

# A REVIEW PAPER ON PREDICTION ANALYSIS: PREDICTING STUDENT RESULT ON THE BASIS OF PAST RESULT

Priyanka Dhamija<sup>1\*</sup>, Rainu Nandal<sup>2</sup>, Harkesh Sehrawat<sup>3</sup>

\*Corresponding Author

Department of Computer Science and Engineering, University Institute of Engineering and Technology,  
Maharshi Dayanand University, Rohtak, Haryana, India

<sup>1</sup>pd9493@gmail.com

<sup>2</sup>rainu\_nandal@yahoo.com

<sup>3</sup>sehrawat\_harkesh@yahoo.com

**Abstract**— In today's world competition is increasing day by day. In field of higher education as competition is increasing so the student self-harm rate is increasing. The reason for this is because students are not able to cope with studies and under pressure they do self-harm. Data mining is a technique which can be used to decrease this self-harm rate. In this paper we want to explain that using data mining we can predict result of students beforehand by using previous year result or any other factors in early stages of any course. This technique is called Prediction Analysis and this an example of use of data mining in the field of education. It can also be called as educational data mining. Any Data mining can be used for this prediction analysis and we will be using WEKA tool of data mining to predict result.

**Keywords**- Data mining, prediction analysis, WEKA, JAVA, tool of data mining, Pattern Generation, Train Set, Test Set.

## I. INTRODUCTION

### A. Data Mining

Data Mining is a part of computer science which is used to implement methods like machine learning, neural network and artificial intelligence. Data mining is the technique or method of extracting patterns or information from the already present information.

Data Mining came into view around 1990's. It can be used to implement classification, regression, clustering and implementing association rules.



Fig. 1. Data Mining [13].

### B. Prediction Analysis

Trying to find the result or answer of any question of present or future on the basis of past information is known as prediction.

In the field of education by using data mining we will try to predict result of students of on the basis of performance in previous year exam. On the basis of previous year result a pattern will be made or arranged using which result of current year will be predicted.

The basis of their performance in previous year exam can be taken as their class behavior, exam result, class attendance, extra-curricular activities or behavior with other student and teachers.

Using all these attributes or properties a pattern will be arranged or information will be extracted using data mining on previous result by training that information under any technique of data mining and will be stored. Later on that stored pattern information of this current result set will be tested and the result will be predicted.

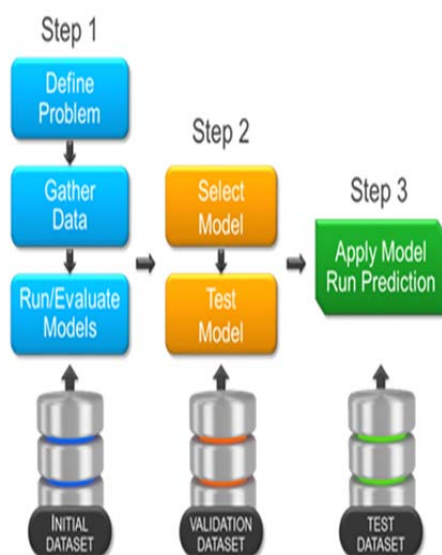


Fig. 2. Prediction Analysis using Data Mining [14]

## II. LITERATURE REVIEW

Data mining is used to filter relevant information from all the data present on web like hypertext and all other resources of web etc. Data mining has many techniques like classification, clustering, artificial neural network etc. under which many algorithms are used as k-means, decision tree, ZeroR and mlp training method as well as k-fold and association mining rules.

L. Romdhae, N. Fadhel and B. Ayeb [1] in 2010 showed that to predict student success rate used data mining techniques based on both regression and classification can be applied. The authors first showed prediction result using regression analysis and Bayes method. Later on he also showed the use of decision tree under j48 method of classification to predict result and they also studied about how clustering technique can be used to predict result.

K. Umamaheswari and S. Niraimathi [2] August 2013 used data mining technique to predict the shortlist students eligible for interview after prediction they arranged students acc. To their merit obtained using predictor they used various data mining technique for this work some are classification and clustering technique data mining. They used data mining technique to keep an eye on the performance of students acc. To them classification technique of data mining provide more accurate result than clustering technique but also say that clustering is an important technique.

S. Aher and L.M.R.J. Lobo [3] in 2011 used classification and clustering technique of data mining to predict student result. The author used ZeroR algorithm of classification technique to predict result and also predicted result using DBSCAN algorithm of clustering of data mining. They used WEKA to predict student result under these both techniques.

Sumam Sebastian [4] in June 2015 evaluated student result on the basis of artificial neural network under data mining. He also used weka tool of data mining to predict result. He used real time set of marks of 300 students to predict result. He used k-fold and association rule mining technique to predict result. Acc. to him k-fold technique provide better result than association rule mining. He conclude it by showing the comparison of both

techniques. In association rule mining he used mlp training method and in k-fold he used 10-fold technique to predict result.

Milos Ilic and Petar Spalevic [5] in march 2016 used classification to predict result rate of students .They predicted result by using weka tool of data mining under classification they used j48 and ZeroR algorithm to predict result. They also showed the use of IBK algorithm of classification. They used two different training sets and on their basis generated two different model later on that two different model they used two different test sets. they showed the comparison of different algorithms ontwo different data sets.and concluded that accuracy of prediction deponce upon the accuracy of training data sets .acc. to them as accurate the train set will be as accurate the result predicted will be by the model generated using training set.

Amirah Mohamed Shahiria, Wahidah Husaina and Nur'aini Abdul Rashida [6] showed a case study how data mining algorithms can be used to predict student result. They reviewed various research papers from 2002 to 2015 and explained their basis which techniques are mostly used to predict result. Acc. to them classification and artificial neural network are two basic techniques used to predict result. They explained neural network (98%) gives highest accuracy in prediction which is followed by decision tree under classification followed by 91% accuracy. Then are k-nearest neighbor with 85% and naïve Bayes with 75% accuracy which is lowest.

Parneet Kaura ,Manpreet Singhb and Gurpreet Singh Josanc [7] in 2015 predicted student result of 164 high school students using data mining techniques as well as slow learning students present in that group of students. They used weka tool to predict result and to show comparison of various techniques to predict result. Acc. to them multilayer perceptron is best algorithm for prediction with maximum accuracy of 85% in their training set. They also showed the comparison some other five data mining techniques including mlp then also multilayer perceptron has given best result as compared to others.

R. R. Kabra and R. S. Bichkar [8] December 2011 predicted the performance of engineering students. They used decision tree algorithm technique to predict student performance.they used the data of around 340 students of engineering field to prdict their performance in their first year exam. They generated pattern usin decision algorithm to create a training model for prediction.accuracy of model generated was 60% .they used confusion matrix to check whether how many student will pass or how many will fail.

### III. METHODOLOGY USED

**WEKA** :- Waikato Environment for Knowledge Analysis

WEKA is data mining tool used for data mining .Data mining can also be defined as process generating new information on the basis of past or previous information. WEKA is a tool based on JAVA language.

JAVA is an object oriented and platform independent programming language developed by sun microsystems. Both JAVA and WEKA has GNU (General Public License).

It was developed by university of Waikato, New Zealand.as it is based on java so it is platform independent tool which has many types of machine learning tool inbuilt in it like classification, clustering and many other data mining algorithms.



Fig. 4. a snap shot of WEKA Tool[17]

#### IV. GENERAL APPROACH

Generally in any study course student take admission they try to cope with all the competition by themselves. They give exam where either they pass or fail. Due to this pressure of getting good result they get buried under studies and when they don't get desired result sometime this can lead to self-harm .

As Now a day's student self-harm rate is increasing day by day rapidly. This is due to increasing competition between students and increasing complexity of learning.

But institutes can help student and can also decrease this self-harm rate and remove study pressure from student by the use of technology i.e. Data Mining .Data mining has many uses in many fields and is a part computer science world whose use is increasing day by day .This use of data mining in the field of education is known as Educational Data Mining commonly known as EDM.

#### V. PROPOSED WORK

By using Data mining in the field of education i.e. Educational Data Mining we can predict result of students beforehand and try to help those students who have low marks according to predictor. This can decrease the ratio of self-harm in students due to low marks and study pressure up to some extent which can be seen as a social help.

In this scheme (as shown in fig 5.1) a training set on the basis of past result of student will be created The basis to create that past result can be anything either their marks or their behavior in class and with fellow students. Using that set a pattern will be generated and that pattern can be used to find the result of students at early stages of any course.

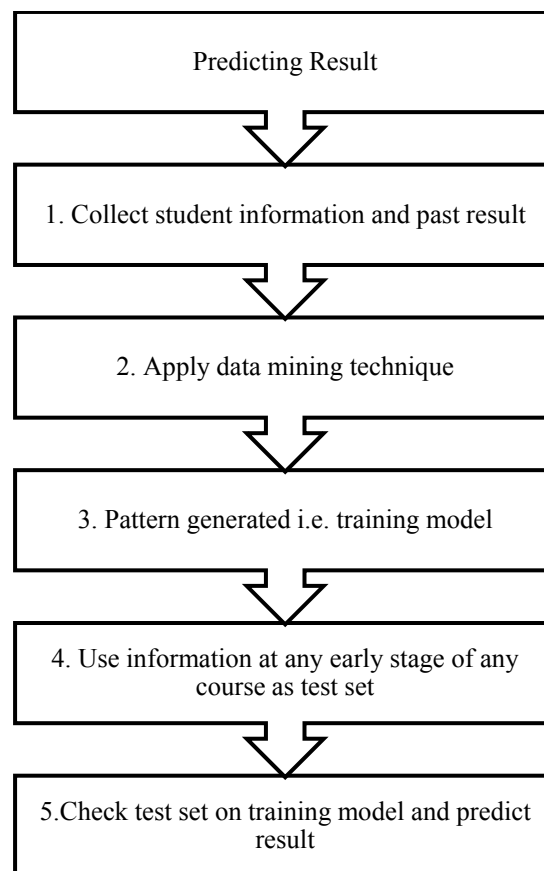


Fig. 4. Proposed Work using Data Mining to predict result[19]

## VI. CONCLUSION

We are trying to predict result of students at early stages only of any course on the basis of previous result using prediction analysis of Data mining. For prediction we are using WEKA tool which is free and is machine learning tool.

In WEKA tool we will as input give the past result set on the basis of which a model will be created which provide a particular patter of student result needed to pass exam. The past result set can be made on the basis of any attribute or property of students either their marks or behavior in class. Later on to predict result at any stage of course test set based on same attribute will be made and tested on the model created using WEKA .By doing this institute can help those students before final exam whose result has been predicted fail or low marks.

## REFERENCES

- [1] L. Romdhae, N. Fadhel, B. Ayeb, "An efficient approach for building customer profiles from business data", Expert System with Applications, vol. 37, 2010, pp. 1573-1585.
- [2] K. Umamaheswari, S. Niraimathi "A study on student data analysis using data mining techniques", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, Issue 8, August 2013, pp. 117-120.
- [3] S. Aher, L.M.R.J. Lobo, *Data mining in educational system using weka*, *International Conference on Emerging Technology Trends, 2011*, pp. 20-25
- [4] International Journal of Computer Applications (0975 – 8887) Volume 119 – No.23, June 2015 36 *Evaluating Students Performance by Artificial Neural Network using WEKA* Sumam Sebastian M-Tech Computer and Information Science College of Engineering Poonjar Jiby J Puthiyidam Assistant Professor Dept. of Computer Science and Engineering College of Engineering Poonjar
- [5] INFOTEH-JAHORINA Vol. 15, March 2016. 684 *Students' success prediction using Weka tool* Milos Ilic, Petar Spalevic Electrical and Computing Engineering University of Pristina, Faculty of Technical Science Kosovska Mitrovica, Serbia, Mladen Veinovic, Wejdan Saed Alatresh Singidunum University Belgrade, Serbia
- [6] The Third Information Systems International Conference *A Review on Predicting Student's Performance using Data Mining Techniques* Amirah Mohamed Shahiria,\* , Wahidah Husaina , Nur'aini Abdul Rashida , aSchool of Computer Sciences Universiti Sains Malaysia 11800 USM, Penang, Malaysia
- [7] *Performance Prediction of Engineering Students using Decision Trees* R. R. Kabra S.G.R. Education Foundation's College of Engineering and Management, Ahmednagar, India. R. S. Bichkar G. H. Raisonni College of Engineering and Management, Pune,
- [8] Z. J. Kovacic, "Early prediction of student success: Mining student enrollment data", Proceedings of Informing Science & IT Education Conference (InSITE) 2010.
- [9] M. Ramaswami and R. Bhaskaran, "A CHAID based performance prediction model in educational data mining", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 1, No. 1, January 2010.
- [10] N. Thai Nghe, P. Janecek, and P. Haddawy, "A Comparative Analysis of Techniques for Predicting Academic Performance", 37th ASEE/IEEE Frontiers in Education Conference, October 2007.
- [11] O. Oyelade, O. Oladipupo, I. Obagbuwa, "Application of k-Means clustering algorithm for prediction of students' academic performance" (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, num. 1, 2010, pp. 292-295.
- [12] A. Kumar, G. Uma, "Improving academic performance of students by applying data mining techniques", European Journal of Scientific Research, no. 4, 2009, pp. 526-534
- [13] [https://www.sas.com/en\\_us/insights/analytics/data-mining/\\_jcr\\_content/socialShareImage.img.png](https://www.sas.com/en_us/insights/analytics/data-mining/_jcr_content/socialShareImage.img.png)
- [14] Fig 1 taken from: <http://predictive-5a88.kxcdn.com/wp-content/uploads/2013/12/Predictive-Modeling.jpg>
- [15] Fig 2 taken from: V. P. Bresfelean, "Analysis and Predictions on Students' Behavior Using Decision Trees in Weka Environment", Proceedings of the ITI 2007 29th Int. Conf. on Information Technology Interfaces, June 25-28, 2007.
- [16] P. Cortez, and A. Silva, "Using Data Mining To Predict Secondary School Student Performance", In EUROIS, A. Brito and J. Teixeira (Eds.), 2008, pp.5-12.
- [17] *India Classification and prediction based data mining algorithms to predict slow learners in education sector* Parneet Kaura ,Manpreet Singhb ,Gurpreet Singh Josanc
- [18] Fig 3 taken from: <http://www.cs.waikato.ac.nz/ml/>
- [19] *Predicting Students' Performance using Modified ID3 Algorithm* Ramanathan L1 , Saksham Dhanda 2, Suresh Kumar D 3
- [20] Fig 4 taken from: Made by author