A Survey on Predicting Student Performance

A.Dinesh Kumar ,Dr.V.Radhika

Sri Krishna Arts and Science College
Coimbatore, India

Abstract-Education is a vast and important subject, one that cannot be addressed adequately in this space. In recent years the amount of data stored in educational database is growing rapidly. The stored database contains hidden information about improvement of student's performance and behavior. The ability to predict the student performance in education is very important in educational environments. Student's academic performance in educational environment is based upon the psychological and environmental factor is predicted by different educational data mining techniques.

Keywords: Educational data mining, Prediction, classification.

I. INTRODUCTION

Data mining is also called Knowledge Discovery in databases (KDD), is the field of determine useful information from huge amounts of data. Data mining has been applied numerous arenas including bank, medical, census data, airline passenger records and super market, bioinformatics, business, genetics etc.,

Currently the data mining techniques has been used in educational environments. Applying data mining technique in Educational setting is called as educational data mining (EDM).

Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in [1].

Educational data mining (EDM) is concerned with developing, researching, and applying computerized methods to detect patterns in large collections of educational data that would otherwise be hard or impossible to analyze due to the enormous volume of data within which they exist.[2].

There are many popular methods are there in Educational Data Mining. Some of them are widely used as prediction, classification, clustering, and regression.

Prediction attempts to form patterns that permit it to predict performance or learning outcomes based on the data from the available data. The objective is to predictive a student's course activities. The most often used techniques for this type of goal are classification, clustering and association. In Educational Data mining, prediction has been used for foretelling student performance and for identifying student behaviors. The goal of prediction is to infer a target attribute or single aspect of other aspects of the data [16].

Classification is one of the most frequently studied problems by DM and Machine learning researchers. Classification is a form of data analysis that extracts models describing data classes. A classifier, or classification model, predicts categorical labels (classes) [17]. The organization of the paper presented in section II.

II. RELATED WORK

Ryan S.J.D. Baker and Kalinayacef reviewed the history and trends in the field of educational data mining (EDM) 2009. They focused about increased importance on prediction, the development of work using existing models to make scientific discoveries. [1]

Data Mining can be used in educational field to develop our understanding of learning process to focus on categorizing, mining and assessing variables related to the learning process of students as described by Alaa el-Halees [4]. Mining in educational environment is called educational data mining.

Bharadwaj and Pal [2] found that the university student's data like attendance, class test, seminar and assignment marks from the student's prior database, to predict the performance at the end of the semester.

Hijazi etal [5] conducted a study on the student performance by selecting a sample of 300 students (225 males, 75 females) from a group of colleges affiliated to Punjab university of Pakistan. The hypothesis that was stated as Student's attitude towards attendance in class, hours spent in study on daily basis after college, students' family income, students' mother's age and mother's education are significantly related with student performance" was framed. By means of simple linear regression analysis, it was found that the factors like mother's education and student's family income were highly interconnected with the student academic performance

Pandey and Pal [6] conducted study on the student performance based by selecting 600 students from different colleges of Dr. R. M. L. Awadh University, Faizabad, India. By means of Bayes Classification on category, language and background qualification, it was found that whether new comer students will performer or not.

Pandey and Pal [7] conducted study on the student performance based by selecting 60 students from a degree college of Dr. R. M. L. Awadh University, Faizabad, India. By means of association rule they find the interestingness of student in selecting class teaching language.

Romero and Ventura [8], have a survey on educational data mining between 1995 and 2005. They concluded that educational data mining is a promising area of research and it has a specific requirements not presented in other domains. Thus, work should be oriented towards educational domain of data mining.

Er.Rimmy Chuchra and [9] applied decision tree, clustering, Neural network techniques to evaluate student performance based by selecting the student from Sri Sai University Engineering, Phagwara, India. By using these techniques he found that teachers can easily evaluate the student performance.

Bharadwaj and Pal [10] conducted study on the student performance based by selecting 300 students from 5 different degree college conducting BCA course of Dr.R.M.L. Awadh University, Faizabad, India. By means of Bayesian Classification method on 17 attributes, it was found that the factors like student's grade in senior secondary exam, living location, medium of teaching, mothers qualification, fathers qualification, students other habit, family annual income and student's family status were highly correlated with the student academic performance.

Yadav and Pal [11] obtained VBS University student's data like Branch, Category, Student grade in high school, Admission type, medium and family size from the previous student database to predict the students who are likely to fail with the help of ID3,C4.5 and CART algorithm. It was observed that C4.5 is the best algorithm for predict student result.

R.Shanmuga Priya [12] conducted study on improving the student's performance using Educational Data Mining based by selecting 50 students from Hindustan College of Arts and Science, Coimbatore, India. By using decision tree classification on 8 attribute, it was found that the class test, seminar, attendance, lab practicals are used to predict the student performance. This prediction will help to the teacher to give special attention of students and improve student confidence on their studies.

C.Marquez, et al[13] had done research on identifying the factors that affect the low performance of students at different educational levels. They obtained 670 middle school students data from Zacatecas, Mexico. By using 10 classification algorithm and 15 selected attributes they found sociological, economic or educational characteristics that may be more relevant in the prediction of low academic performance in school students.

Ahmed and Ibrahim [14] describes the most useful data mining techniques is classification; the classification task is used to predict the final grade of students by using decision tree method. Their study will help to identify the students which need special attention to reduce failing ration and taking appropriate action.

Sajadin et al [15] conducted a research on analyze the relationships between student's behavioral and their success and to develop the development of student performance predictor by using Smooth Support Vector Machines (SSVM) classification and kernel k-means clustering techniques. They find out there is a strong relation between mental condition of student and their final academic performance.

Dorina Kabakchieva [18] has conducted a research is focused on the development of data mining models for predicting student performance, based on their personal, pre university and university-performance characteristics by using One Rule Learner, Decision Tree, Neural Network and K-Nearest Neighbour classification methods, she found Neural Network model is predict with higher accuracy than other three models. The trends of EDM methods are explained in section III

III. TRENDS IN EDUCATIONAL DATA MINING METHODS

Romero and Ventura's survey of Educational Data Mining research from 1995 to 2005, 60 papers was stated that developed EDM methods to answer research questions of applied interest. Relationship mining methods of various types were the most prominent type of EDM research between 1995 and 2005. 43% of papers in those years involved relationship mining methods.

Prediction was the second most prominent research area, with 28% of papers in those years involving prediction methods of various types. Human judgment/exploratory data analysis and clustering followed with (respectively) 17% and 15% of papers [8]. The full distribution of methods across papers is shown in Figure 1.
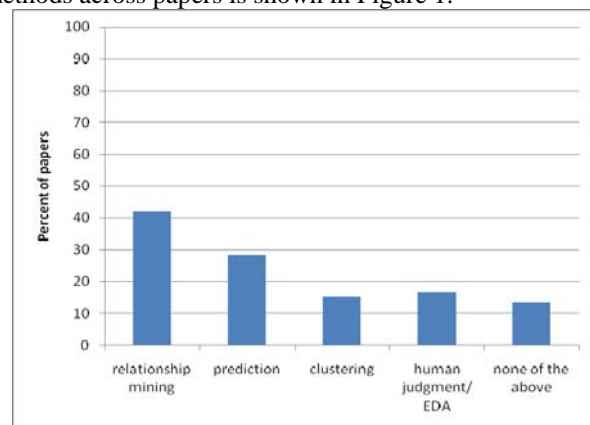

Figure-1.

A very different pattern is seen in the papers from the first two years of the Educational Data Mining conference [Baker et al. 2008; Barnes et al. 2009], as shown in Figure 2.
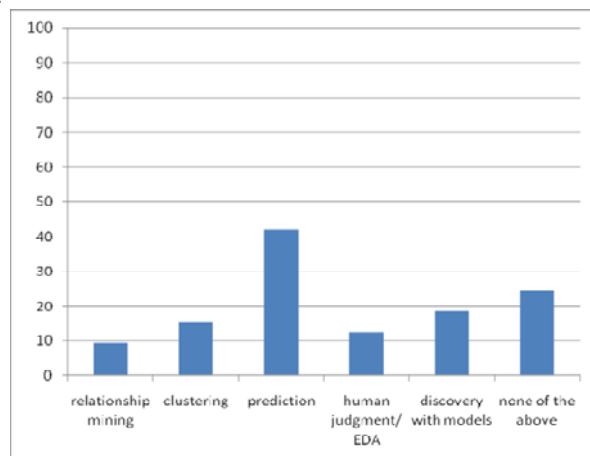

Figure-2

Whereas relationship mining was leading between 1995 and 2005, in 2008- 2009 it slipped to fifth place, with only 9% of papers involving relationship mining. Prediction, which was in second place between 1995 and 2005, moved to the leading position in 2008-2009, representing 42% of EDM2008 papers. Human judgment/exploratory data analysis and clustering remain in approximately the same position in 2008-2009 as 1995-2005, with (respectively) 12% and 15% of papers [1].

IV. CONCLUSION

This paper surveys the predictive model in data mining and current trends in prediction in data mining. Prediction on student academic performance is one of the most modern topics in educational data mining. In Previous research some of the factors were affected the student performance, such as student family income, student family size, mother's qualification, student learning behavior are highly correlated with student performance. The main objective of Educational Institution is to provide the best quality education to its students and to improve their behavior. Prediction of student academic performance is helps to teachers to predict about student success and failure in examination. The prediction model acts like a warning system to identify potential weak students, so that the teachers can take an appropriate action towards them. Otherwise, they can give proper advices to prevent failure in the examinations. Predict the student performance and help them to improve their performance.

REFERENCES

1. BakerRSJd, Yacef K. The state of educational datamining in 2009: A review and future visions. J EduData Min 2009.
2. Romero C, Ventura S, Pechenizky M, Baker R. Handbook of EducationalData Mining. Data Mining and Knowledge Discovery Series. Boca Raton,FL: Chapman and Hall/CRCPress; 2010.
3. B.K. Bharadwaj and S. Pal. "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security (IJCSIS), 2011.
4. Alaa el-Halees, "Mining students data to analyze e-Learning behavior: A Case Study",2009.
5. Hijazi, S. T., and Naqvi, R.S.M.M.," Factors Affecting Student's Performance: A Case of Private Colleges", Bangladesh e-Journal of Sociology,2006.
6. Pandey, U. K. and Pal, S., "Data Mining: A prediction of performer or under performer using classification", (IJCSIT) International Journal of Computer Science and Information Technology, 2011.
7. Pandey, U. K. and Pal, S., "A Data Mining View on Class Room Teaching Language", (IJCSI) International Journal of Computer Science Issue, Vol. 8, Issue 2, March -2011.
8. Romero, C. and Ventura, S. (2007) 'Educational data mining: A Survey from 1995 to 2005', Expert Systems with Applications (33), pp. 135-146.
9. Er.Rimmy Choura"Use of Data Mining Techniques for the evaluation of student performance: A Case Study", Vol 1 Issue 3 October 2012.
10. B.K. Bharadwaj and S. Pal. "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security(IJCSIS), Vol. 9, No. 4, pp. 136-140, 2011.
11. S.K.Yadav and S.Pal. "Data Mining: A prediction for performance improvement of Engineering students using classification", World of Computer Science and Information Technology Journal (WCSIT), Vol. 2, No. 2, pp. 51-56, 2012
12. K.Shanmuga Priya" Improving the student's performance using Educational data mining", International Journal of Advanced Networking and Application, Vol.4,pp-1680-   1685 (2013)
13. C.Marquez-Vera, C.Romero and S.Ventura."Predicting School Failure Using Data Mining" -2011
14. Din Ahmed, Ibrahim Sayed. "Data Mining: A Prediction for student's performance using classification method", World journal of Computer Application and Technology-2014
15. Sajadin Sembering, M.Zarlis, " Prediction of student academic performance by an application of data mining techniques", International conference on management and Artificial Intelligence-2011.
16. Romero C, Ventura C," Data mining in education", IEEE Trans Sys Man Cybern. Vol 3-2013.
17. Romero C, Ventura C," Data Mining algorithm to Classify Students", IEEE Trans Sys Man Cybern.
18. Dorina Kabakchieva," Student Performance Prediction by Using Data Mining Classification Algorithms", International Journal of Computer Science and Management Research Vol 1 Issue 4 November 2012.