



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Analysis and Prediction of Student's Academic Performance Based on Educational Data Mining

Prof. Vinay Nalawade, Jamdade Pragati Ravindra, Lawand Akshata Mohan, Narkhedkar Ajay

Kailas, Raut Madhura Sunil

Department of Computer Engineering, S. B. Patil College of Engineering, Indapur, Maharashtra, India

ABSTRACT: Education plays a pivotal role in producing qualified human power that accelerates economic development and solves the real problems of a community. Students are also expected to spend much of their time on their education and need to graduate with good academic results. However, the trend of graduating students is not proportional to the trend of enrolled students and an increasing number of students commit readmission, suggesting that they did not perform well in their academics. Thus, the study aimed to identify the determinants of academic performance among university students.

KEYWORDS--: Data mining, Performance analysis, Machine Learning.

I. INTRODUCTION

Education is regarded as a promoter of human development and seen by many to be in the center of any society's life and concern. It is a social artifact embodying aspiration about the welfare and development of the society it deems to serve. To Botswana, education is expected to contribute towards the social, cultural, political and economic welfare and development of citizens (RNPE, 1994). According to Botswana educational goals, children who complete secondary education are expected to have acquired lifelong skills and be competitive in the global village when it comes to their employability (RNPE, 1994). This therefore, calls for students to excel academically or hopefully perform to the satisfaction of the nation. In Botswana basic education is free to all children. The government is the sole sponsor for the education of the students from primary to secondary school levels. To some they receive government

sponsorship at tertiary level. Since government committed itself to provide basic education for all, the Ministry of Education and Skill Development has been receiving a lion's share in both recurrent and development budget. Since 2007/2008 budget the Ministry of Education and Skill Development has been allocated over P5 billion of the recurrent budget. The 2013/2014 budget the Ministry of Education and Skill development is allocated P7.93 billion or 22.98 per cent of the ministerial recurrent budget. Considering government hefty investment in education, its output about the quality of students has not been commensurate with the expenditure. The students' academic performance has been declining at an alarming rate since 2010. Table 1 below shows yearly students' academic performance for 2010, 2011 and 2012, which indicate a serious decline. This has caused a concern for both the government and the public.

II. PREVIOUS WORK

1. Low Accuracy Detected
2. Limited dataset used

III. PROPOSED SYSTEM

Student's performance in the educational process can literally be defined as something that is obtained from changes in the behavior of students based on their experiences, besides that learning outcomes are also a realization of the potential or capacity possessed by students [1]. These learning outcomes from students can be seen from their behavior, both behavior in the form of

understanding knowledge, thinking skills, or motor skills [2]; an outcome of the process of changing student behavior after attending lessons [3]. The concrete form of student's performance can be seen from their understanding of the knowledge being studied, their expertise in processing information and making decisions based on certain thoughts or motor skills [4]. Based on those understandings, student's performance can be observed and measured in the realm of students' knowledge, attitudes, and skills after following a series of lessons. Student's performance depends on the teaching and learning process they go through, so that learning outcomes can be used as considerations in improving the quality of the learning process. On the other hand, a similar terminology is learning achievement which is a measure of student achievement after participating in learning activities in the form of an assessment scale (either letters, numbers or certain symbols [5]. Giving a weight or rating for student's achievement in learning requires the preparation of suitable assessment indicators, and needs to be ensured of the validity and the reliability [6]. This value can then be used as a description of student's performance in a certain period of time [7]. Student's performance is obtained after passing various measurements (in various forms of assessment) after students carry out several learning processes.

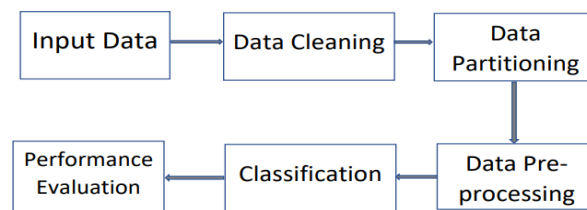


Fig. System Architecture

IV. MATHEMATICAL MODULE

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. The major purpose of the study was to investigate the teachers' and students' views, ideas and experiences concerning factors which contribute towards students' low academic performance. It also explored strategies needed to alleviate the problem if any.

$\beta = (\beta_0, \beta_n)$ are the estimated regression

coefficient. Random forest (RF): Random Forest uses bagging method to generate trees in which its prediction is more accurate than that of any individual tree [19]. Random forest was also used to avoid over fitting on the training set and limiting errors due to bias hence yield accurate and useful results. RF can handle outliers and noise in the data and gains high classification accuracy. RF generate numerous decision trees in the training phase and output class label.

Data Field	Description
Gender	The student's gender.
Nationality	The student's nationality
Place of Birth	The student's country of birth.
Stage ID	Educational level student belongs to (Elementary, Middle, or High School)
Grade ID	The grade year of the student.
Section ID	The classroom the student is in.
Topic	The topic of the course.
Semester	The semester of the school year. (F for Fall, S for Spring)
Relation	The parent responsible for student.
raised hands	How many times the student raises his/her hand on classroom
Visited Resources	How many times the student visits a course content
Announcements View	How many times the student checks the new announcements
Discussion	How many times the student participate on discussion groups
Parent Answering Survey	Parent answered the surveys which are provided from school or not
Parent School Satisfaction	Whether or not the parents were satisfied. "Good" or "Bad". Oddly this was not null for parents who did not answer the survey. It is unclear how this value was filled in.
Student Absence Days	Whether or not a student was absent for more than 7 days
Class	Our classification field. 'L' is for students who got a failing percentage (Less than 69%), 'M' for students who got a low passing grade (Between 70% and 89%), and 'H' for students who achieved high marks in their course (90% to 100%)

Table 1: his table depicts the student related variables and their description

Support Vector Machine	70.73
Random Forest	73.65
Naïve Bayes	67.90

Table 2: his table depicts the various visualization algorithm accuracy

V. CONCLUSION

The system focuses on the student academic growth analysis using machine learning techniques. For analysis Decision tree and random forest classifier This process can help the instructor to decide easily about performance of the students and schedule better method for improving their academics. In future additional features are added to our dataset to acquire better accuracy.

VI. ACKNOWLEDGMENT

We would also like to give special thanks to Dr.Shirkande S.T. (Principal at SBPCOE,Indapur) and Dr.Gavali A.B. for sharing their pearls of wisdom with us to do the research regarding the “Analysis and prediction of student’s Academic performance Based on educational Data mining”.

REFERENCES

1. N.V.Krishna Rao, Dr.N.Mangathayaru, Dr.M.Sreenivasa Rao,” Evolution and Prediction of Radical MultiDimensional E-Learning System with Cluster based Data Mining Techniques”, International Conference on Trends in Electronics and Informatics,2017, PP.701-707.
2. Pushpa S.K, Manjunath T.N, “Class result prediction using machine learning”, InternationalConference On Smart Technologyfor Smart Nation,2017,pp1208- 1212. Micheal Bowles, Machine Learning in Python: Essential Techniques for Predictive Analysis. John Wiley Sons, Inc. 2015.
3. Trishul Chilimbi, Yutaka Suzue, Johnson Apacible, and Karthik Kalyanaraman. Project Adam:Building an efficient and scalable deep learning training system. In 11th USENIX Symposium on Operating Systems Design and Implementation (OSDI 14), pages 571–582, 2014.
4. Jack Clark. Google turning its lucrative web search over to AI machines, 2015. www.bloomberg.com/news/articles/2015-1026/googleturning-its-lucrative-websearch-over-to-aimachines.
5. J. Xu, K. H. Moon, and M. Van Der Schaar, “A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs,” IEEE J. Sel. Top. Signal Process., vol. 11, no. 5, pp. 742–753, 2017.
6. K. P. Shaleena and S. Paul, “Data mining techniques for predicting student performance,” in ICETECH 2015 - 2015 IEEE International Conference on Engineering and Technology, 2015, no. March, pp. 0–2.
7. M. Shahiri, W. Husain, and N. A. Rashid, “A Review on Predicting Student’s Performance Using Data Mining Techniques,” in Procedia Computer Science, 2015.
8. Y. Meier, J. Xu, O. Atan, and M. Van Der Schaar, “Predicting grades,” IEEE Trans. Signal Process. vol. 64, no. 4, pp. 959–972, 2016.
9. P. Guleria, N. Thakur, and M. Sood, “Predicting student performance using decision tree classifiers and information gain,” Proc. 2014 3rd Int. Conf. Parallel, Distrib. Grid Comput. PDGC 2014, pp. 126–129, 2015.
10. P. M. Arsad, N. Buniyamin, and J. L. A. Manan, “A neural network students’ performance prediction model (NNSPPM),” 2013 IEEE Int. Conf. Smart Instrumentation, Meas. Appl. ICSIMA 2013, no. July 2006, pp. 26–27, 2013.



Impact Factor: 8.379



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details