

## International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X

*IJCSMC, Vol. 4, Issue. 3, March 2015, pg.681 – 683*

### **RESEARCH ARTICLE**



# A Study on Data Mining Techniques and Genetic Algorithm in Education Sector

Saddam Khan<sup>1</sup>, Sunny Gupta<sup>2</sup>, Yogesh Kumar Sharma<sup>3</sup>, Dr. Radhakrishna Rambola<sup>4</sup>

<sup>1</sup>SCSE & Galgotias University, India

<sup>2</sup>SCSE & Galgotias University, India

<sup>3</sup>SCSE & Galgotias University, India

<sup>4</sup>SCSE & Galgotias University, India

<sup>1</sup> [sksaddam127@gmail.com](mailto:sksaddam127@gmail.com); <sup>2</sup> [gupta08aug@gmail.com](mailto:gupta08aug@gmail.com), <sup>3</sup> [yogesh.sgi1992@gmail.com](mailto:yogesh.sgi1992@gmail.com), <sup>4</sup> [dr.rambola@gmail.com](mailto:dr.rambola@gmail.com)

---

**Abstract**— *This paper is examine the analysis of students' and prediction of students' performance. Data mining techniques are playing vital roles in Higher education institution. This paper is reviewing some data mining techniques like association rules, classification technique, clustering, outlier detection and genetic algorithms in Education sector. This paper is also analysis that how much these techniques are beneficial in Higher education.*

**Keywords**— *Educational data mining, Association rules, classification, clustering, outlier detection, genetic algorithm.*

---

### I. INTRODUCTION

Educational data mining is a new method that discovers the knowledge from the educational environment. Educational data mining techniques is playing a vital role to improve the students' performance. Several techniques which are uses in educational data mining are – Artificial Neural Network, Support Vector Machine, naïve Bayesian, Decision tree, etc. The objective of education sector is to improve the students' performance and provide quality education. A way to achieve the highest level of quality in education sector is by discovering knowledge from educational institution. And learn the main attributes that may improve the students' performance in all kinds of education sector [7].

Association rules is defining as a relationship between items which is suitable in given data set. To make relationship among students' data association role play a vital role in Higher education. Association rules also play a big role to prediction of students' performance. It helps to searching rules in the form of if-then rule. It can help to predict the class of item and classify the item sets. In higher education, for relationship among students' and his course association rules are uses.

A process of data mining which is classifies the class label or grade from data set is called the classification. In higher education for prediction the student performance we classify the students' grade in various class labels. Classification method have various technique like decision tree, Naïve Bayesian, artificial neural network, support vector machine, etc., which is used in educational data mining to classify the data set of higher educational institutional. A rule based classifier is also uses for classify the data set. It is show the relationship between data set attributes and label of class.

In Higher Education when student need to be in a group a data mining technique which name is clustering. Clustering is a process of data mining in which we can create groups of objects according to their similarity on the basis of many variables like

students' performance, age, sex, etc. The main goal of clustering is finding the high quality cluster. It includes intra-cluster distance are minimized and inter-class are maximized. In this paper we review k-means clustering technique. Selection of suitable cluster centre which will be centroid is the main aim of k-means clustering.

Outlier detection, existing data object is do not obey with the general performance or model of data. These data objects are totally different or unpredictable with the remaining set of data. It can be caused by quantity or execution error. Various data mining algorithms are try to minimize the effect of outliers and these algorithms are remove them all together but the result is loss the important hidden information. Outliers may be particular interest in the case of fraud discovery [12].

Genetic algorithm is techniques which make students predication more accurate. Using genetic algorithm in higher education the error rate is low and it is used for finding optimal solution of a problem in higher education. Genetic algorithm is uses two operator crossover and mutation operators which is based on the Darwinian Survival of the fittest principle. This algorithm can be used for finding optimal solution to complex problem in higher education and it is utilize the principle of natural selection.

## II. RELATED WORK

Performance of student is great concern in university courses. Qasem A. Al-Radaideh, Emad M. Al-Shawakfa and Mustafa I. Al-Najjar is use the data mining processes, mainly classification, to help provided in quality of the higher educational system by assessing student data. Authors purpose, CRISP framework for data mining is used for mining student related academic data. Authors are mainly uses Classification rules – decision tree. Decision tree are uses to generated rules to make a final grade in a course [1].

Improve the current trends in the higher education systems is very important. P. Veeramuthu and Dr. R. Periasamyim says that by using data mining methods, we can choose reliable students and then we can motivates them to know about deep knowledge of higher education system. For these both the authors are using association analysis to find the interesting relationship, if – then rules for classify each item in set of data, and k-means cluster are uses for best groups of data of similar objects [2].

In educational management system Educational Data mining (EDM) is playing an significant role. EDM play a vital role in course management system also. It deals with its theoretical and practical features. In today world, for learning resources are available but it is very huge. [3] Dr. B. L. Shivakumar, Mr. V. Muruganathan is deploying application of EDM for the University. For this cluster are developed using RHadoop and for association rules Rapid miner is uses. In this whole eLearning CMS system are used in which 350 faculties and 8500+ learners.

Abeer Badr El Din Ahmed and Ibrahim Sayed Elaraby is use ID3 algorithm for data classification. ID3 algorithm is decision tree method. In [4] classification task is used for predict the grade of student. In [4] the paper, various technique are used for identify the fail students. After finding fail students the teacher need to special attention to reduce failing rate [4].

In education sector data mining techniques are play vital role on business point of view. Dr. Mohd. Maqsood ali observes that various data mining technique like decision tree, Regression analysis, Neural network and cluster analysis which contribute to improve business in education sector to offer competitive courses, improve students' academic performance and teachers' performance [5].

The process to provide quality education is describes by Brijesh Kumar Baradwaj and Saurabh Pal (2011) to provide quality education higher educational institutions. To achieve highest level of quality the authors are uses two ways one is by discovering knowledge for prediction regarding enrolment of students' in a particular course, detection of abnormal values in the result sheets of the students another ways is classification task to evaluate student's performance and decision tree that are used for data classification [6].

In [9] Behrouz Minaei-Bidgoli and William F. Punch is uses combination of four classifiers to leads to a significant improvement in classification performance. For minimize the error rate and improvement in prediction the authors is uses feature weighting vector using genetic algorithm. Genetic algorithm produces many different groupings of features, to extract new features and improve prediction accuracy. It also shows that when the number of features is few; feature weighting is works better than just feature selection [9]. To find association rule in [9] Behrouz Minaei-Bidgoli and William F. Punch are apply Evolutionary Algorithms that is classify the students and problems.

To analyse much accurate information of students' performance Sajadin Sembiring, M. Zarlis, Dedy Hartama, Ramliana S, and Elvi Wani says that there are various prediction model exist with different approaches but there is no predictors available which can identify that whose student will be an academic topper, a drop out, failure or an average performer accurately. For determine the relation between behaviour of students and success of students [10] is using the kernel method. They are used smooth support vector machine for prediction and k-means clustering for grouping of students [10].

To improvement of students' performance Monika Goyal and Rajan Vohra describe that data mining techniques like clustering, decision tree and association rule which are play important role in higher education institution. This techniques are also help to prediction of variety of courses, to measure their retention rate and the grant fund management in higher educational institutional [11].

In [14] recommender systems are focus to enhance the process of teaching and learning to improve the standard of education. Hana Bydzovska, Lubomir Popelinsky are using data mining technique like machine learning and classification

technique to predict the student performance and compare the both techniques. Analysis of study related data and social behaviour data the authors are uses social network to achieve results [14].

### III. CONCLUSIONS

This Paper is analyzing various techniques like association, classification, cluster and genetic algorithms here which are uses in Education sector to predict students' performance and find some interesting result. Association rules are uses when a relationship finds between students in education sector. To classify the data and predict the data classification plays a vital role. Decision tree is better in term of speed and ANN is better in term of accuracy and error rate but it is very slow. To collect groups of students cluster are uses. Genetic algorithms play a vital role in higher education and reduce the error rate.

### REFERENCES

- [1] Qasem A. Al-Radaideh, Emad M. Al-Shawakfa and Mustafa I. Al-Najjar, "Mining Student Data Using Decision Trees", The 2006 International Arab Conference on Information Technology, ACIT, Nov. 2006, Jordan.
- [2] P. Veeramuthu and Dr. R. Periasamy, "Application of Higher Education System for Predicting Student Using Data mining Techniques", IJIRAE, Volume 1, Issue 5, June 2014.
- [3] Dr. B. L. Shivakumar, Mr. V. Murugananthan, "A Novel Application Framework for Educational Data Mining towards Automated Learning System", 2014 International Conference on Intelligent Computing Applications, IEEE, 2014.
- [4] Abeer Badr El Din Ahmed and Ibrahim Sayed Elaraby, "Data Mining: A prediction for Student's Performance Using Classification Method", World Journal of Computer Application and Technology 2(2): 43-47, 2014.
- [5] Dr. Mohd Maqsood Ali, "Role of data mining in education sector", IJCSMC, Vol. 2, Issue. 4, pp. 374 – 383, April 2013.
- [6] Brijesh Kumar Baradwaj , Saurabh Pal, "Mining Educational Data to Analyze Students' Performance", IJACSA, Vol. 2, No. 6, 2011.
- [7] Mohammed M. Abu Tair, Alaa M. El-Halees, "Mining Educational Data to Improve Students' Performance: A Case Study", ICT Journal, Volume 2 No. 2, February 2012.
- [8] Kumar, V. and Chadha, A. (2011) 'An Empirical Study of the Applications of Data Mining Techniques in Higher Education', International Journal of Advanced Computer Science and Applications, vol. 2, no. 3, pp. 80-84.
- [9] Behrouz Minaei-Bidgoli and William F. Punch, "Using Genetic Algorithms for Data Mining Optimization in an Educational Web-Based System", Springer-Verlag Berlin Heidelberg, pp. 2252–2263, 2003.
- [10] Sajadin Sembiring, M. Zarlis, Dedy Hartama, Ramliana S, Elvi Wani, "Prediction of student academic performance by an application of data mining techniques", IACSIT, IPEDR vol.6, (2011).
- [11] Monika Goyal, Rajan Vohra, "Applications of Data Mining in Higher Education", IJCSI, Vol. 9, Issue 2, No 1, March 2012.
- [12] Jiawei han and Micheline Kamber, "Data Mining Concepts and Techniques, Second Edition, Morgan Kaufman, Elsevier, 2006.
- [13] Arun K Pujari, "Data Mining Techniques", 3rd edition, Universities Press (India) Private Limited, pp. 98-308, 2013.
- [14] Hana Bydzovska, Lubomir Popelinsky, "Predicting Student performance in Higher Education", 24th International Workshop on Database and Expert Systems Applications, 2013.