IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 01-04

Academic Assesment & Performance Prediction Using Data Mining

Shivam Tiwari[1], Barkha Bhardwaj[2], Namrata Dhanda[3], Bramah Hazela[4]
[1234]*Department of Computer Science & Engineering, Amity School of Engineering & Technology, Amity University, Lucknow, Uttar Pradesh, India*

Abstract: *These days, the most critical test for every association like an instructive foundation, a worldwide enterprise, and so forth is to support their notoriety by giving quality administrations. Numerous associations are directing persuasive, social occasions to improve their representatives or understudies execution. The essential test to deal with is to limit the disappointment proportion of a recipient and give quality support of every single recipient. With the goal that establishment needs an expectation framework which is utilized to foresee recipient potential execution based on their own execution. By which organization can decide the fitting arrangements for each recipient execution expectation. Based on the forecast, foundation sorts their recipient and makes suitable arrangements for every single order. In this report, I am talking about a type of forecast techniques like a neural system, an information mining calculations named as Linear Regression, support vector machine, Naive Bayes, nearest neighbour, decision tree, and so on and furthermore examined the parallel significance of a calculation in recipient Performance by Prediction System. Here I will talk about the past research identified with this point. I have referenced in this report the different outcomes by utilizing an alternate calculation. Dissecting the outcomes and give a future part of the calculation. In the interest of the past investigation, the talk is given about the expectation framework and different systems to unpack the recipient execution.*

Keywords: *Quality Administration, linear regression, information mining, prediction system, Support vector machine, naïve Bayes.*

I. Introduction

Nowadays, education is a challenge. There are so many areas which affect the student's academic performance directly as well as indirectly. To overcome these challenge institutes are organizing so many activities on their campus as well as out of the campus. This is the bigger challenge for each and every institute to improve their student's academic performance and give quality education to all. Because the performance of their student directly affect the reputation of the institute so that many types of research are going on educational data. The data mining technique having all information related to education called "Educational data mining [EDM]" [6].

With the help of education data mining [EDM] institute analyse the student's academic data and predict the student's academic performance by using a data mining algorithm. It is a curious need for the education system to predict the result of the student and making policies on behalf of them. This technique will always help the system to maintain its reputation as well as by predicting the result institutes can make good policies for their students. By using that technique overall result will be improved as well as the quality education would be delivered to all the students as per the categorization is done. Here Discussion some of the various researches which have been done on this problem. And also analyse the conclusion.

On the basis of prediction, the institute can categorize their students and make appropriate policies for each categorization. There are many more researches are done on the basis of some attributes like Student behaviour, Demographics, Information, Socioeconomic, and their physical and mental strength [7].
Here in Table 1 [1], given a table to explain some measures on behalf of their attributes [1].

ATTRIBUTES	MEASUREMENTS
Student Behaviour	Access to each feature extracted from e-learning logged data.
Demographics	Hobbies, family size.
Information	Age, Name, Gender, Address.
Socio Economic	Background of student
Psychometric	Attitude, Ability

Table 1. Attribute measure table

There are various data mining techniques such as Regression, Classification, Clustering, Decision tree, Naive Bayes, Bayesian Network, and Neural Network for predicting student's academic performance in their use.

In this report, I am discussing a data mining algorithm named as Linear Regression and application of Linear Regression algorithm in Student's Academic Performance Prediction System. Here I will discuss the previous research related to this topic. I have mentioned in this report the various results by using a different algorithm. Analysing the results and give a future aspect of the algorithm. On behalf of the previous study, the discussion is given about this topic.

II. Literature Review And Discussion

In the series of various study related to student's academic performance prediction, there are many more conclusions /results are founded for different data mining algorithm. Various results are founded in different research. Here the discussion is present on behalf of various research work have been done in the past. Here I am discussing the various methods result and find out the best conclusion. Some of the research conclusion is given below on behalf of them discussing Linear Regression and comparison which is already done is to be discussed.

Now, talking about various researches which gave a better conclusion for the student's academic performance prediction. The various results are founded in various researches. Researches explanation and their results are shown individually below [8].

In[1], There is various algorithm are used in this research such as Decision tree, Neural Network, Naive Bayes, K-Nearest Neighbour, and Support Vector Machine. In this paper, a study was done on Malaysian educational data. Here the problem is that the educational data is not arranged in a proper manner like institute was not able to find out the real factors that affect the student's performance directly as well as indirectly so there is much need of a student academic performance prediction system. The main reason for the performance ratio was gradually decreasing in nature. Because of that, the reputation of the institutes was falling down so the research was done that is kept in mind for the problem. As I have already said that in this research various algorithms are used and also calculated the error ratio between actual and result in data. And then we come to know the best algorithm for this situation. Here various attributes were taken like internal assessments, Psychometric factors, External assessment, CGPA, Student Demographic, High school background, Scholarship, Social Network interaction, extracurricular activities, and Soft skills.

In result, the Neural Network algorithm was founded most successful with 98% accuracy. The results of other algorithms are given as -

In figure 1, Decision Tree with 91% accuracy, followed by Support Vector machine with 83% accuracy. Here K-Nearest Neighbour was founded with the same accuracy as Support Vector machine [SVM]. Naive Bayes was found with the least accuracy which is equal to 76%.

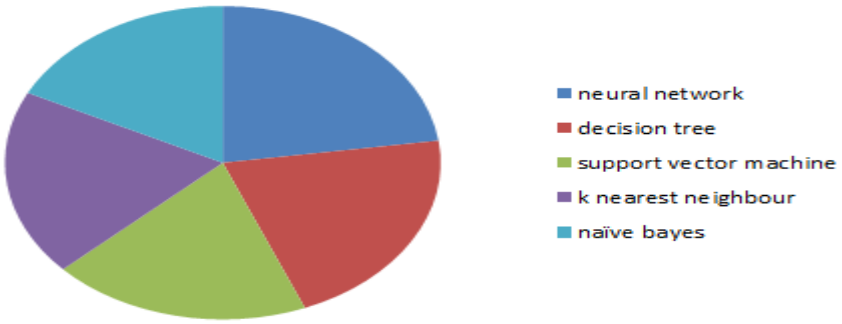

Figure 1. Data science algorithms

So here we can say Neural Network gave the best accuracy result thus the conclusion is that Neural Network may give the best result for prediction student's academic performance.

In [2] Decision Tree is used to predict the student's academic performance. Here 1548 records of students are used. Data set is in such format 292 "excellent", 536 are in "very good" category, 477 are in "good", 188 are in "acceptable" and rest are in "fail" category. After getting the result of entropy "Midterm" is selected as the root node of the tree. In this research Decision tree (ID3) method is used. In this research taken as attributes are Midterm Mark, Lab Test Grade, Student Participate, Homework, Performance, Department and

Final Grade. By using these attributes take ID3 method in work and got accuracy results. Here there are 8 case study is taken in picture. Here no comparison is given for accurate results. Only prediction point is calculated.
In [3] Neural Network, Decision tree, Binary classification, and Naive Bayes classification are used to predict the academic performance of the student. Here sex, age, and past years' results are taken as attributes to calculate the prediction.
In result Neural Network with 94% accuracy and binary classification with 72% accuracy.

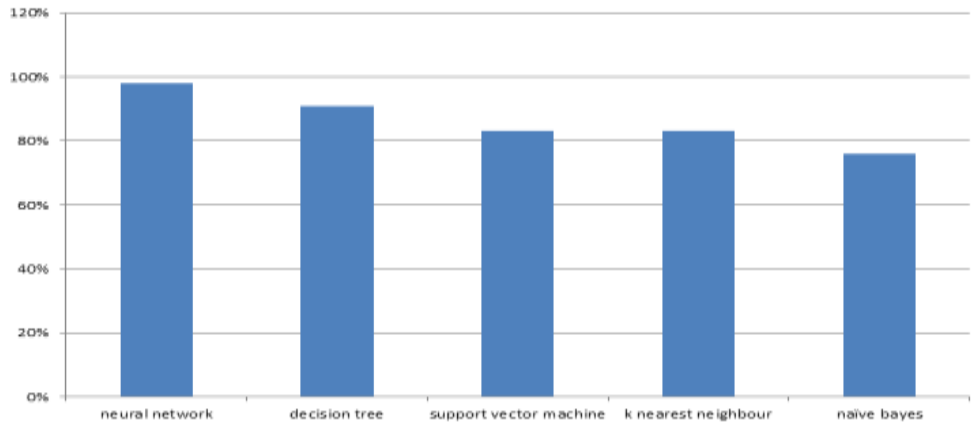

Figure 2. Effectiveness of data science Methodologies

In [4], Here using some traditional classification named as (1) Bagging ensemble method, (2) boosting ensemble method, (3) Ensemble filtering. Here all of above methods give similar results. Here the most important result is the two factors student's behaviour and absent days are the key to improve student's academic performance. In this paper the comparison is also taken in picture between various algorithm and their results. Here taken the evolution measures Accuracy, Precision and Recall. As I have pointed in beginning considering the attribute (1) Results of all features like previous percentage got by the students. (2) Behaviour (3) number of absents. Here writer considered that or focused the main 2 attributes Behaviour, Absents. According to this paper if no. of absents is maximum the performance of the students will be decreasing day by day. Here comparison between Naive Bayes and k-nearest neighbour in this comparison Naive Bayes gives 74.2 accuracy by taking "result of all features like previous percentage" taking an attribute whereas for this attribute k-nearest neighbour gives 76.1 accuracy. Other comparison is also made here that is Decision tree and Artificial intelligence. In this comparison the result is for the attribute "Behaviour" is 55.8 for Decision tree and 55.3 for Artificial intelligence. In this research ensemble filtering technique is used and that gives 84.3 accuracy which is better than the result taken by the Artificial Neural Network that is 78.6.

The overall conclusion of this paper is the student's performance will always depend upon mainly 2 factors, (1) behaviour and (2) number of absents by the students [9]. And here focused to improve the accuracy of the results obtained from the various algorithms by using some filters.

III. Motivation

As per the above study, it is cleared that most of the time Neural Network having the highest accuracy. This prediction system having highly needed requirement for each and every institute, student, etc. Applying an algorithm and then analyse the result, there is the primary need of data set. After those applying different algorithms, we can start analysing different results and find out the best algorithm for the dataset.

Here now I am getting various types of research-based knowledge like how the problem analysis is to done and how we get the best accuracy results. As per the above studies, I have learned so many research strategies. Find out the results is not enough; it is very important to compare the result with the previous study or research done. With the help of this study, I started self-study to develop a new system which can be used for predicting the student's academic performance. In the primary stage, I started to learn Algorithm that is "Linear Regression".

In my study, I have learned Linear Regression can also be used in predicting the student's academic performance. Here like an equation that is:

y = Mx+C;[Here "y'" is dependent to "x";we can calculate each "y" for every "x".]

The motivation is to find the solution in its own way. I have learned so many things from this above study. Here as per "Predicting the Student's academic performance" is taken as a subject, then we got to know that There is a much need of that kind of automated system that can be able to predict the performance of the

student's. Now a day there is the basic need of each and every college, school or student's level, a system like that. By finding out this type of better system Institute will be helped to make better policies for their students. The performance of the student always directly proposal to the reputation of the institute. In this commercial world, there is a race between the institute as well as students to come out first. So, by taking in mind that scenario, now every institute will be going to have a big need of that system. This system will always help students on how to increase the performance of their academic session.

As per the above study I also got the various aspects to predict the student's academic performance like we can take: Library record in hour, attendance in percentage, address in form of day scholar or hosteller and previous year's results in percentage, these can be taken as the attributes to help out the prediction system. Taken motivation by the study I am now able to do some of my research in this area. I pointed above the attribute list.

IV. Conclusion

In this examination, we have taken in the equivocal conduct of the recipient execution by breaking down their inclination through a forecast framework. Here, we spoke much about the different information science calculations parallel with their adequacy. some of may be the most appropriate for particular employment like for foreseeing recipient execution according to the subtleties we talked about above, the neural system gives the greatest effectiveness. That will assist the association with making a better strategy for the recipient. It will improve the method for administration conveyance as far as their individual areas. Then again, there are some other information science calculations that can likewise be useful while guessing the recipient inputs. In general, we can say that it will improve the recipient potential execution by examining past sources of info. The recipient execution is the way to progress full positive input of the other recipient of the specific association. In this manner, we can say in the event that the recipient execution is estimated previously, at that point association can keep up their notoriety by making a few proper arrangements.

According to the future extent of this investigation, we can discover the new better instructive information mining calculation and parallel attempt to make a progressively appropriate system for anticipating the recipient potential execution with the best exactness in the outcomes. Here it will dependably rouse me for future research here. Conceivably, information mining is dependably a superior stage for research in the field of information science. This investigation will propel me for further looks into that I will precede in my next.

References

[1]. The Third Information Systems International Conference, A Review on Predicting Student's Performance using Data Mining Techniques, Amirah Mohamed Shahiri ,Wahidah Husain, Nur'aini Abdul Rashid, a school of Computer Science University Sains Malaysia 11800 USM,Penang, Malaysia.

[2]. Data Mining: A prediction for Student's Performance Using Classification Method, Abeer Badr El Din Ahmed, Ibrahim Sayed Elaraby.

[3]. Ahmed, A. B. E, Ibrahim S. E. "Data Mining: A prediction for Student's Performance Using Classification Method." World Journal of Computer Application and Technology Volume 2(2) (2014).

[4]. 2nd International Conference on Electrical & Electronic Engineering (ICEEE), 27-28 December 2017,RUET, Rajshahi, Bangladesh, "Predict Student's Academic Performance and Evaluate the Impact of Different Attributes on the Performance Using Data Mining Techniques", Md. Hasibur Rahman, Md. Rabiul Islam.

[5]. Predicting Students Performance in Final Examination using Linear Regression and Multilayer Perceptron, Febrianti Widyahastuti and Viany Utami Tjhin ,978-1-5090-4688-1/17/$31.00 copyright 2017 IEEE.

[6]. Asif R, Merceron A, Pathan MK. Predicting student academic performance at degree level: a case study. International Journal of Intelligent Systems and Applications. 2014;7：49.

[7]. J. MM, S. H. Prediction of students' academic performance: Adapt a methodology of predictive modeling for a small sample size. 2014 IEEE Frontiers in Education Conference (FIE) Proceedings2014. p. 1-3.

[8]. Goga M, Kuyoro S, Goga N. A recommender for improving the student academic performance. ProcediaSocial and Behavioral Sciences. 2015;180：1481-8.

[9]. Altaher A, BaRukab O. Prediction of Student's Academic Performance Based on Adaptive Neuro-Fuzzy Inference. International Journal of Computer Science and Network Security (IJCSNS). 2017;17：165.