Student Grade Analysis and Prediction for Developing Countries Using XAI

1st Shimul Mandal Tamo
Department of Computer Science and Engineering
BRAC University
66 Mohakhali, Dhaka 1212, Bangladesh
shimul.mondol.tamo@g.bracu.ac.bd

2nd Md. Fazle Rabbi Linkon
Department of Computer Science and Engineering
BRAC University
66 Mohakhali, Dhaka 1212, Bangladesh
fazle.rabbi.linkon@g.bracu.ac.bd

3rd Sharia Hoque
Department of Computer Science and Engineering
BRAC University
66 Mohakhali, Dhaka 1212, Bangladesh
shahria.hoque@g.bracu.ac.bd

4th Annajiat Alim Rasel
Senior Lecturer
Department of Computer Science and Engineering
BRAC University
66 Mohakhali, Dhaka 1212, Bangladesh
annajiat@bracu.ac.bd

***Abstract*—Students have the flexibility to learn anything with ease in the world of Massive Open Online Courses (MOOC) and open education systems because the learning information is readily available. However, this resource may encourage complacency in students. As a result, making predictions about student performance in advance becomes challenging. This study uses a univariate linear regression model to inform the student of his or her performance in advance. To forecast the final grade in one subject, we gathered the results of its internal exam components. For reliable results, the internal marks are adjusted to 100 (exam grade for a specific subject. Knowing the necessary number of points for the internal assessment also benefits students."**

***Index Terms*—Grade, Marks, Linear regression, Decision tree, prediction, analysis.**

I. Introduction

Student performance prediction in current educational systems is deteriorating day by day. Predicting student performance can assist students and their teachers in monitoring a student's progress. Today, a lot of institutions use a continuous evaluation method. These systems help pupils by enhancing their academic achievement. The ongoing evaluation mechanism is designed to benefit normal students. Unit tests or class tests are conducted regularly under the continuous assessment system. It is necessary to participate in every unit exam or class test to maintain consistency in performance for the subject's final grade.

II. Attribute Information

- school - student's school (binary: 'GP' - Gabriel Pereira or 'MS' - Mousinho da Silveira)
- sex - student's sex (binary: 'F' - female or 'M' - male)
- age - student's age (numeric: from 15 to 22)
- address - student's home address type (binary: 'U' - urban or 'R' - rural)
- famsize - family size (binary: 'LE3' - less or equal to 3 or 'GT3' - greater than 3)
- Pstatus - parent's cohabitation status (binary: 'T' - living together or 'A' - apart)
- Medu - mother's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - " 5th to 9th grade, 3 - " secondary education or 4 - " higher education)
- Fedu - father's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - " 5th to 9th grade, 3 - " secondary education or 4 - " higher education)
- Mjob - mother's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), '$at_$home' or 'other')
- Fjob - father's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), '$at_$home' or 'other')
- reason - reason to choose this school (nominal: close to 'home', school 'reputation', 'course' preference or 'other')
- guardian - student's guardian (nominal: 'mother', 'father' or 'other')
- traveltime - home to school travel time (numeric: 1 - $<$ 15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - $>$1 hour)
- studytime - weekly study time (numeric: 1 - $<$ 2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - $>$ 10 hours)
- failures - number of past class failures (numeric: n if 1$<$ = n $<$ 3, else 4)
- schoolsup - extra educational support (binary: yes or no)
- famsup - family educational support (binary: yes or no)
- paid - extra paid classes within the course subject (Math or Portuguese) (binary: yes or no)
- activities - extra-curricular activities (binary: yes or no)
- nursery - attended nursery school (binary: yes or no)

- higher - wants to take higher education (binary: yes or no)
- internet - Internet access at home (binary: yes or no)
- romantic - with a romantic relationship (binary: yes or no)
- famrel - quality of family relationships (numeric: from 1 - very bad to 5 - excellent)
- freetime - free time after school (numeric: from 1 - very low to 5 - very high)
- goout - going out with friends (numeric: from 1 - very low to 5 - very high)
- Dalc - workday alcohol consumption (numeric: from 1 - very low to 5 - very high)
- Walc - weekend alcohol consumption (numeric: from 1 - very low to 5 - very high)
- health - current health status (numeric: from 1 - very bad to 5 - very good) absences - number of school absences (numeric: from 0 to 93)

III. Methodology

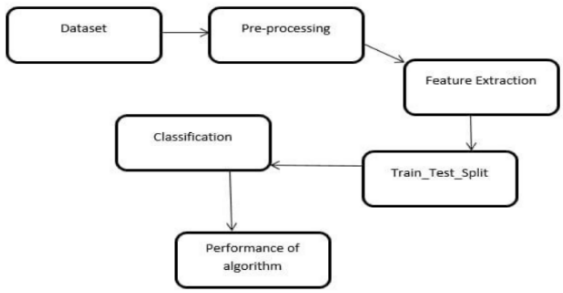
Since universities are prestigious places of higher education, students' retention in these universities is a matter of high concern. It has been found that most of the students' drop-out from the universities during their first year is due to lack of proper support in undergraduate courses. Due to this 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).

Flow-chart:



IV. Result Analysis

V. Conclusion

References

[1] Stage, S.A. and Jacobsen, M.D., 2001. Predicting student success on a state-mandated performance-based assessment using oral reading fluency. School Psychology Review, 30(3), pp.407-419.

[2] Campbell, A.R. and Dickson, C.J., 1996. Predicting student success: A 10-year review using integrative review and meta-analysis. Journal of Professional Nursing, 12(1), pp.47-59.

[3] Polyzou, A. and Karypis, G., 2016. Grade prediction with models specific to students and courses. International Journal of Data Science and Analytics, 2(3), pp.159-171.

[4] Iqbal, Z., Qayyum, A., Latif, S. and Qadir, J., 2019, February. Early student grade prediction: an empirical study. In 2019 2nd international conference on advancements in computational sciences (ICACS) (pp. 1-7). IEEE.

[5] Iqbal, Z., Qayyum, A., Latif, S. and Qadir, J., 2019, February. Early student grade prediction: an empirical study. In 2019 2nd international conference on advancements in computational sciences (ICACS) (pp. 1-7). IEEE.

[6] Iqbal, Z., Qayyum, A., Latif, S. and Qadir, J., 2019, February. Early student grade prediction: an empirical study. In 2019 2nd international conference on advancements in computational sciences (ICACS) (pp. 1-7). IEEE.