A Novel Approach to Enhance Teaching and Learning Through Mining and Learning Analytics

1N. Menaga, 2Dr. B. Umadevi
1Research Scholar, 2Assistant Professor & Head,
1P.G. & Research Department of Computer Science
1Raja Doraisingam Govt. Arts College, Sivaganga, Tamil Nadu, India.
2P.G. & Research Department of Computer Science
2Raja Doraisingam Govt. Arts College, Sivaganga, Tamil Nadu, India.

Abstract
Information and communication technologies brought-in tools and techniques in the field of education that introduced new concepts of teaching and learning. Learning management System is one of the key tools used in educational institutes to facilitate e-learning. There is remarkable digital divide among urban and rural areas. Use of e-learning programs has been intensifying, with high demand for education, particularly in developing countries. The research addresses the problem of e-learning service quality in higher education environment. So it is more important to assess the quality of e-learning offerings, as it is critical to the program success and survival. In general the objective is to have a cost-effective learning environment equipped with latest technologies to provide learners an opportunity to get insight into new information and communication technologies and e-learning environment.

Index Terms - E-Learning, Service Quality, Satisfaction, Learning Content.

I. INTRODUCTION

E-Learning is the use of internet-based courses or programs that deliver instruction using pedagogical tools as part of a formal educational program. Due to drastic improvement in information and communications technology (ICT) the E-learning has obtained a remarkable position. E-Learning differs extensively from traditional classroom education by its synchronous and asynchronous behavior. With synchronous E-Learning, students and faculty interact simultaneously at a specific time using the internet for direct communication [1]. With asynchronous E-Learning, the students and faculty interact at different times by placing messages or coursework in files that are accessed at different times [2]. With either approach, E-Learning provides a learning approach allowing an educational institutions without physical presence of the student or faculty in the campus. Many higher education institutions in the Middle East are increasing their emphasis on ICT education and access because of the need for these institutions to compete with universities outside the region in retaining students. This can be partially attributed to the difficulties with the practical development and implementation of E-Learning programs.

This study analyses the effect of student attitudes on the acceptance of E-Learning. The data mining technology is applied to predict the behaviour of E-Learning Quality of students assessment based on the subjective norms and perceptions. It also identifies key variables that can be measured and analyzed to support an empirical assessment of the effect of the variables on the intention to adopt E-Learning.

The paper is structured as follows. The section one describes about introduction. Section 2 deals about background study and its related works. The methodology of the research work is explained in section three. In section four portraits the Experiment results. Finally the paper is concluded in last section.

II. BACKGROUND AND RELATED WORKS

Data mining has its own set of techniques that can be used to mine relevant and attractive knowledge from data. Data mining has several techniques such as association rule mining, classification and prediction, and clustering. Classification techniques are supervised learning techniques that classify data item into predefined class label [3]. It is one of the most useful techniques in data mining to build classification models from an input data set [4]. The classification techniques commonly build models that are used to predict future data trends.

2.1 Literature Review

In this paper, diversified approach using various data mining techniques are collected to analyse and predict the e-learning quality of student’s at various levels. The study related to data mining for extracting and predicting the e-learning rate prediction in various models and the comprehensive literature review of various researchers’ works are stated below:

N. Venkatesan designed to provide an up-to-date snapshot of the current state of research and applications of Data Mining methods in education and e-learning process [5]. Educational data mining concerned with developing methods for discovering knowledge from educational domain [6]. He used data mining algorithms to discover pedagogically relevant knowledge contained
in databases obtained from Web-based educational systems [7]. He also analysed the performance of the Students’ based on their learning capability, detection of irregular learning behaviors, e-learning system navigation and interaction optimization, clustering according to similar e-learning system usage and systems’ adaptability to students’ requirements and capacities [8].


eLearning tool and pre-process the collected record for the best solutions to understand the behavioural patterns of the students. Their study gave a clear-cut idea of various features extracted for analyzing the behavioural pattern of a student with eLearning tool. They also highlighted new developments including exploration of data from massive open online courses (MOOCs).

D. Masoumi and B. Lindstrom addressed the concerns regarding enhancing and assuring quality in e-learning, a comprehensive e-quality framework is developed by taking into account the pros and cons of the previous models, [10] frameworks, and studies of e-quality. This e-quality framework offers a structured set of factors and benchmarks as a tool for practical quality work with e-learning in virtual institutions.

III. METHODOLOGY
E-learning analytics provides sophisticated investigative facility to improve the learning and education. It draws from other fields of study including business intelligence, web analytics, academic analytics, educational data mining, and action analytics [11].

The teaching and learning objectives are:

1) To identify students at poor learning and to improve their educational standards
2) To provide suggestions in relation to reading material and learning activities.
3) To detect the need for, and measure the results of, pedagogic improvements.
4) To adapt course offerings.
5) To identify teachers with good performance and who needs new teaching methods.
6) To assist in the student recruitment process.

The applications of various data mining techniques [12] which is adopted as a methodology to analyse and predict the E-Learning ability of the students. The most predictive [13] data models applied are artificial neural networks [14], Naive Bayes, K-Nearest Neighbour (KNN), Support Vector Machine, logistic regression, classification trees, classification and regression trees and discriminant analysis. The following section describes the popular models which are used to analyse the learning behaviour and characteristics of Students’. The steps involved in for analyzing the learning behaviour of Student’s is explained in Figure 1.

3.1 K-means Clustering
The K-means algorithm is one of the clustering algorithms with it popularity is in widespread manner. K-means is a partitioning algorithms and it has two main advantages one is is very easy to implement and it takes little time to run, which makes it suitable for large data sets. The K-means is the most common partitioned clustering algorithm which is used to partition n observations into K clusters in which each observation belong to cluster with nearest mean. It is simple, non supervised iterative learning method. The idea behind classifying set of data objects into K number of clusters where K is fixed initially. It first fixes initial group centroids. Then assign each object to the group that has closest centroid. Once all the objects are assigned it recalculate positions of centroids [15]. Again repeat the same process until centroids not change.

3.2 DBSCAN
DBSCAN (Density Based Spatial clustering of application with noise) is to create clusters with minimum size and density. Density is defined as the minimum number of points within a certain distance of each other. It requires two parameters: epsilon (eps) and minimum points (minPts). DBSCAN does not require you to know the number of clusters in the data a priori. DBSCAN does not have a bias towards a particular cluster shape or size [16]. DBSCAN is resistant to noise and provides a means of filtering for noise if desired. DBSCAN does not respond well to high dimensional data. As dimensionality increases, so does the relative distance between points making it harder to perform density analysis. DBSCAN does not respond well to data sets with varying densities.

3.3 Random Tree
Random Tree is a supervised Classifier; it is an ensemble learning algorithm that generates many individual learners. It employs a bagging idea to produce a random set of data for constructing a decision tree. In standard tree each node is split using the best split among all variables. Random trees have been introduced by Leo Breiman and Adele Cutler. The algorithm can deal with both classification and regression problems. Random trees is a collection (ensemble) of tree predictors that is called forest. The
classification works as follows: the random trees classifier takes the input feature vector, [17] classifies it with every tree in the forest, and outputs the class labels that received the majority of “votes”. In case of a regression, the classifier response is the average of the responses over all the trees in the forest.

IV. EXPERIMENT RESULTS
The data set is collected from various College students’ who are studying at Sivagangai district. At the outset a sample of 300 data is taken for analysis with different attributes. The description of Attributes is explained in Table 1.

The data set has been pre-processed; The K-Means and DBSCAN algorithms are applied over the data to partitioning the data into (K = 5) different clusters. The Random initialization method and the ten as the initial seed value are used to form the clusters. In K-Means, the Euclidean Distance and in DBSCAN, the Manhattan Distance metrics are used to measure distance between of an observation and the initial cluster centroids.

From the cluster data, the E-Learner [18] behaviour is analysed through the classification [19] technique Random Tree. In the classification the E-Learning Quality is taken as the class variable. In the classification, the class (K=3) is randomly chosen attribute at each node to construct a tree. The outcomes of the results are used to identify or categorize the different types of E-Learners. The interpretations are given in Tables 2 – 4 and Figures 2– 6.

<table>
<thead>
<tr>
<th>Attribute -Names</th>
<th>Type</th>
<th>Possible-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaboration</td>
<td>Number</td>
<td>1 - Brainstorm 2 - groups-of-students 3 - Real-Time-Students 4 - Online-Chatting</td>
</tr>
<tr>
<td>Connectivity</td>
<td>Number</td>
<td>1 - web-browser 2 - Chromebooks 3 - iOS 4 - Android devices</td>
</tr>
<tr>
<td>Personalisation</td>
<td>Number</td>
<td>1 - laptop 2 - desktop 3 - Tablet 4 - Smartphone 5 - Read 6 - Listen 7 - Watch</td>
</tr>
<tr>
<td>Cluster Name</td>
<td>Collaboration</td>
<td>Connectivity</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Cluster-0</td>
<td>1.4032</td>
<td>2.2742</td>
</tr>
<tr>
<td>Cluster-1</td>
<td>2.2807</td>
<td>1.5614</td>
</tr>
</tbody>
</table>

**Figure 2 : Clustering by K-Means**

**Figure 3 : Clustering by DBSCAN**

**Table 2 : Clustering by K-Means**
Table 3: Clustering by DBSCAN

<table>
<thead>
<tr>
<th>Cluster-Name</th>
<th>Collaboration</th>
<th>Connectivity</th>
<th>Personalisation</th>
<th>Interactivity</th>
<th>Extended-Oppotunities</th>
<th>Flexibility</th>
<th>Motivation</th>
<th>E-Learning-Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster-0</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>Good</td>
</tr>
<tr>
<td>Cluster-1</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>Satisfied</td>
</tr>
<tr>
<td>Cluster-2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>Poor</td>
</tr>
<tr>
<td>Cluster-3</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>Average</td>
</tr>
<tr>
<td>Cluster-4</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>VeryGood</td>
</tr>
</tbody>
</table>

Table 4: Error-Accuracy by Classification for Cluster data

<table>
<thead>
<tr>
<th>Classification for Cluster data</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random-Tree (K-Means)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Random-Tree (DBSCAN)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
V. CONCLUSION

E-learning is a new dimension of education system. The traditional approach in education is monotonous as well as predefined. In this paper the attributes of e-learning are analyzed through the K-means and DBSCAN algorithms for grouping of similar data. The Random Tree classification algorithm is applied over the clustered data for predicting E-Learning quality of the students. The “E-Learning Quality” attribute is considered as the class label for the classification. The outcome of the classification algorithm generated the tree for the clustered data. From the tree generated for K-means data the e-learning qualities are classified as satisfied, poor, Average, good and Very good. Similarly for DBSCAN the e-learning qualities are predicted as Good, Satisfied, poor, Average and Very Good. It substantiates that the essential care and importance should be given to the e-learning students in accordance with their categories.

REFERENCES


Authors’ Biography

N. MENAGA is an M.Phil Research Scholar in PG & Research Department Of Computer Science, Dorasingam Government Arts College, Sivaganga, Tamilnadu, India. Her research concentration in Data mining, Machine Learning and its applications. She published paper in international journal.

Dr. B. UMADEVI has received her Doctoral degree in Computer Science from Manonmaniam Sundaranar University, Tirunelveli, India. Currently working as Assistant Professor & Head- P.G and Department of Computer Science, Raja Dorasingam Government Arts College, Sivangangai-Tamilnadu, India. She has over 23 years of Teaching Experience and published her research papers in International, National Journals and Conferences. Her research interests include Data Mining, Soft Computing and Evolutionary Computing. She got the Best Paper Award for her publication in the International Conference on Computational Intelligence and Computing Research held on 27th Dec 2013 at VICKRAM College of Engineering and Technology.