RESEARCHERID
THOMSON REUTERS

[Srinivasan* *et al.*, 6(7): July, 2017]
IC™ Value: 3.00

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

AN ANALYSIS ON PREDICTION OF SUPERIOR SCHOOLING APPRENTICE ACCOMPLISHMENT IN INSTRUCTIVE DATA MINING

Srinivasan B.[1] Abinaya V.[2]
ASSOCIATE PROFESSOR[1] RESEARCH SCHOLAR[2]
Department Of Computer Science & Engineering
Gobi Arts and Science College, Gobichettipalayam, Erode, TamilNadu 638 453, India
DOI: 10.5281/zenodo.834574

ABSTRACT

Educational Data Mining is used to predict the potential learning behavior of the learner. One of the biggest challenges that higher education faces today is predicting the paths of students. Predicting the performance of a student is a great concern to the higher education managements. [15] The possibility of this paper is to identify the factors influencing the performance of students in final examinations and find out a suitable data mining algorithm to predict the grade of students. [15] This work will help the educational institutions to classify the students who are at risk and to provide better further training for the weak students. Data mining is the better tool to predict the end result of the student. [13]

KEYWORDS: Educational data mining, Data Mining Algorithm, Naïve Bayesian Classifier.

I. INTRODUCTION

Measuring of academic performance of students is challenging since students academic performance hinges on diverse factors like private, socio-monetary, mental and other environmental variables. The scope of this paper is to predict the student marks and what are the factors that influence the performance of the learner. [1]

In Tamilnadu, the higher secondary education consists of two years of schooling, preceding ten years of basic education and followed by higher education. The higher secondary education is important in a student's life because it is one of the factors that are going to decide the future of the student. Based on the mark in higher secondary examination, they are going to get college education. [15]

Data mining provides many tasks that could be used to study the student's performance. In this paper the classification task is used to evaluate performance of a student and as there are many approaches that are used for data classification, the Decision Tree and Naive Bayes, Multilayer Perception methods is used in data classification. [15][13]

Educational data mining is one of the applications of data mining. Data mining is used to find the hidden pattern from a huge data set and then apply that hidden patterns for the decision making in future. Its application is not limited to education but also covers fields like sales, retail, transportation, sports, marketing etc…In education; these data mining techniques are used to predicate slow learners, dropout, under-performer etc… and hence provide timely help to those students who are the problem in learning. [9]

This study is more useful for identifying weak students and the identified students can be individually assisted by the educators so that their performance is better in future. This study investigates the accuracy of some classification techniques for predicting performance of a learner. [15]

The main objectives of study are,

- Classification of extremely influencing predictive variables on the academic performance of higher secondary students.
- Discover the most excellent classification algorithm on learner data.
- Predict the position at superior secondary oral exam.

Figure: Shows the Cycle of applying Data Mining in Educational Setting. To design, plan and build the Educational Systems and Maintenance of the Educational Systems. Interactions of data and course information are used in Data Mining Techniques Such as Classification and Clustering, Outlier and Association. The students are participe and communicate to the Educational Systems and Data Mining techniques are showing recommendations to students and showing Discovered Knowledge to Educators.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

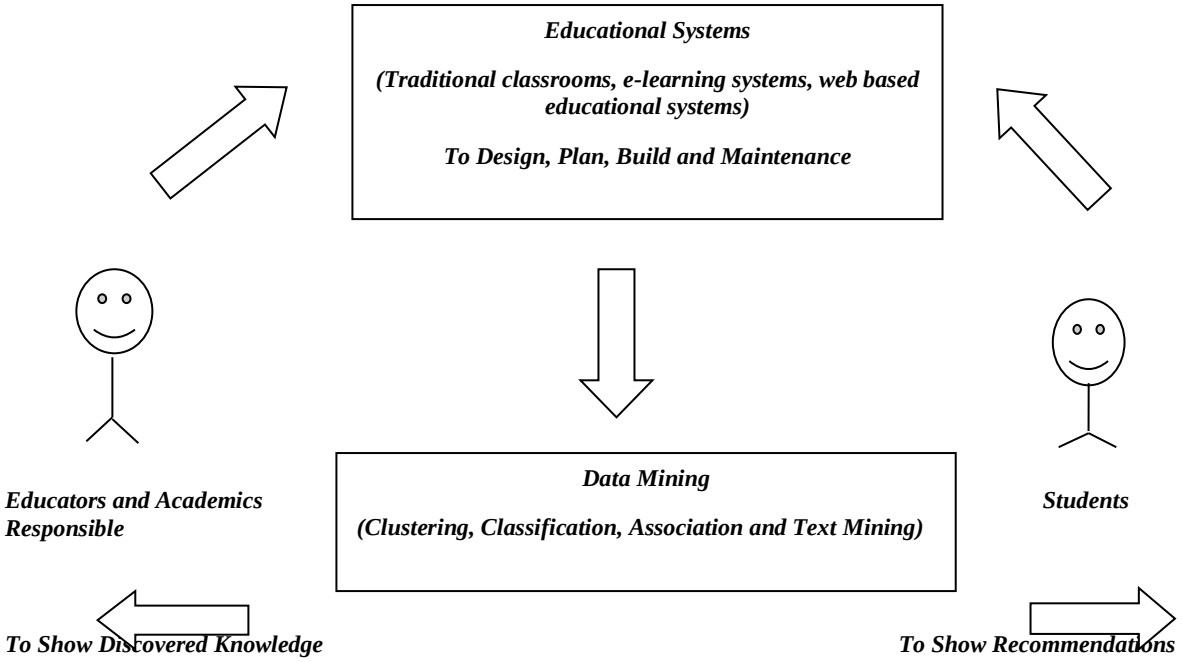

Figure 1. Flow Cycle of Educational Data Systems

II. REVIEW OF LITERATURE

A number of reviews pertaining to not the diverse factors like personal, socio-economic, psychological and other environmental variables that influence the performance of students but also the methods that have been used for the performance prediction are available in the literature and a few specific studies are listed below for position. [15]

The literature review reveals that these problems have been of interest for various researchers during the last few years. The development of data mining models for predicting student performance at various levels, and comparison of those models, are discussed in a number of research papers. In 2000 the results of a study are described aimed at finding weak students and involving them in additional courses for advanced support by extracting association rules from data. [6]

Use of data mining in education is tremendous. But still, lots of researchers data mining techniques for the betterment of education. As already told it's a broad field and not limited to the present discussion like the prediction of slow learner in a class. [9]

M.Ramaswami and R.Bhaskaran[14] proposed CHAID prediction model to analyze the interrelation between variables that are used to predict the outcome of the performance at higher secondary school education. The features like medium of instruction, marks obtained in secondary education, location of school, living area and type of secondary education were the strongest indicators for the student performance in higher secondary education. The CHAID prediction model of student performance was constructed with seven class predictor variable.

Nguyen Thai-Nghe, Andre Busche, and Lars Schmidt-Thieme[11] presented to applied machine learning techniques to improve the prediction results of academic performances for two the real folder study. Three methods have been use to compact with the class imbalance problem and all of them show satisfactory results. They first re balanced the datasets and then used both cost-insensitive and sensitive learning with SVM for the small datasets and with Decision Tree for the superior datasets. The models are originally deploy on the local web.

Arockiam et al. [2] presented FP Tree and K-means clustering technique for finding the similarity between urban and rural students training skills. FP Tree mining is useful to filter the patterns from the dataset. K-means clustering is used to determine the programming skills of the student. The study clearly indicate that the country and the town students differ in their programming skills. The huge proportions of urban students are

good in programming skill compared to rural students. It divulges that academicians provide extra training to urban students in the programming subject.

Cortez and Silva [3] has proposed to predict failure in the two core classes (Mathematics and Portuguese) of two secondary school students from the Alentejo region of Portugal by utilizing 29 predictive variables. Four data mining algorithms Such as Decision Tree (DT), Random Forest (RF), Neural Network (NN) and Support Vector Machine (SVM) were applied on a data set of 788 students, who appear in 2006 oral exam. It was reported that DT and NN algorithms had the predictive accuracy of 93% and 91% for two-class dataset (pass/fail) correspondingly. It was also report that both DT and NN algorithms had the predictive accuracy of 72% for a four class dataset.

Ramaswami et al. [14] proposed in focus on developing predictive data mining model to identify the slow learners and study the influence of the dominant factors on their educational presentation, using the well-liked CHAID decision tree algorithm.

Dorina Kabakchieva [4] presented to define EDM as techniques which are helpful in finding the unknown facts from a database, which are impossible to find manually and hence that information, are effectively used in the education setting. It is used to increase the student retention rate, improve the educational standard, and help administrator for setting new rules and regulation for improving educational standard.

Lalit Dole et.al, [6] proposed to analyze the student's data for predicting their future learning behavior and hence the result. They also predict the student result and warn them that they are at risk of failure in final examination and provide timely help to them.

V. Ramesh. et. al, [15] presented the survey methodology to make the final dataset with some significant variable of students and with experimental methodology tried to found only those variables which influencing the final result of the student. They applied SMO, J48, REPTree, Naïve Bayes and multilayer perception techniques for their experimentation purpose. After analysis, they found that factor like parent's occupation plays a very important role in student performance.

Applying EDM techniques for knowledge discovery is important for the teachers, management, and student. They all are using this knowledge for the improvement of the education system. [9] Teachers are using this knowledge for improving their teaching standard and the student is using to improving their learning skill. Management of the institution is using this knowledge for improving infrastructure standard; provide basic facilities to the student and decision making. [9][12]

III. CONCLUSION

There is lot of drawbacks in education system like midterm evaluation system use. It is really not understood that why midterm evaluation is taken throughout the year. In this paper, classifications are used for prediction on the dataset of students, to predict and analyze student's performance as well slow learners among them.

Data mining techniques allow a high level extraction of knowledge from raw data, offering interesting possibilities for the education domain. Furthermore, we intent to enlarge the experiments to collect additional features like psychological factors which disturb the student, motivational hard work in use by the teacher and e-learning materials available to the students. A hybrid model, which utilizes the full impact of variable selection and its consequences, is being worked out.

In future, integration of data mining techniques with DBMS and E-Learning techniques is merged together on different datasets to find accuracy and predictors of desired results. Educators with no expertise in data mining can also apply their hands in these fields. Also some new factors can be applied to improve the student's performance, learning and retention capabilities among them.

IV. REFERENCES

[1] Arun K Pujari, "Data Mining Techniques", University Press, Education 2001.
[2] Arockiam. L, Charles. S, Arul Kumar, "Deriving Association between Urban and Rural Students Programming Skills", International journal on Computer Science and Engineering, 2010, Vol. 02, No. 03
[3] Cortez.P, and Silva. A, "Using Data Mining to Predict Secondary School Student Performance", In EUROSIS, A. Brito and J. Teixeira (Eds.), 2008, pp.5-12.
[4] Dorina Kabakchieva (2012), "Student performance by using data mining classification Algorithms", International Journal of Computer Science and Management Research Issues Vol. 1, Issue 4, November 2012.
[5] Jaiwei Han, Michelinne Kamber, "Data Mining: Concepts and Techniques".

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

[6] Lalit Dole, Jayant Rajurkar, "A Decision Support System for Predicting Student Performance", International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 2, Issue 12, December 2014.

[7] Monika Goyal and Rajan Vohra, "Applications of Data Mining in Higher Education", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012.

[8] Moriana.J.A, Alos.F, Alcala.R, Pino.M.J, Herruzo.J, and Ruiz.R, "Extra Curricular Activities and Academic Performance in Secondary Students", Vol.4, No.1, 2006, Electronic Journal of Research in Educational Psychology, pp.35-46.

[9] Mukesh Kumar, Shankar Shambhu & Punam Aggarwal, "Recognition of Slow Learners Using Classification Data Mining Techniques", Imperial Journal of Interdisciplinary Research (IJIR) ISSN: 2454-1362 Vol-2, Issue-12, 2016.

[10] Mrs. Mythili.M.S, Dr. Mohamed Shanavas.A.R, "An Analysis of students performance Using Classification algorithms", IOSR Journal of Computer Engineering, Volume 16, Issue 1, January 2014.

[11] Nguyen Thai-Nghe, Andre Busche, and Lars Schmidt-Thieme, "Improving Academic Performance Prediction by Dealing with Class Imbalance", 2009 Ninth International Conference on Intelligent Systems Design and Applications.

[12] Dr. Nithya.P, Umamaheswari.B, Umadevi.A(2016), "A Survey on Educational Data Mining in Field of Education", Department of CS, PSG CAS, Coimbatore. Tamil Nadu, India. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 5 Issue 1, January 2016.

[13] Parneet Kaur,Manpreet Singh,Gurpreet Singh Josan, "Classification and prediction based Data Mining Algorithms to predict slow learners in education sector", Scholar, Department Of CSE, Punjab Technical University, Jalandhar 144603, India, Assistant Professor, Department CSE&IT, GNDEC, Ludhiana, Punjab, India, Assistant Professor, Department Of CSE & IT, Punjabi University, Patiala, Punjab, India. 3rd International Conference on Recent Trends in Computing 2015(ICRTC-2015)

[14] Ramaswami.M and Bhaskaran.R, "A CHAID Based Performance Prediction Model in Educational Data Mining", International Journal of Computer Science Issues Vol. 7, Issue 1, No. 1, January 2010.

[15] Ramesh.V, Parkavi.P, Ramar.K(2013), "Predicting student performance : A statistical and Data mining approach", International journal of computer applications, Volume 63-no. 8, PP 35-39.

CITE AN ARTICLE

B, Srinivasan , and Abinaya V. "AN ANALYSIS ON PREDICTION OF SUPERIOR SCHOOLING APPRENTICE ACCOMPLISHMENT IN INSTRUCTIVE DATA MINING." *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY* 6.7 (2017): 873-76. Web. 25 July 2017.