

**INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH
TECHNOLOGY**

**CLASSIFYING THE STUDENTS AT RISK THROUGH THE PERFORMANCE
REFLECTED IN PRE-REQUISITE AND NATURE OF MODULE IN HIGHER
EDUCATIONAL SYSTEM**

Mr.Sivakumar Venkataraman*, Mrs. Subitha Sivakumar

*Teaching Fellow, Faculty of Computing, Botho University, Gaborone, Botswana
Head of Department, Software Engineering Department, Botho University, Gaborone, Botswana

DOI: 10.5281/zenodo.51851

ABSTRACT

Educational institutions are growing diametrically to stand internationally in providing honorable merit employable graduates. Future prediction should be essential factor in an institute to take additional effort in fallout spaces. The existing student's performance data and the present nature of the module from the institutions are used to analysis and predict the earlier performance for a module. Data mining techniques like feature selection, clustering and predicting methods are been used in evaluating future prediction. The prediction are categories the features in three groups like conquer group, modest group and letdown group. Based on this groups, the institutions can have an appropriate indication on the upcoming results and to help the students in right time. Institutions can focus on the modest and letdown groups more than the conquer group students who are at risk.

KEYWORDS: Feature selection, clustering, predicting, data mining, student's performance data.

INTRODUCTION

Most Institutions are fronting competitors in launching the new different programs in the education to beat the real market. To stand in the international market the institutions needs to increase the student's performance or a strategy in rising the student's performance. Performance can be affected by how student perceive a module, falls on his learning behavior, modules under same category, level of difficult, past historical performance on related modules and pre-request modules. There is need to identify a group of students who may drop the performance and need to have a special care in their teaching and learning process.

This work helps the institutions in predicting the future outcome for individual modules in a program based on the student performance data and nature of the module. Feature selection, clustering methods and predicting methods are implemented in performance data to predict the student's at risk. Feature selection method is used as pre-processing technique to eliminate the redundant data, unrelated data and also used to decrease the number of attributes. Clustering and predicting method are applied in the features obtained after the feature selection to relate the predication. Prediction results are categorized as three outputs like conquer group, modest group and letdown group. Conquer group are stated as retention students those progress to next module, modest group are stated as unbalanced students and letdown group are stated as fallout students. Students falling in modest and letdown groups should be trained with additional care to reach the goal.

THE LITERATURE REVIEW

Student Performance data collected from the institutions are more multifaceted. Data mining is an important technique in computer based information system to handle huge number of data and to predict the results [1]. Educational data mining deals with the huge dataset that regulates impulsive problems in pattern recognition [2]. During the implementation of the data mining technique, [3, 4, 5, 6] recommends the feature selection should be first process as pre-processing method to eliminate, reduce the redundant and unrelated data and also to speed-up the procedure.

Similarly, feature selection helps the institutions to increase the performance, too retrieve the essential features and to select feature subset from the huge dataset based on certain standards [7, 2]. Pattern recognition is a major task in positioning the attributes in the groups and is also a challenge to discover the similar patterns from a huge amount of data [8]. Similar categories of scholars can be identified by analyzing the student's performance patterns and is also centered on the structure or nature of the module [9]. Clustering method is used to gather identical attributes into a group and the unrelated attributes to separate groups [10, 11]. The attributes belong to each group are termed as clusters, but entirely varied from the attributes in other clusters [12]. Trimming the attributes within each cluster will make more apparent and will show the similarities of each attribute where to have a perfect result in prediction [13].

PROPOSED WORK

Process Model for earlier prediction shown in figure 1 has four stages as input dataset stage, pre-processing stage, prediction stage and group stage.

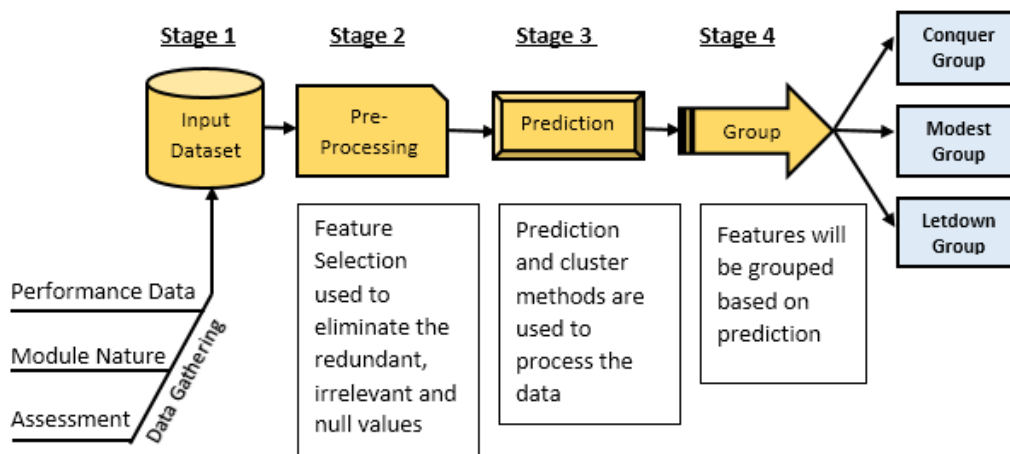


Figure 1: Process Model for earlier prediction

Stage 1: Students Performance Dataset, nature of the module and assessment for the module are the data gathering source provided as Input dataset in the first stage. Student performance data is the main source which has the features like module code, student id, batch code, attendance, internal marks, mid-exam marks, total marks, result, and supplementary marks. Nature of module and assessment is the other source which has the features like module code, module name, module assessment and nature of module assessments.

Stage 2: Input Dataset supplied will have duplicate values, irrelevant values and null values which to be eradicated. Feature Selection is used as pre-processing to eliminate those values and retains the rest of values in the dataset. Feature selection helps in increasing the prediction performance and also helps to speed-up the process.

Stage 3: In the third stage, the input dataset is processed based on prediction method and cluster method. The prediction technique will predict the important features and cluster method is used to group the predicted features as groups with related to expectation.

Stage 4: Fourth stage is the result where the groups are classified as three groups like Conquer group, modest group and Letdown group. Students appear in conquer group will have the capability to pass the current module, students appear in modest group has possibility as may or not pass and students appear in letdown group has less possibility in passing the module.

Students appear in letdown group should be considered as serious trainees and have to take care with more attention. Students in modest group have to be trained with appropriate guidance.

Steps to can be handled for students in letdown or modest groups

- Need to analysis where the students are lacking.
- Need to conduct support sessions
- Group of students can be taken care by separate tutor or leaders apart from regular sessions or timings.
- More activities or practice focused on examination should be done
- Regular monitor on student progress should be recorded to find their improvement

CHALLENGES

As a new prediction model, the following will be the major challenges during the implementation of the system.

1. Getting the appropriate Input data from the students
2. Tutors should spend more time in monitoring the students
3. Staff members should be involved in additional teaching development activities
4. Time consuming and cost of development will increase

CONCLUSIONS

This article is based on the predication from the student performance. The new model will be very useful for the institution to improve their results and for the students also to improve their leanings in their lacking subject. This model will help the institutions to identify the students who are lacking and can treat them as some special students by having extra classes or extra activities. The author(s) recommends to use this model to be effective and efficient to predict the students who are at risk in performance and can train them to overcome from that risk.

REFERENCES

- [1] E. V. George , W. F. Thomas and K. George , "The use of computer-based information systems by German managers to support decision making," *Information & Management* , vol. 41, p. 763–779, 2004.
- [2] Z. Yishi , Y. Chao , Y. Anrong , X. Chan , Z. Xingchi and Z. Zigang , "Neurocomputing," Feature selection for classification with class-separability strategy and data envelopment analysis, 2015.
- [3] M. Fan, H. Qinghua and Z. William, "Feature Selection with test cost constraint," *International Journal of Approximate Reasoning* , vol. 55, pp. 167-179, 2014.
- [4] F. Shobeir , S.-Z. Hamid and F. Farshad , "Bias and stability of single variable classifiers for feature ranking and selection," *Expert Systems with Applications*, vol. 41, p. 6945–6958, 2014.
- [5] V. Bolón-Canedo, N. Sánchez-Maró and A. Alonso-Betanzos, "Distributed feature selection: An application to microarray data classification," *Applied Soft Computing*, vol. 30, pp. 136-150, 2015.
- [6] F. M. C. David and Z. Hamidreza, "Data association mining for identifying lighting energy waste patterns in educational institutes," *Energy and Buildings*, vol. 62, p. 210–216, 2013.
- [7] Y. Ke and Z. David, "Feature selection and analysis on correlated gas sensor data with recursive feature elimination," *Sensors and Actuators B: Chemical*, vol. 212, pp. 353-363, 2015.
- [8] F. Alicia , G. Álvaro , L. Federico and P. Álvaro , "Pattern Recognition in Latin America in the “Big Data” Era," *Pattern Recognition*, vol. 48, p. 1185–1196, 2015.
- [9] B. Marija and M. Zivadin , "A web-based intelligent report e-learning system using data mining techniques," *Computers and Electrical Engineering*, vol. 39, p. 465–474, 2013.
- [10] P. Berkhin, "A Survey of Clustering Data Mining Techniques," *Grouping Multidimensional Data*, pp. 25-71, 2006.
- [11] K. Håkon and U. Kjetil , "Data reduction via clustering and averaging for contingency and reliability analysis," *Electrical Power and Energy Systems*, vol. 43, p. 1435–1442, 2012.
- [12] B. Sasan and M. Mohammad , "Developing an approach to evaluate stocks by forecasting effective features with data mining methods," *Expert Systems with Applications*, vol. 42, pp. 1325-1339, 2015.
- [13] H. Chiao-Tzu , L. Wen-Tsann , W. Shen-Tsu and W. Wen-Shan , "Planning of educational training courses by data mining: Using China Motor Corporation as an example," *Expert Systems with Applications*, vol. 36, pp. 7199-7209, 2009.