INTERNATIONAL JOURNAL OF
RECENT TRENDS IN MULTIDISCIPLINARY RESEARCH
March- April 2022, Vol 2(02), 06-08.
ISSN No: 2583-0368

Construction Extrapolative Models For Data Mining Schemes

Booraja P[1], Preethi S[2]
[1,2] *Dept. of CSE, Annai Teresa College of Engineering, Tamilnadu, India.*

Article Type:Research

OPENACCESS
Article Citation: BOORAJA P, PREETHI S, "Construction Extrapolative Models For Data Mining Schemes", International Journal of Recent Trends In Multidisciplinary Research, March-April 2022, Vol 2(02), 06-08.

Accepteddate:Mar18,2022

Published date : Mar25,2022

Published by 5th Dimension Research Publication.

Abstract: This paper focused in on building perceptive models for data mining projects and knowledge discovery functionalities. The objectives are data selection and transformation, Generation of a prediction models using classification data mining techniques, ID of different characteristics which affects retention and performance of students. The survey used dataset from the students pursued the BS Computer Engineering program.Decision tree classifiers such as ID3, J48 and Truck were used to build models. Results of the study showed that when the attribute evaluation was conducted using WEKA (Waikato Environment for Knowledge Analysis),the College Entrance Test (CET) got the highest significant regard among the perceived attributes in expecting the retention and execution of students while J48 got the highest accuracy rating when classifying instances. However, further research on factors or attributes that influence retention and performance of students should be explored and to include other programs in the School to deal with the accuracy of the results of collection.

Index Terms: Data Mining, ID3, J48, CART, CET, GWA, HSGPA, SCHLR

1. Introduction

Probation status is given to any student who procures last grades under 3 in half 74% of the total academic units in which he is chosen for the semester. Dismissed status is given to student who,at the end of the semester obtained a grade below 3 in 75% but less than 100% of unquestionably the quantity of educational units enrolled during the semester. Permanent disqualification status is given to student who gets last grade under 3 of every 100% of the academic units chose during the semester. Student in this academic status won't ever from now onward be allowed to pursue any program in all grounds of the University. A student may regain a status of good standing by passing more than 50% of the units in which heen rolled during the semester, otherwise, he will be dropped from the roll of the school where he belongs. Any student dropped from the roll of one college shall not be taken ownership of another school in the University. However, in extraordinary cases, where the student's aptitude is different from where he has failed. He may be recommended by the Guidance Office to the College Dean or Campus Head to be allowed to pursue another unit where his aptitude may be developed. Re-admission to any school of the University shall be allowed only once. Nore-admission of dismissed or prohibited student will be considered by the Dean without favorable recommendation from the Director of the Student Organizations or his supported representative.(ISU and You Student Handbook).

This general retention policy was formulated, approved And implemented to encourage students to study seriously and trive to earn college degrees. Early identification of vulnerable students who are leaned to drop their courses is crucial for the advancement of any up keep strategy. This would allow education institutions to undertake timely and proactive measures.Early recognizing confirmation of in peril students can be the recipient of academic and legitimate assistance to increase their chance of staying in the course and eventually complete the program.

The propelling changes in higher education system in the Philippines aimed at improving the performance of state Universities in changing financial condition. Thusly, retaining good students who will become well spring of HR to improve the economy of the country is a huge endeavor for the University. It is a challenge every higher education institution must accept and monitoring students‟‟ performance is not a basic task for an academic establishment with large number of students. The execution of support policy is sometimes too drastic for the student to accept and therefore early exposure of in peril students could cause the students to try harder to stay in the program. Thus, this study was concerned to build prediction models for performance and up keep of students using decision tree classifiers.

As this study was concerned to build prediction models for performance and support of students using decision tree classifiers the objectives are:

1. Data selection and transformation;
2. Generation of prediction models using classification data mining techniques;

3. Identification of different qualities which affects retention and performance of students; and
4. Comparison of accuracy on the classification techniques use in the prediction models.

2. Work Done/Contributions

A. Frame work of the Study

The audit used the Data Discovery Process (KDP) as portrayed by [7].The KDP figure was modified to suit the objectives of the survey. The modified version is presented on Fig.1.The process starts from preprocessing of data used in the study, followed by selection and change of data where data are changed into a type practical to the item used in data mining and interpretation and evaluation of the results to gain knowledge.

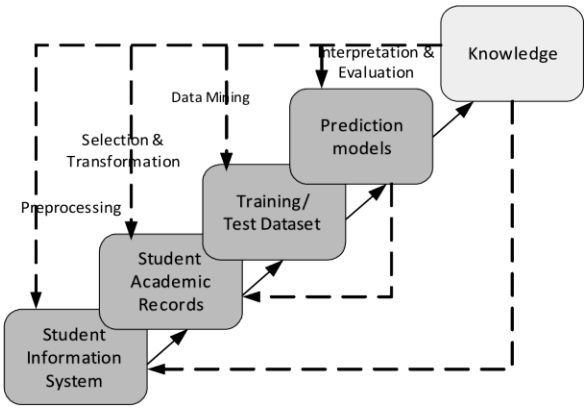

Fig.1. Frame work of the Study

B. Data Mining Process Frame work

Data mining process frame work in Fig.2 shows the data mining process used in the study. The data set is produced after data preprocessing. This fills in as commitment to the data mining device for the utilization of the picked classification algorithms such asID3, J48 and CART to produce prediction models. The prediction models are then evaluated based from their classification accuracy consistent with the results obtained from planning dataset. The data discovered can then be used for decision making.

C. Methodology Proposed Model

The proposed model makes estimate about support of students based on university entrance examination result, general point typical in optional school and past semester average.

If CET=good, HSGPA=first; GWA=high, then performance= high
On the off chance that CET=average, HSGPA= second; GWA=Medium, then performance= typical
If CET=poor, HSGPA=second, GWA=low, then performance= low
In case CET=good, HSGPA=third, GWA=Med, then performance= ordinary
If CET=poor, HSGPA=third, GWA=low, then retention=low

Application

The data aggregated from school students was analyzed using a decision tree algorithms such as ID3, J48 and CART. The following steps were performed in sequence:

Application Software

Weka 3.6 a popular data mining software was used to analyze the dataset.

Data Mining Process

The following steps consist of a data mining process:

a. Data Preparation /Preprocessing

In this step, data were cleaned removing those within complete information and completing those which can still be completed.

b. Data Selection and Change

The course of data decision and change involve careful assessment on which fields will be normal for data mining. Gotten factors from students related information were selected. Table I shows the marker and response variables which were derived from the data base.

c. Model Construction / Generation

Data Visualization

Fig. 3 shows that 64% of students got a GWA of >2.5,during first semester, 26% got a GWA of 2.5 or better and only 10% of students got a GWA of 2.0 or better.

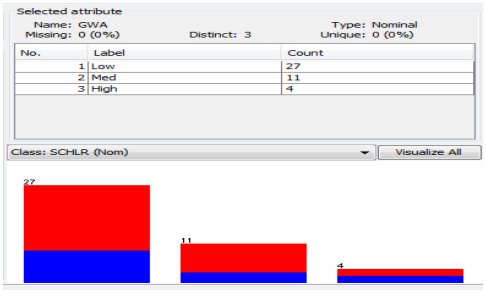

Fig.2.GWA Indicator

Fig. 4 shows that 24% of students got a score of 60 or higher during the college entrance test given by the university, 60% got a score of less than 60 but not lower than 40 and 16% got a score of less than 40 but not lower than 20.

3. Conclusion

Based from the delayed consequence of the audit the School Entrance Test (CET) got the highest significant value among the identified attributes in predicting the retention and performance of students while J48 got the most essential accuracy rating when classifying instances compared to ID3 and CART. However, further research on factors or attributes that influence support and execution of students should be investigated and to recollect various undertakings for the University to improve the accuracy of the results of classification.

References

1. *P.Cortez,and A.Silva,"Using data mining to predict secondary school student performance." In the Proceedings of 5th Annual Future Business Technology Conference, Porto,Portugal, 5-12,2008.*
2. *A. Galit, A. Hershkovitz, R. Mintz, and R. Nachmias, "Examining online learning processes based on log files analysis: a case study."Research, Reflection and Innovations in Integrating ICT in Education2007.*
3. *J. Han, and M. Kamber, "Data Mining: Concepts and Techniques," 2nd edition. The Morgan Kaufmann Series in Data Management Systems,Jim Gray,Series Editor,2006.*
4. *J. Han and M. Kamber, "Data Mining: Concepts and Techniques."Simon Fraser University,Morgan Kaufmann publishers,ISBN1-55860-489-8.2001.*
5. *S.T.Hijazi,andR.S.M.M.Naqvi,"Factors Affecting Student"sPerformance: A Case of Private Colleges," Bangladesh e-Journal of Sociology, Vol.3,No.1, 2006.*
6. *Z. N. Khan, "Scholastic achievement of higher secondary students in science stream," Journal of Social Sciences, Vol. 1, No. 2, pp. 84-87,2005..*
7. *Z. J. Kovacic, "Early prediction of student success: Mining student enrollment data," Proceedings of Informing Science & IT Education Conference2010.*