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

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

Mazen Qteishat, Jafar Alqatawna and Mohammad Al-Maaitah examined the factors contributing to attitudes towards E-Learning in higher education among students in Jordan. The research developed a TAM-EL (Technology Acceptance Model for E-learning) for predicting the intention to adopt E-Learning using the constructs of the [9] Technology Acceptance Model.

A.S. Arunachalam, and T.Velmurugan focused on collecting different samples and queries from different colleges who uses ELearning tools 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], Naïve 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 centriods. Then assign each object to the group that has closest centroid. Once all the objects are assigned it recalculate positions of centriods [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.

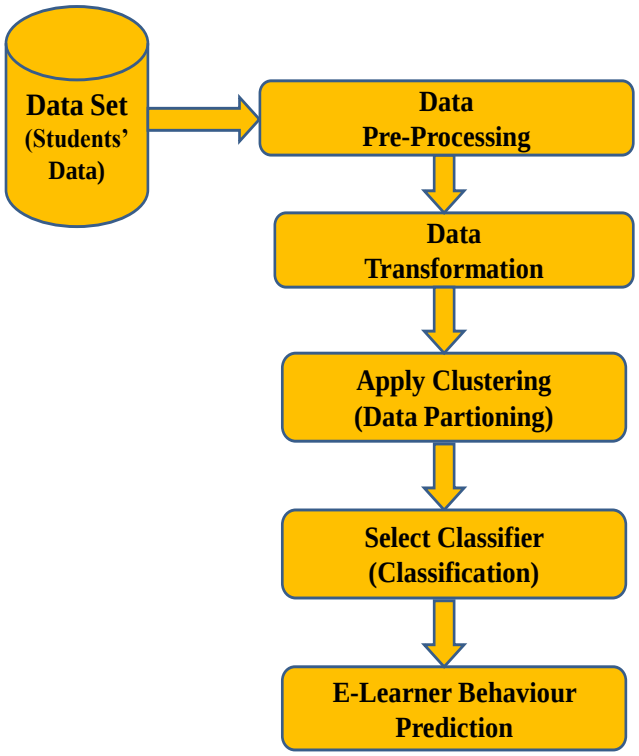

Figure 1: Steps involved in Predicting the Behaviour of E-Learner

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 partioning 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 anlaysed 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 1 : Attribute Description

Attribute -Names	Type	Possible-Values
Collaboration	Number	1 - Brainstorm 2 - groups-of-students 3 - Real-Time-Students 4 - Online-Chatting
Connectivity	Number	1 -web-browser 2 - Chromebooks 3 - iOS 4 - Android devices
Personalisation	Number	1 - laptop 2 - desktop 3 - Tablet 4 - Smartphone 5 - Read 6 - Listen 7 - Watch

Interactivity	Number	1 - Passive 2 - Limited 3 - Moderate 4 - Full
Extended-Opportunities	Number	1 - Apprenticeships 2 - Community service 3 - Independent study 4 - Online courses 5 - Internships
Flexibility	Number	1 - Andragogy 2 - Technological 3 - Interface Design 4 - Evaluation 5 - Resource 6 - Institutional
Motivation	Number	1 - Say Less 2 - Increase The Challenge 3 - Delay Judgment 4 - Appeal To Emotion 5 - Change The Difficulty 6 - Learners Take Control
E-learning -Quality	Nominal	Good, VeryGood, Average, Satisfied, Poor

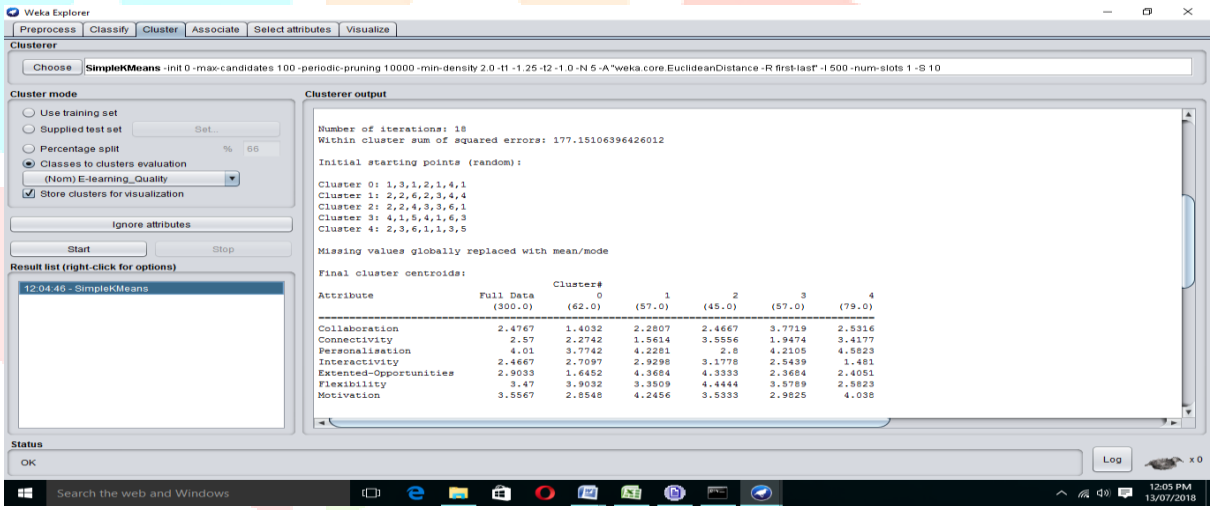

Figure 2 : Clustering by K-Means

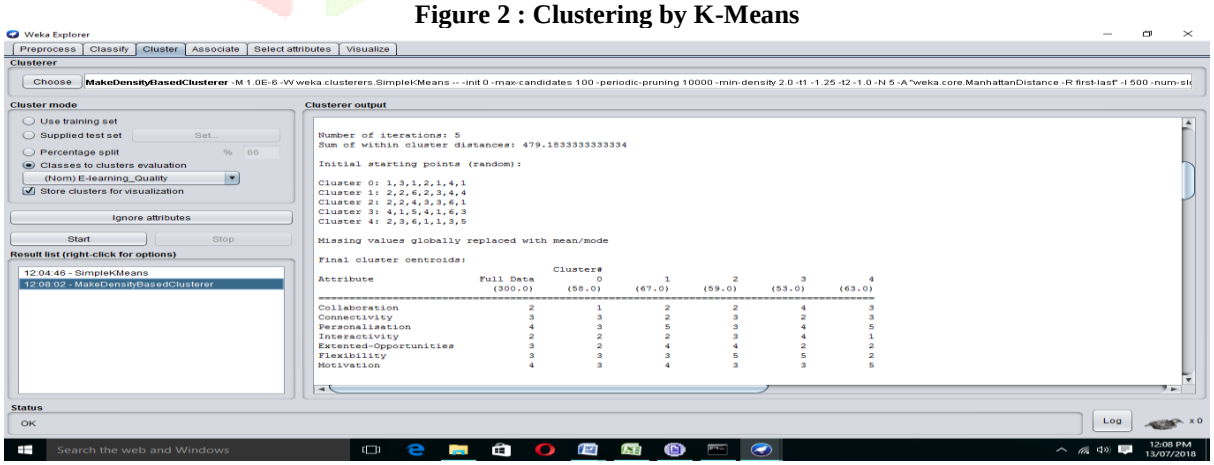

Figure 3 : Clustering by DBSCAN

Table 2 : Clustering by K-Means

Cluster-Name	Collaboration	Connectivity	Personalisation	Interactivity	Extended-Opportunities	Flexibility	Motivation	E-Learning-Quality
Cluster-0	1.4032	2.2742	3.7742	2.7097	1.6452	3.9032	2.8548	Satisfied
Cluster-1	2.2807	1.5614	4.2281	2.9298	4.3684	3.3509	4.2456	Poor

Cluster-2	2.4667	3.5556	2.8	3.1778	4.3333	4.4444	3.5333	Average
Cluster-3	3.7719	1.9474	4.2105	2.5439	2.3684	3.5789	2.9825	Good
Cluster-4	2.5316	3.4177	4.5823	1.481	2.4051	2.5823	4.038	VeryGood

Table 3 : Clustering by DBSCAN

Cluster-Name	Collaboration	Connectivity	Personalisation	Interactivity	Extended-Opportunities	Flexibility	Motivation	E-Learning-Quality
Cluster-0	1	3	3	2	2	3	3	Good
Cluster-1	2	2	5	2	4	3	4	Satisfied
Cluster-2	2	3	3	3	4	5	3	Poor
Cluster-3	4	2	4	4	2	5	3	Average
Cluster-4	3	3	5	1	2	2	5	VeryGood

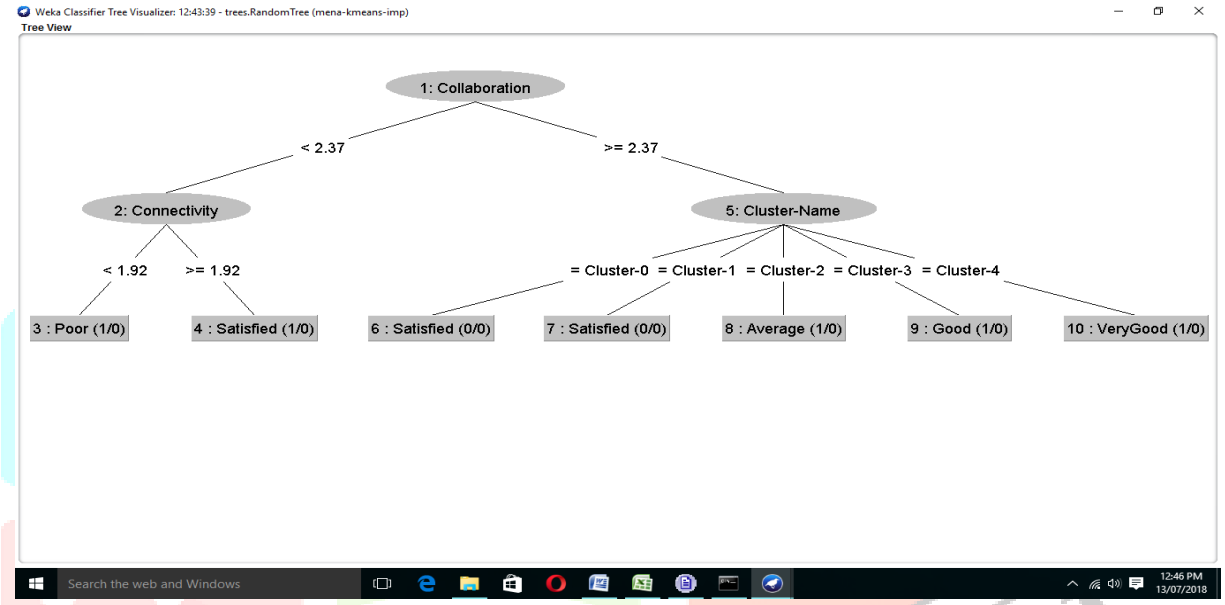

Figure 4 : Predicting E-Learning Behaviour from Random Tree for K-Means clustering data

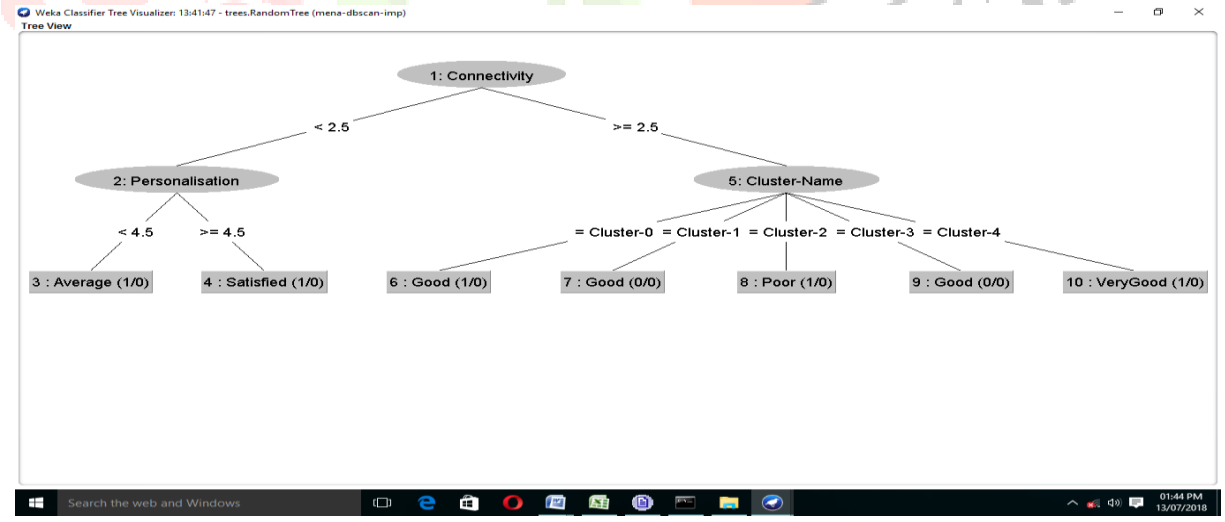

Figure 5 : Predicting E-Learning Behaviour from Random Tree for DBSCAN clustering data

Table 4 : Error-Accuracy by Classification for Cluster data

Classification for- Cluster data	MAE	RMSE
Random-Tree (K-Means)	0	0
Random-Tree (DBSCAN)	0	0

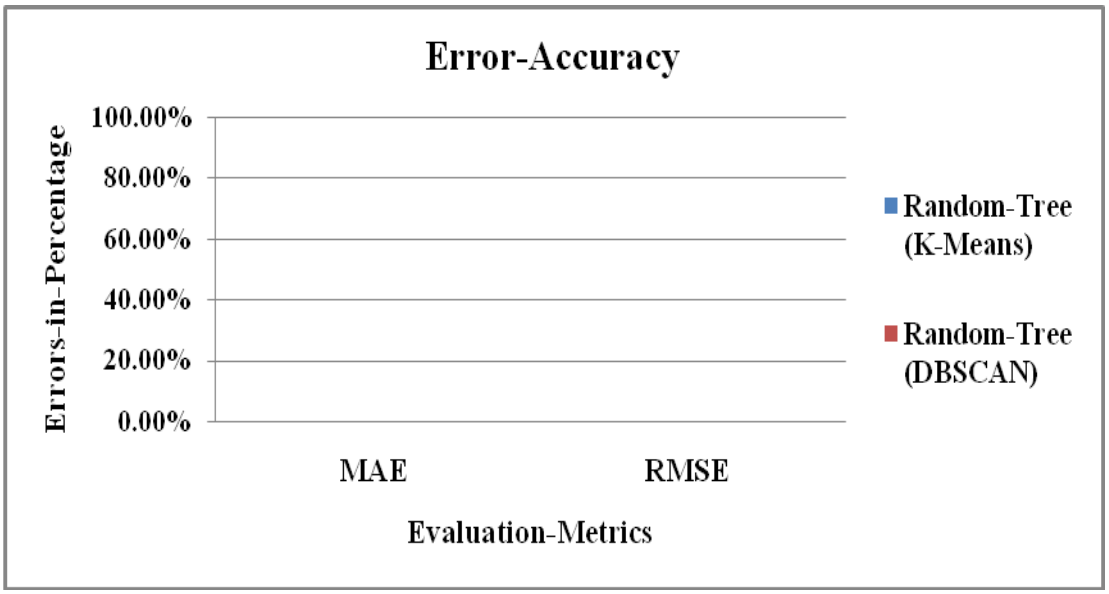

Figure 6 : Error-Accuracy by Classification for Cluster data

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, Satified, poor, Average and VeryGood. It substantiate that the essential care and importance should be given to the e-learning students in accordance with their categories.

REFERENCES

[1] Marie Bienkowski, Mingyu Feng and Barbara Means, "Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief", U.S. Department of Education, October 2012.

[2] Mazen Qteishat, Haitham Alshibly, Ja'far Alqatawna and Mohammad Al-Ma'aitah," Factors Influencing the Adoption of E-Learning in Jordan: an Extended TAM Model", European Journal of Business and Management, ISSN 2222-1905 (Paper) ISSN 2222-2839 (Online) Vol.5, No.18, 2013.

[3] Rahul B. Diwate , Amit Sahu," Data Mining Techniques in Association Rule : A Review", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (1) , 2014, 227 -229.

[4] Syeda Farha Shazmeen, Mirza Mustafa Ali Baig, and M.Reena Pawar," Performance Evaluation of Different Data Mining Classification Algorithm and Predictive Analysis", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p-ISSN: 2278-8727Volume 10, Issue 6 (May. - Jun. 2013), PP 01-06.

[5] Sahil P. Karkhanis, Shweta S. Dumbre," A Study of Application of Data Mining and Analytics in Education Domain", International Journal of Computer Applications (0975 – 8887) Volume 120 – No.22, June 2015.

[6] Dr. N. Venkatesan," Role Of Data Mining Techniques In Educational And E-Learning System", Apjr international journal ,Volume: II, Issue: VII ,August 2013.

[7] Sachin R. Barahate," Educational Data Mining as a Trend of Data Mining in Educational System", International Conference & Workshop on Recent Trends in Technology, Proceedings published in International Journal of Computer,(TCET) 2012.

[8] Shimaa Abd Elkader Abd Elaal,"E-Learning Using Data Mining", Chinese-Egyptian Research Journal, Vol. 5 (1) , 2014, 227 -229.

[9] Sujeet Kumar Sharma, Jyoti Kumar Chandel," Technology acceptance model for the use of learning through websites among students in Oman", International Arab Journal of Information Technology · January 2013.

[10] Davoud Masoumi, Berner Lindström," Quality in e-learning: A framework for promoting and assuring quality in virtual institutions", Journal of Computer Assisted Learning · February 2012.

[11] L.Calvet Liñán, and A. Juan Pérez, " Educational Data Mining and Learning Analytics: differences, similarities, and time evolution" RUSC. Universities and Knowledge Society Journal, (2015). 12(3). pp. 98-112.

[12] Umadevi B. and Sundar D, Alli Dr P., January 2013, "An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50", *International Journal of Data Mining & Knowledge Management Process (IJDKP),*Vol.3.

[13] Umadevi B, Sundar D., Alli Dr.P., 10th June 2014, "Novel Framework For The Portfolio Determination Using PSO Adopted Clustering Technique", *Journal of Theoretical and Applied Information Technology* , Vol. 64- No.1.

[14] Umadevi B.and Sundar D., Alli Dr.P., 2013, "An Optimized Approach to Predict the Stock Market Behavior and Investment Decision Making using Benchmark Algorithms for Naive Investors",*Computational Intelligence and Computing Research (ICCIC), IEEE International Conference*, 26-28 Dec. 2013,Page(s):1 – 5978-1-4799-1594-1INSPEC Accession Number: 14061140.

[15] Birendra Kumar Sahu," K-Means Clustering Algorithm - Case Study
",CTO & Vice President – Technology**,** Published on May 18, 2018.

[16] K.Vanitha, G.Roch Libia Rani," Analysis of Classification and Clustering Algorithms using Weka For Banking Data", International Journal of Advanced Research in Computer Science, 1 (4), Nov. –Dec, 2010, 104-107.

[17] Dr.R.Geetha Ramani ,G. Sivagami and Shomona Gracia Jacob," Feature Relevance Analysis and Classification of ParkinsonDisease Tele-Monitoring Data Through Data MiningTechniques", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 3, March 2012ISSN: 2277 128X.

[18] N. Menaka, Dr. B. Umadevi," A Perspective Review on E-Learning Promotions and Development", International Journal of Computer Science and Mobile Applications,Vol.6 Issue. 6, June- 2018, pg. 16-22.

[19] Dr. B. Umadevi, and R. Dhanalakshmi, "A Comprehensive Survey of Students Performance Using Various Data Mining Techniques". International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064. Volume 6 Issue 4. April 2017.

Authors' Biography

N. MENAGA is an M.Phil Research Scholar in PG & Research Department Of Computer Science, Doraisingam 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 Doraisingam Government Arts College, Sivagangai-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.