iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.43819

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

Implementation of Predicting Student Performance Using Linear Regression Algorithm

Ms. Kavita Ashok Chauthe[1], Dr. S. S. Dhande[2]
[1]*PG Scholar,* [2]*Professor, Computer Science & Engineering*
Sipna College of Engineering and Technology, Amravati, Maharashtra, INDIA

Abstract: ***This data approach student achievement in secondary education of two Portuguese schools. The data attributes include student grades, demographic, social and school-related features) and it was collected by using school reports and questionnaires. Two datasets are provided regarding the performance in two distinct subjects: Mathematics (mat) and Portuguese language (por). In [Cortez and Silva, 2008], the two data sets were modeled under binary/five-level classification and regression tasks. Important note: the target attribute G3 has a strong correlation with attributes G2 and G1. This occurs because G3 is the final year grade (issued at the 3rd period), while G1 and G2 correspond to the 1st and 2nd period grades. It is more difficult to predict G3 without G2 and G1, but such prediction is much more useful (see paper source for more details).***
Keywords: Decision Tree Algorithm, Naïve Bayes Algorithm, Support Vector Machine

I. INTRODUCTION

Since universities are prestigious places of higher education, students' retention in these universities is a matter of high concern. It has been found that most of the students' drop-out from the universities during their first year is due to lack of proper support in undergraduate courses. Due to this reason, the first year of the undergraduate student is referred as a "make or break" year. Without getting any support on the course domain and its complexity, it may demotivate a student and can be the cause to withdraw the course. There is a great need to develop an appropriate solution to assist student's retention at higher education institutions. Early grade prediction is one of the solutions that have a tendency to monitor students' progress in the degree courses at the University and will lead to improving the students' learning process based on predicted grades.

Using machine learning with Educational Data Mining can improve the learning process of students. Different models can be developed to predict students' grades in the enrolled courses, which provide valuable information to facilitate students' retention in those courses. This information can be used to early identify students at-risk based on which a system can 1 suggest the instructors to provide special attention to those students. This information can also help in predicting the students' grades in different courses to monitor their performance in a better way that can enhance the students' retention rate of the universities.

Using various packages such as cufflinks, seaborn & matplotlib to represent the data along with different attributes graphically or pictorially to analyses the dataset for predicting the Final Grade(G3).

II. LITERATURE SURVEY

[1]. Classification is one of the most researched questions in machine learning and data mining. A wide range of real problems have been stated as classification problems, for example credit scoring, bankruptcy prediction, medical diagnosis, pattern recognition, text categorization, software quality assessment, and many more. The use of evolutionary algorithms for training classifiers has been studied in the past few decades. Genetic programming (GP) is a flexible and powerful evolutionary technique with some features that can be very valuable and suitable for the evolution of classifiers. This paper surveys existing literature about the application of genetic programming to classification, to show the different ways in which this evolutionary algorithm can help in the construction of accurate and reliable classifiers.

[2]. This study presents a learning behavior diagnosis system to study students' learning status from learning portfolios. The proposed linking layer enables the proposed system to work on various e-learning platforms without reprogramming. Additionally, the use of a supervisory agent enables teachers and students to obtain their learning status or information provided by the proposed system in both Web and e-mail. Furthermore, the computer engineering curriculum operating systems was adopted to evaluate the proposed system. Evaluations of confidence between learning status and learning achievement yield positive experimental results.

[3]. Data mining techniques are applied in higher education more and more to give insights to educational and administrative problems in order to increase the managerial effectiveness. However, most of the educational mining research focuses on modeling student's performance. In this paper, data mining is utilized to analyze course evaluation questionnaires. Here, the most important

©IJRASET: All Rights are Reserved

variables that separate “satisfactory” and “not satisfactory” instructor performances based on students’ perception are found. Hopefully, these can help instructors to improve their performances. In addition, irrelevant variables that do not differentiate “satisfactory” and “not satisfactory” instructor performances are also listed. Different dimensions of course and instructor effectiveness are measured with course evaluation questionnaires in higher education institutions and these findings may be used to improve measurement instruments.

[4]. This paper represents the data mining techniques used for analysing pupil performance. Educational institutions contain an enormous amount of academic database containing student details. These student databases along with other attributes are taken into consideration like family background, family income, etc. It will help us by identifying promising students and by providing us a chance to pay heed and to refine those students who likely get low marks. For answer, we prepare a structure which will analyse the pupil’s performance from their last performances using concepts of Data Mining under Classification. Classification Algorithms like Decision Tree, Naïve Bayes and Support Vector Machine can help us for predicting student’s performance. This prediction helps parents and teachers to keep track of student’s performance and provide required counselling. These Analysis also help in providing scholarship and other required training to the student. We are actually trying to enhance student’s acquirement and success more effectively in a way using educational data mining techniques. It can bring the benefits & influence of novice, teachers and educational institutions. Experimental answers show that suggested procedure significantly outperforms prevailing procedure due to the misuse of family incomes and students’ personal data component sets. Results of this examination can act as policy improvement technique in higher education.

III. SYSTEM DIAGRAM

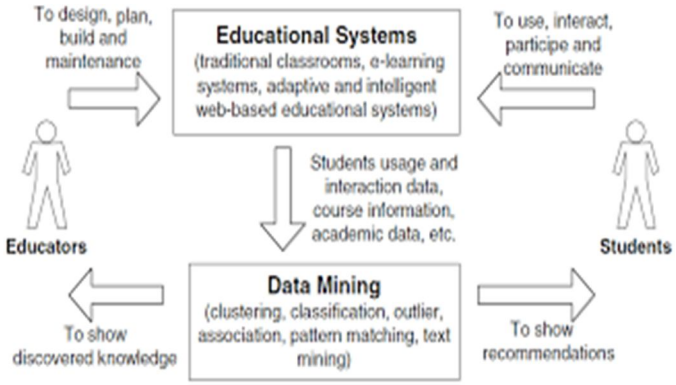


Fig : Predicting Student Performance using data mining techniques

IV. SYSTEM METHODOLOGY

Since universities are prestigious places of higher education, students’ retention in these universities is a matter of high concern. It has been found that most of the students’ drop-out from the universities during their first year is due to lack of proper support in undergraduate courses. Due to this reason, the first year of the undergraduate student is referred as a “make or break” year. Without getting any support on the course domain and its complexity, it may demotivate a student and can be the cause to withdraw the course.There is a great need to develop an appropriate solution to assist students retention at higher education institutions. Early grade prediction is one of the solutions that have a tendency to monitor students’ progress in the degree courses at the University and will lead to improving the students’ learning process based on predicted grades.

Using machine learning with Educational Data Mining can improve the learning process of students. Different models can be developed to predict students’ grades in the enrolled courses, which provide valuable information to facilitate students’ retention in those courses. This information can be used to early identify students at-risk based on which a system can 1 suggest the instructors to provide special attention to those students. This information can also help in predicting the students’ grades in different courses to monitor their performance in a better way that can enhance the students’ retention rate of the universities.

Using various packages such as cufflinks, seaborn & matplotlib to represent the data along with different attributes graphically or pictorially to analyse the dataset for predicting the Final Grade(G3).

V. SYSTEM RESULTS

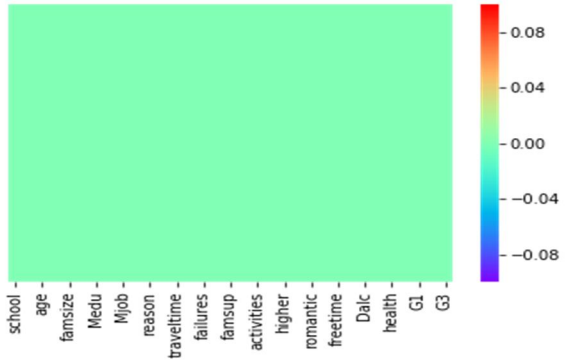


Fig 1. Pictorial representation of any null data present in the dataset.

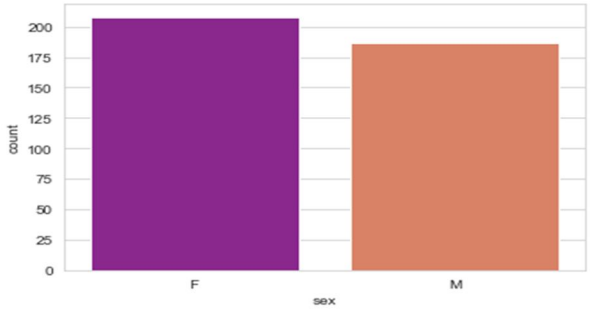


Fig 2. Count Plot for Student Sex Attribute

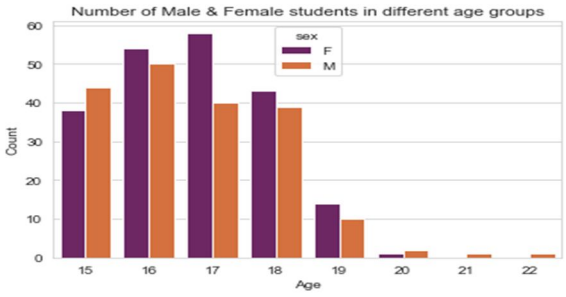


Fig 3. Count PLot for Male & Female students in different age groups.

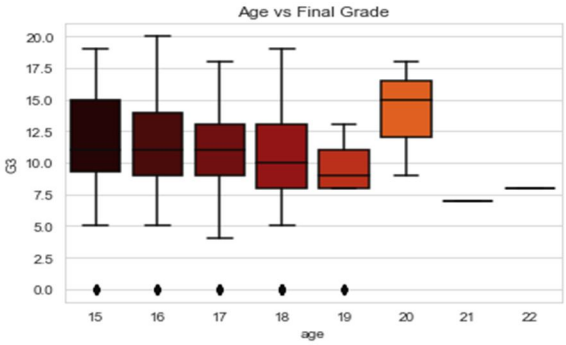


Fig 4. Count Plot for students from Urban & Rural Region.

VI. CONCLUSION

We concluded the proposed systems which monitor students' learning situations on a regular time schedule set by teachers. We have successfully completed the system that collects learning records of students and determines abnormal learning status of students. Each student uses the e-learning system, and their actions are recorded in a database to be analyzed. Data mining function modules, such as the Bayesian approach, cluster, decision tree, etc., will be imported into the proposed system according to teachers' need in the future.

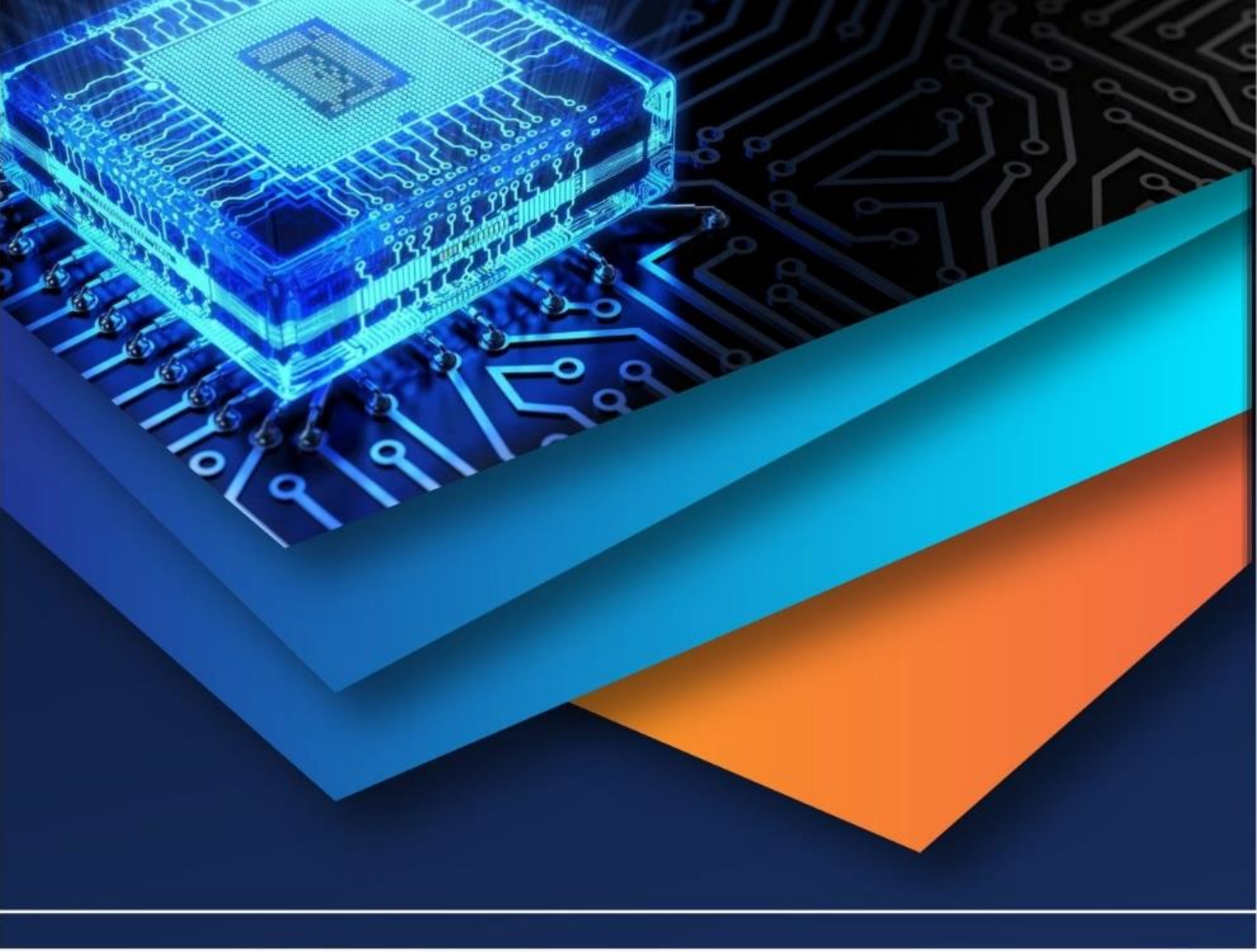
VII. ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my **Dr. S. S. Dhande** who has in the literal sense, guided and supervised me. I am indebted with a deep sense of gratitude for the constant inspiration and valuable guidance throughout the work.

REFERENCES

[1] Pedro G. Espejo, Sebasti´an Ventura, and Francisco Herrera "A Survey on the Application of Genetic Programming to Classification".
[2] Bin-Shyan Jong, Member, IEEE, Te-Yi Chan, and Yu-Lung Wu "Learning Log Explorer in E-Learning Diagnosis".
[3] Mustafa Agaoglu "Predicting Instructor Performance Using Data Mining Techniques in Higher Education"
[4] Predicting Instructor Performance Using Data Mining Techniques in Higher Education Astha Soni1, Vivek Kumar2 , Rajwant Kaur3 , D. Hemavathi4 Dept. of Information Technology, SRM Institute of Science & Technology
[5] Han, J. 2012. Micheline Kamber, & Jian Pei."Data Mining: Concepts and Techniques".
[6] Aziz, A. A., Ismail, N. H., & Ahmad, F. 2013." Mining Students' Academic Performance" inJournal of Theoretical & Applied Information Technology,53(3).
[7] Ahmad, F., Ismail, N. H., & Aziz, A. A. 2015. "The Prediction of Students' Academic Performance Using Classification Data Mining Techniques." inApplied Mathematical Sciences,9(129), 6415-6426.
[8] El-Halees, A. 2009. Mining Student's Data to Analyze ELearning Behavior: A Case Study.
[9] Tair, M. M. A., & El-Halees, A. M. 2012." Mining Educational Data to Improve Students' Performance: A Case Study" in International Journal of Information,2(2), 140-146.
[10] Garc´ıa, E. P. I., & Mora, P. M. 2011." Model Prediction of Academic Performance for First Year Students". InArtificial Intelligence (Micai), 2011 10th Mexican International Conference (pp. 169-174).
[11] Bhardwaj, B. K., & Pal, S.2012." Data Mining: A Prediction for Performance Improvement Using Classification".Arxiv Preprint Arxiv:1201.3418.
[12] Al-Radaideh, Q. A., Al-Shawakfa, E. M., & Al-Najjar, M. I. 2006." Mining Student Data Using Decision Trees." InInternational Arab Conference on Information Technology, Yarmouk University, Jordan.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)