted organization
Population	University, Engineering Institution (Private/Government), Students
Intervention	Data Mining Techniques/method used for prediction of student presentation and progress in education
Outcome	Student presentation accuracy, finalize prediction techniques
Context	University, Schools and Colleges(Private and Government)

From the over table, the whole thing is understandable about the research goal institute, techniques undertaken through the review, related outcome and precious institutions. Allowing for the above criteria in intelligence when frame in the research information, to restrict the range of this study with these investigate data.

i. Try to recognize those student attribute which are useful for predicting student educational presentation.

ii. Try to classify those data mining techniques which are regularly used for predicting student educational presentation.

After research formation we need to do the direct study on the associated subject and then need to locate out the research gaps between different works done the different researcher by using data mining techniques. Before start the literature review, everything should be clear in the mind of the investigator that what they want to search and how the search can be done.

MATERIALS FOR LITERATURE REVIEW:

Searched databases: Springer Link, Google Scholar, JET, IJEECSE, Elsevier and Science direct other computer science journals. Pointed sentences and keywords: Predicting student presentation, uses data mining techniques, use of data mining in education, Educational Data Mining attitude or techniques, Calculation of student result using data mining techniques. Publication periods are in use into concern: 2007 to July 2017. Types of text searched: papers, PDF, Full-length article with summary and keywords. Search Items: Journal articles, National and International Conferences, Topic connected communities like Educational data mining, Student's Learning Ability using Data Mining Techniques.

IMPORTANT FACTORS OF STUDENTS LEARNING ABILITY

The calculation of SAP is based on different factors of student's like an character, society, emotional and environmental values. During previous few years a lot of researches have been accepted to predict student's educational presentation. So in this segment, we are taking few research articles into concern and then analyze them for difficult student's factors which change the student academics calculation. Almost 40-50 research articles, book chapters are consider for evaluation.

Farhana Sarkar and Hugh C Davis (2013) in his research show that the institutional interior data sources (IDS) and exterior data source (EDS) gave the top result than the form based on only institutional interior student database [4].

In a new revise, M. D. Angelene (2013) used internal evaluation Test mark, Assignment submission and Grade, accurate answer, self-assurance and concentration in the exacting course and Degree aim for prediction of student's educational presentation [5].

Abier Badre El Den Ahamed et. al. (2014) in his revise used the path of the student, HSD, mid-term marks, Lab experiment grade, seminar presentation, assignment, presence, training, student involvement for calculation SAP [6].

Fadhelah Ahamad and Aziwa Abdula Azez (2015) collect information from the database of educational Department, UniiSZA that stored in Informix Database Management Systems (DBMS). They additional used nine dissimilar parameters like gender, event and homeland, GPA, family proceeds, university entrance mode, grades Malayalam Language, English, and Mathematics [7].

Maashael A. Al-Barraak and Monaa S. Al-Razgaan (2015) collect dataset of student's from the Information Technology branch at Kiin Saaud University, Saudi Arabia for their investigation. They additional used the different attribute for the prediction like student ID, student name, student marks in three different quiz's, Mid-Term1, Mid-Term2, assignment, class, semester exam, and overall points obtained in Data structure lessons of computer science branch [8].

Ediin Osmaanbegović and Mirzaa Suljic (2012) collect data from survey in the middle of first-year students and the information taken for the period of the enrollment at the University of Tuzla. They further used the difficult attribute for the prediction like Gender, Family, Travel, High School, GPA, Entrance exam, scholarship, instance, Materials, the Internet, Grade value, salary [9].

Rahela Asiff and Mahamood K. Pathaan (2014) in his revise they used four academic batches of Computer Science & Information Technology (CS&IT) branch at NED University, Pakistan. They used HSC marks, marks in MPC, Maths marks in HSC, marks in a variety of subject studied in the regular classes of a programming language, CSA, Logic design, OOPS, RDBMS, ALP, SAD, Data Structure, Visual Basic etc for their study [10]. Mohammed M.

Jyoty Bansodee (2016) for predicting student academic presentation collected information from Shah and Anchor Kutchhi Polytechnic, Chembur, Mumbai. They consider student attributes like parent's learning, parent' s profession, group, SSC board, entrance type, SSC medium, SSC class, first-semester result, second-semester, third-semester, forth-semester, the fifth-semester and sixth-semester result as most essential attributes [15].

Mukesh Kumar (2017) collect data to consider student attributes like parent's learning, parent' s profession, group, SSC board, entrance type, SSC medium, SSC class, first-semester result, second-semester, third-semester, forth-semester, the fifth-semester and sixth-semester result as most essential attributes[25].

After review more or less 20-25 research papers, we found that in most of the cases, student's factors which influence the SAP are gender, high school grade, student's parental education, financial environment, living place, standard of teaching, student's family situation, students' earlier semester marks, class examination grade, seminar presentation, assignment presentation, general ability, present in class and lab effort, Interest in exacting course, Study

performance, Engage Time and Family Support for study, admission style, earlier schools marks, accommodation form, parent's ability, parent's profession. All these attributes fall into different categories like individual, family, educational, Institutional and communal.

III. RESULTS AND DISCUSSION

In Educational data mining field, making a prediction about student academic presentation is usually completed. To build a predictive modeling we need to obtain different data mining techniques into concern like classification, clustering, association rule mining and regression analysis. In more or less every research paper, the only classification algorithm is taken into concern for predicting student academic presentation. There are so many classification techniques offered for prediction but we are taking into concern only decision tree, Naive Bayes, Support Vector Machine (SVM), Artificial Neural Networks (ANN), K-Nearest Neighbor, C4.5 and J48 etc.

TABLE-4.1 GAVE A SHORT RESULT OF DIFFERENT RESEARCH PAPERS WITH THEIR AUTHOR'S NAME, MAJOR ATTRIBUTES SUPPORTIVE FOR PREDICTION ACCURACY WITH DIFFERENT DATA MINING ALGORITHM USED.

Authors	Attributes which affect prediction accuracy	DT	NB	KNN	ANN
F Sarkar et.al.	Internal attributes + Students fist semester mark(Models1)	-	-	-	76.5
F Sarkar et. al.	Internal + External attributes + Student's first semester mark(Model2)	-	-	-	74.5
F Ahamad et. al.	Gender, race, homeland, GPA, Family proceeds, SPM grades	-	69.0	-	65.8
Maashael A et. al.	First Midterm Examination(Predict Students Failure)	-	94	-	89.8
R Asif et.al.	HSC, MPC and HSC marks, Pre-unit marks, marks in different courses	72	89.6	74	68.6
M Koutina et. al.	Gender, Age, Marital Status, No of children, Occupation, Job related with PC, Bachelor, Another master, Comp literacy, Bachelor in informatics	69.5	100	100	-

From the exceeding table, we find that Maaria Koutina, in his research found the 100% exactness with Naive Bayes (NB) and K-Nearest Neighbor algorithm (KNN) [24]. She represented their effect in Table 3 under "Total accuracy (%) of re-sample information and attribute selection". For prediction student academic performance they used attributes like Gender, Age, Marital Status, Number of children, Occupation, Job related with the computer, Bachelor, Another master, Computer literacy, Bachelor in informatics.

IV. DISCUSSION ON PREDICTING STUDENT'S STUDY

In this exacting segment, we will talk about the major finding of our meta-analysis. In this meta-analysis, we find that generally used data mining algorithm for SAP is Decision Tree (DT), Naive Bayes (NB), Artificial Neural Networks (ANN) and K-Nearest Neighbor (KNN). In Decision tree algorithm the maximum and minimum exactness for predicting student's academic presentation are 99.9% and 66.8% respectively. To find the maximum prediction exactness Maria Goga and Nicolae Goga used the combination of student's aspect like family, PEP, EES, end of first gathering result [21]. In Naive Bayes algorithm, the maximum and minimum precision for predicting student's academic presentation are 100% and 63.3% correspondingly. Maaria Koutina et. al. used the different grouping of student's attribute like Gender, Age, Marital Status, Number of children, Occupation, Job connected with the computer, Bachelor, a different master, Computer literacy, Bachelor in informatics for receiving maximum exactness [24]. In rule-based algorithm, the maximum and minimum accuracy for predicting student's academic performance are 96.7% and 55.0% correspondingly. Fig. 4.1 shows the prediction exactness that uses classification method grouped by algorithms for predicting student's presentation since 2012 to 2017.

TABLE4.3 STUDENT ACADEMIC REPRESENTATION OF RESULT ANALYSIS

Data Mining Techniques	DT	NB	KNN	NN
Highest Accuracy	99.9%	100%	100%	89.8%
Lowest Accuracy	66.8%	63.3%	74%	67.6%
Average Accuracy	83.35%	81.65%	87%	78.7%

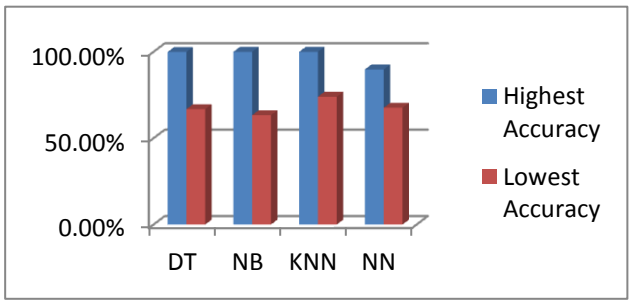

Fig. 4.1 Student Academic Performance Prediction Grouped by Algorithm used

V. CONCLUSION

At current research in educational data mining generate lots of awareness in the research society. Because predicting student academic presentation, predicting educational failure student in next to future, predicting institution placement and admission in a new educational year is most useful for educators and management and educational plan maker. It also used for humanizing the teaching-learning procedure in the institution as well. This paper has review a lot of research papers, the object on predicting student's academic presentation with selected quality and an logical algorithm used. In most of the cases, CGPA and the internal marks of the student in academic are essential attributes for prediction of effect. In one of the study paper author's find 100% exactness for their prediction with a arrangement of different attributes like Gender, Age, Marital Status, Number of children, career, Job related with the computer, Bachelor, a new master, Computer literacy, Bachelor in informatics. It will really help to advance our education structure to ensure the usual presentation of the student.

VI. REFERENCES

[1]. Mihai Dascalu and Elvira Popescu et. al., Predicting Academic Performance Based on Students' Blog and Microblog Posts, Springer International Publishing Switzerland 2016 K. Verbert et al. (Eds.): EC-TEL 2016, LNCS 9891, pp. 370–376, 2016. DOI: 10.1007/978-3-319-45153-4_29.

[2]. U. bin Mat, N. Buniyamin, P. M. Arsad, R. Kassim, An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention, in: Engineering Education (ICEED), 2013 IEEE 5th Conference on, IEEE, 2013, pp. 126–130.

[3]. Randa Kh. Hemaid and Alaa M. El-Halees, Improving Teacher Performance using Data Mining, International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015.

[4]. Farhana Sarker, Thanassis Tiropanis and Hugh C Davis, Students' Performance Prediction by Using Institutional Internal and External Open Data Sources, http://eprints.soton.ac.uk/353532/1/Students' mark prediction model.pdf, 2013.

[5]. D. M. D. Angeline, Association rule generation for student performance analysis using an apriori algorithm, The SIJ Transactions on Computer Science Engineering & its Applications (CSEA) 1 (1) (2013) p12–16.

[6]. Abeer Badr El Din Ahmed and Ibrahim Sayed Elaraby, Data Mining: A prediction for Student's Performance Using Classification Method, World Journal of Computer Application and Technology 2(2): 43-47, 2014.

[7]. Fadhilah Ahmad, Nur Hafieza Ismail and Azwa Abdul Aziz, The Prediction of Students' Academic Performance Using Classification Data Mining Techniques, Applied Mathematical Sciences, Vol. 9, 2015, no. 129, 6415 - 6426HIKARI Ltd, www.m-hikari.com http://dx.doi.org/10.12988/ams.2015.53289.

[8]. Mashael A. Al-Barrak and Mona S. Al-Razgan, predicting students' performance through classification: a case study, Journal of Theoretical and Applied Information Technology 20th May 2015. Vol.75. No.2.

[9]. Edin Osmanbegović and Mirza Suljic, DATA MINING APPROACH FOR PREDICTING STUDENT PERFORMANCE, Economic Review – Journal of Economics and Business, Vol. X, Issue 1, May 2012.

[10]. Raheela Asif, Agathe Merceron, Mahmood K. Pathan, Predicting Student Academic Performance at Degree Level: A Case Study, I.J. Intelligent Systems and Applications, 2015, 01, 49-61 Published Online December 2014 in MECS (http://www.mecs-press.org/) DOI: 10.5815/ijisa.2015.01.05.

[11]. Mohammed M. Abu Tair, Alaa M. El-Halees, Mining Educational Data to Improve Students' *Literature Survey on Student's Performance Prediction in Education using Data Mining Techniques* Performance: A Case Study, International Journal of Information and Communication Technology Research, ISSN 2223-4985, Volume 2 No. 2, February 2012.

[12]. Azwa Abdul Aziz, Nor Hafieza Ismailand Fadhilah Ahmad, First Semester Computer Science Students' Academic Performances Analysis by Using Data Mining Classification Algorithms, Proceeding of the International Conference on Artificial Intelligence and Computer Science(AICS 2014), 15 - 16 September 2014, Bandung, INDONESIA. (e-ISBN978-967-11768-8-7).

[13]. Kolo David Kolo, Solomon A. Adepoju, John Kolo Alhassan, A Decision Tree Approach for Predicting Students Academic Performance, I.J. Education and Management Engineering, 2015, 5, 12-19 Published Online October 2015 in MECS (http://www.mecs-press.net) DOI: 10.5815/ijeme.2015.05.02.

[14]. Dr Pranav Patil, a study of student's academic performance using data mining techniques, international journal of research in computer applications and robotics, ISSN 2320-7345, vol.3 issue 9, pg.: 59-63 September 2015.

[15]. Jyoti Bansode, Mining Educational Data to Predict Student's Academic Performance, International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169, Volume: 4 Issue: 1, 2016.

[16]. R. Sumitha and E.S. Vinoth kumar, Prediction of Students Outcome Using Data Mining Techniques, International Journal of Scientific Engineering and Applied Science (IJSEAS) – Volume-2, Issue-6,June 2016 ISSN: 2395-3470.

[17]. Karishma B. Bhegade and Swati V. Shinde, Student Performance Prediction System with Educational Data Mining, International Journal of Computer Applications (0975 – 8887) Volume 146 – No.5, July 2016.

[18]. Mrinal Pandey and S. Taruna, Towards the integration of multiple classifiers pertaining to the Student's performance prediction, http://dx.doi.org/10.1016/j.pisc.2016.04.076 2213-0209/© 2016 Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/ licenses/by-nc-nd/4.0/)

[19]. Maria Goga, Shade Kuyoro, Nicolae Goga, A recommender for improving the student academic performance, Social and Behavioural Sciences 180 (2015) 1481 – 1488.

[20]. Anca Udristoiu, Stefan Udristoiu, and Elvira Popescu, Predicting Students' Results Using Rough Sets Theory, E. Corchado et al. (Eds.): IDEAL 2014, LNCS 8669, pp. 336–343, 2014. © Springer International Publishing Switzerland 2014.

[21]. Parneet Kaur, Manpreet Singh, Gurpreet Singh Josan, Classification and prediction based data mining algorithms to predict slow learners in education sector, Procedia Computer Science 57 (2015) 500 – 508.