Application of Machine Learning Method for Students' Learning Behaviour Analysis in Moodle Learning Management System

[1]Musa Wakil Bara, [2]Hamisu Alhaji Ali [3]Mohammed Ali Bizi
[1&2]Department of Computer Science, Mai Idris Alooma Polytechnic Geidam, Yobe Nigeria
[1&2]Department of Computer Science, Federal Polytechnic Damaturu, Yobe Nigeria

ABSTRACT: This research investigates the performance of K-Means Clustering Method to analyze students' e-learning activities with the aim to identify clusters of students who use the e-learning environment in similar ways using data obtained from the log files of their actions as input. The K-Means clustering technique was used to group 53 Computer Science Students from Universiti Teknologi Malaysia into three clusters. Learning behaviors of students in each cluster were analyzed; a relationship between students' learning behaviour and their academic performance (Final Results) was investigated. The analysis shows that students in Cluster1, having the highest interactions frequency with the e-learning, also got the highest final score mean of 91.12%, this followed by Students in Cluster2 with less number of interactions than Cluster1 and final score mean of 75.65%. Finally, students in Cluster3 have least number of interactions than the remaining clusters with least final score mean of 36.57%. The research shows that, students who participate more in Forum activities perform higher, while students with lowest records in Forum activities have the lowest performance. The research found that Forum activity has significant factor on student's course success but it is optional to students and no marks allocated to it. The research suggests that marks should be allocated to Forum activities to encourage students' participations.
KEYWORDS: - *Machine Learning Learning, Management System, ELearning, Data Mining*

I. INTRODUCTION

Electronic mode of performing business and other transactions has dominated every facet of human endeavour. Education sector is not an exception. Most of the developed and of course the developing countries have adopted the use of ELearning (also written as E-Learning) as means of knowledge sharing [1]. The Learning Management system is the platform that supports the delivery of learning materials and also provides everything needed for knowledge delivery, these tasks that are provided by the LMS are called the E-Learning activities. E-Learning has not replaced Face-to-Face method of learning but it complements it. There are various learning management systems in use, these include the Modular Object Oriented Dynamic Learning Environment (Moodle), the A-Tutor, the BlackBoard, etc. Among these LMSs, Moodle is the most commonly used LMS especially in developing countries. Moodle LMS provides access to the courses and facilitates communications between students and tutors or among the students.

Log files are used to store students' activities in the LMS, these activities can be mined using Data Mining methods to study the learning behaviour of the students. Educational data mining is the application of data mining processes that is concerned with development of methods for discovering the uniqueness of data that come from educational database, and use the methods to for understanding learners and their learning skills [2].

This research is concerned with students' involvement in eLearning. The researchers studied students' learning behaviors and analyses the effects of their learning performance upon carrying out the e-learning activities. It also discovers which activity has more significance over the others. The research uses data obtained from students' log files in Moodle LMS, the students are from Computer Science undergraduate taking a Data Structure Course at Faculty of Computing, Universiti Teknologi Malaysia. The research is aimed to investigate the performance of K-Means clustering technique to obtain clusters of students who have similar way of interacting with the e-learning

environment, and analyze the effects of their learning behaviours on their learning performance

II. LITERATURE REVIEW
a. The learning Management System

Learning Management System (LMS) is the latform that supports the management of Elearning activities. Among the functions of the LMS are the provision of learning materials to students, giving the students the ability to access the learning materials, interact with teachers and students, submit assignment and so on [3]. Fig.1 shows how LMS works.
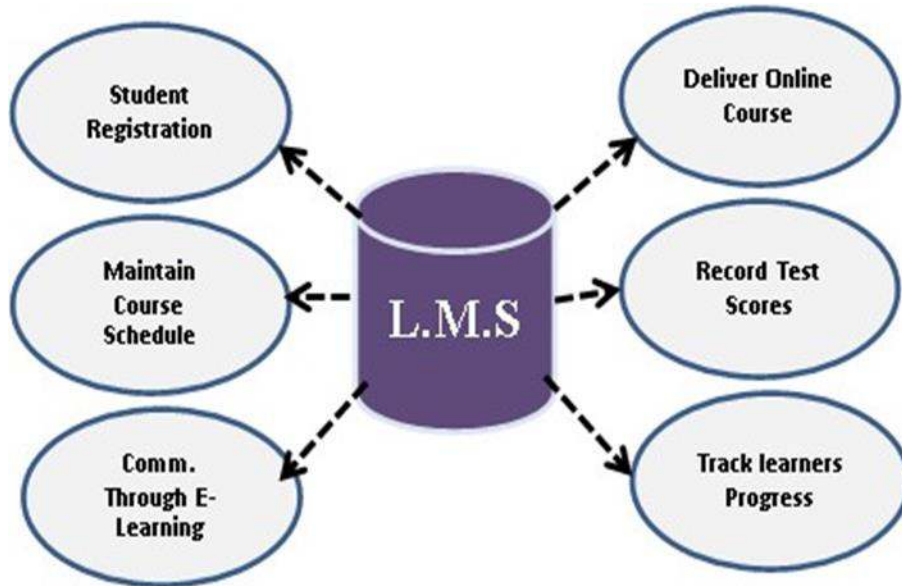

Figure 1. Some Benefits of Learning Management System

e-Learning Management System (eLMS) is also known as Course Management Systems (CMS), and the two terms may be used interchangeably. Many universities have already adopted CMS and some others are planning on introducing CMS software such as Moodle or Blackboard to support their learning operations. However, currently, there has not been much research to explore the influence of Power (electricity) on the use of CMS. To this end, it is vital to learn its perceived usefulness from the perspectives of students. The results of this study could help universities and by extension, higher institutions make better investment decisions and help instructors in using this technology more effectively. Additionally, it can help designers of course management software to improve the learning tools and get higher satisfaction level in

using their final semester result.

the learning environment (Nicholas-omoregbe, 2017).

b. Educational Data Mining

Educational Data Mining (EDM) is an emerging discipline, concerned with development of methods for exploring the uniqueness of data that come from educational database, and use those methods to better understand learners and their learning settings [4]. These methods include classification, clustering, prediction and visualization. While the tools used in EDM include Waikato Environment for Knowledge Analysis (WEKA), Rapid Miner, MATLAB, etc.
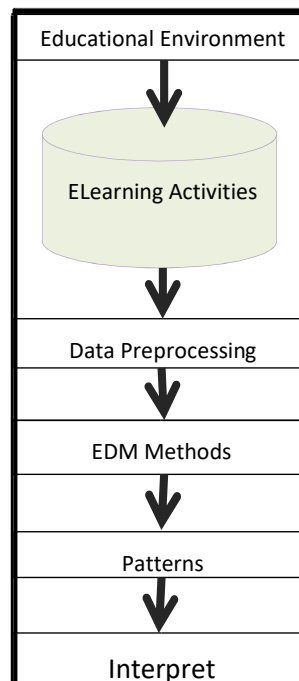
Educational data mining fits various research works in eLearning such identifying students' learning styles, prediction of students' performance as in [5] and determining

Fig. 2 Educational Data Mining Steps

the situation, characteristics and phenomena that affect students' performance. Reference [4] used educational data mining technique to improve the comprehensibility of the learning materials, students were clustered based on their behavioural usage of the e-learning, however, the researchers did not state which activities is more significant than the other.

Fig. 2 depicts the steps of mining student activities in EDM.

c. K-MEANS CLUSTERING TECHNIQUES

In this steps k-means clustering algorithm was applied to the proposed data and get valuable information, k- means is an old and most widely used clustering algorithm by MacQueen in 1967 [6]

Algorithm of Basic K-means Algorithm:

1. Select K points as the initial centroids.

2. Repeat.

3. From K- cluster by assigning all points to the closest centroids.

4. Recomputed the centroid of each cluster.

5. Until The centroids don't change.

A. METHODOLOGY

Some information contained in raw data have no effect on accuracy of the method used during the mining process. The data must first be pre-processed before carrying out our analysis on the collected data. The importance of pre-processing is to make the data clean from noisy data, missing values or inconsistent data. The major tasks involved in data pre-processing include data cleaning in which the raw data is being cleaned; these include removal of noisy data, filling the missing values and identifying outliers; data transformation which involved transforming the cleaned data into format that is readable by the data mining tool. In this research, the cleaned data was transformed to Comma Separated Version (CSV) before been exported to MATLAB for clustering.

B. EXPERIMENT AND RESULT

a. The Dataset

The dataset was collected from the Universiti Teknologi Malaysia (UTM) Moodle LMS log records. It contains the activities of Computer Science undergraduate students from Faculty of Computing, Universiti Teknologi Malaysia during the period of first semester 2014/1015 session. The activities of 52 regular students were monitored from the beginning of the semester to the end. The semester lasted for fourteen weeks. The data was downloaded in excel

format. The interface is the medium that is used to carry out the data collection from the log records.

b. Training the K-Means

The network was then trained using the data at hand and the result obtained is shown in Fig. 3. The result shows three clusters; each cluster contains instances associated to it.
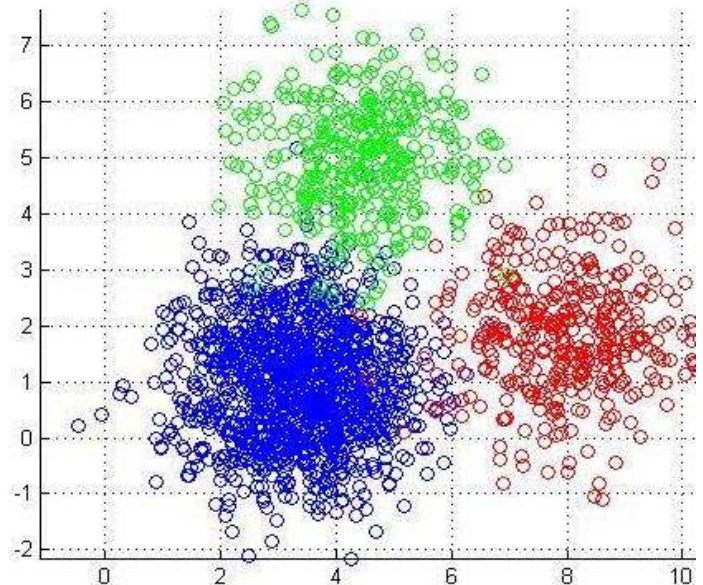
a. Experimental Result

Figure 2. Three (Clusters) of Students According to their Similarity in Learning Behaviour

From the experimental result, the students are grouped into three clusters. Each cluster contains students with similarity in interaction with the e-learning environment.

Cluster1 contains students who have the highest number of actions thus; they are termed as Very Active Students. Students in Cluster2 have more actions than those in Cluster3 and they are termed as Active Students. Cluster3 students have lowest actions hits and are termed Non-Active, this categorization is based on students' interactions and in line with the work of [7] categorization. On the other hand and based on students' course success, Clusters are categorized as High Learners (Cluater1) with the highest marks 91.12% in their Final Scores, Medium Learners (Cluster2) with Final Score marks 75.65% and Low Learners (Cluster2) with the lowest Final Score marks of 36.57. This categorization is in line with [4]. Fig. 4 shows the graphical representation of Actions hits means and the Final Score Means of each Cluster.
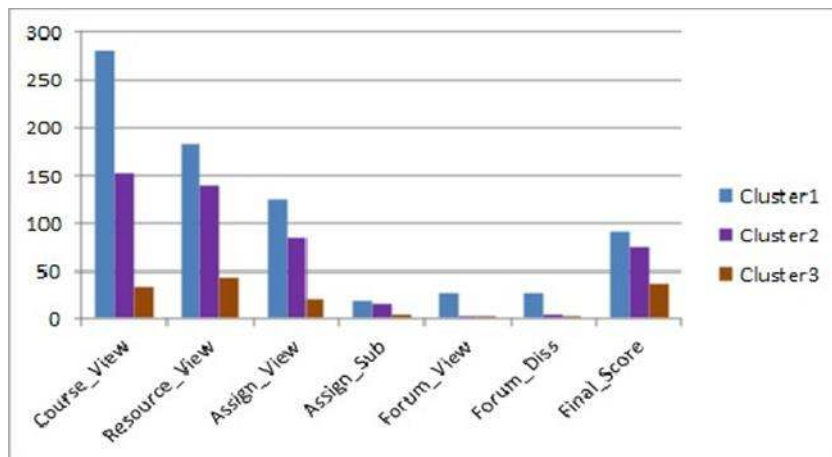
It was learnt that participation in Forum activities by students is optional and there is no mark allocated to it, this is why students feel reluctant to participate. However, if marks can be allocated to Forum participation, students can give it a priority and this would improve their learning performance.

C. RESULT FINDINGS

From the experiments carried out, it was learnt that, students' actions on e-learning activities differ according to their learning behaviours. As a result, the students are grouped into clusters such that each cluster contains students with similarities in learning behavior. The research discovered three categories of students, those that are Very Active (Cluster1), students of this group have higher number of interactions with the E-learning especially participation in Forum activities and they have the highest grades. Active (Cluster2), students of this group have average number of interactions with the ELearning and have average final grades, those that are Non-Active (Cluster3) students of this group have least number of interactions with the E-learning and they all failed the course.

It can therefore be concluded that, students' learning behaviour contributes to their learning performance. It was also learnt that, the E-learning activities differ in contributing to student's learning performance. Students who participate more in Forum (Cluster1) have high chance to outperform the student who do not participate more in Forum. Likewise, students who failed to submit their assignments regularly have chance to fail the course.

D. CONCLUSION

In this research, the implementation of K-Means Clustering Method to group students according to their similarities in interaction with the E-learning was explained. Three clusters were obtained and are termed based on students' learning success as High Learners (Cluster1) with Final Score of 91.12%, Medium Learners (Cluster2) with Final Score of 75.65% and Low Learners (Cluster3) with Final Score of 36.57% in line with [4], as Very Active Students (Cluster1), Active Students (Cluster2) and Non-Active Students (Cluster3) in line with (Bara, 2018) categorization. The researchers also conducted analysis on the correlation between students' learning behaviors and their course success (performance). The result showed that, students that fall in both 'High Learners' and 'Very Active' group (Cluster1) emerged to be best in their course success. The Final grade is A$^+$ or an excellent pass for this group. Cluster2 contains students that are Active and High Learners, and passed with an A grade. Finally, Cluster3 students emerged to be both Low Active and Low Learners with an F grade which means they failed the course.

REFERENCE

[1] L. S. Jamil, "ASSESSING THE BEHAVIOURAL INTENTION OF STUDENTS TOWARDS LEARNING MANAGEMENT SYSTEM , THROUGH TECHNOLOGY ACCEPTANCE MODEL - CASE OF IRAQI UNIVERSITIES," vol. 95, no. 16, pp. 3825–3840, 2017.

[2] M. W. Bara, "Self-Organizing Map Clustering Method for the Analysis of E-Learning Activities," 2018.

[3] J. Broadbent, "Academic success is about self-efficacy rather than frequency of use of the learning management system," vol. 32, no. 4, pp. 38–49, 2016.

[4] I. P. Ratnapala, R. G. Ragel, and S. Deegalla, "Students behavioural analysis in an online learning environment using data mining," *2014 7th Int. Conf. Inf. Autom. Sustain. "Sharpening Futur. with Sustain. Technol. ICIAfS 2014*, pp. 1–7, 2014, doi: 10.1109/ICIAFS.2014.7069609.

[5] A. Tick, "Research on the Digital Learning and E-learning Behaviour and Habits of the Early Z Generation," pp. 33–38, 2018.

[6] R. B. Bhise, S. S. Thorat, and A. K.

Supekar, "' Importance of Data Mining in Higher Education System ,'" vol. 6, no. 6, pp. 18–21, 2013.

[7]   C. Yin, F. Okubo, A. Shimada, M. Oi, and S. Hirokawa, "Analyzing the Features of Learning Behaviors of Students using e-Books," no. April 2014, pp. 617–626, 2015.