



DISCOVERING THE FACTORS AFFECTING THE PERFORMANCE OF E-LEARNING PLATFORM FOR IRAQI POSTGRADUATE STUDIES USING MLP-ANN ALGORITHM

Nazhat Saed
Abdulrazak

University of Technology, Baghdad, Iraq

Maryam Abdul-
Salam Ali*

Iraqi Informatics Institute for Higher Education Studies, Baghdad, Iraq
*Corresponding Author

ABSTRACT

There is no doubt that higher education is a key aspect in the development of educational reality in Iraq, and the focus of Iraqi universities and higher education institutions now directed towards developing their scientific level to improve learning environments for their students. Through modern information technology, we have the potential to create intelligent electronic learning platforms that have many benefits in developing students' scientific level. Data mining technique is used for this purpose. Educational data mining (EDM) is now applied to extract hidden knowledge through a raw data repository from an online learning environment such as a student's demographic data. For instance, age, country, university, previous occupation, etc. And also by analyzing the behavior of students within the e-learning platform and their achievement and grades in the various proposed courses, to assess the impact of these factors on the performance of the e-learning platform in general to determine the effectiveness of the delivery of information to the students. In this research, we will use multilayer perceptron (MLP) algorithm to evaluate the proposed e-learning platform performance for post-graduate studies to predict the most influencing factors in evaluating e-learning platform determine the accuracy of the proposed algorithm, as well as to detect the most influenced factors that affect the results obtained from the proposed algorithm.

KEYWORDS : Higher education, information technology, educational data mining online learning, multilayer perceptron algorithm.

INTRODUCTION

E-learning platforms are now, thanks to the age of technology, a significant tool to develop the educational reality, many universities and scientific institutes are in constant competition to develop their scientific level and to introduce all the modern amenities in various aspects of education to raise the quality and efficiency of learning. E-learning platforms are widely available, but these platforms only deliver the material to the students with some possibilities. There is no continuous monitoring of the behavior of each student on the platform. On the other hand, there is no actual evaluation of these platforms in order to improve their level. How to predict system performance is always a question concerned by faculty members and platform's supervisors. Based on a previous examination of student's results and achievements, it is possible to forecast a future plan to develop the existing e-learning platforms. Nowadays Data mining techniques have been used in a large educational domain. Educational Data mining (EDM) on the other hand, has the focus its the domain of interest towards the educational dataset, it is useful to elicit an unknown knowledge from previously known dataset repository to create an intelligent knowledge representation process. Institutions of higher education have already evaluate their students through their performance and always tries to push the education reality as modern as possible to cope with the information technology process by building many electronic learning platforms to make their student's capable of keeping base with the modern era of technology, This research mainly focuses on studying the different indicators that affect the proposed e-learning platform by exploring student's enrollment data. The data repository is used to create an intelligent decision making and to extract hidden knowledge that is utilized to predict the proposed platform's performance to develop these platforms.

This research offering an electronic platform for students of higher education that provides selective courses of the Department of Computer Science and through their behavior in the platform and their grades in electronic tests we assess the effectiveness and efficiency of the proposed platform in the development of the proposed scientific environment through pre-stored, pre-possessed data, and use these data into

multilayer perceptron algorithm(MLP) to predict the performance of the e-learning platform provided to the students by analyzing the most significant factors affecting the development of the proposed e-learning platform through the behavior of each student inside it. Section 2, concerning with relative previous studies. Section 3, offers a general overview about data mining techniques, artificial neural networks, and MLP algorithm. Section 4, describes the proposed practical work. Section 5, discusses the extracted results and experiences. Section 6 offers a conclusion and the suggested future work.

2 Previous studies

N. V. K. et al., 2017 proposed a new method to assess student achievement. A clustering data mining method is employed to investigate the extensive student dataset. A clustering technique will heighten the searching method pace and examination. The researchers found that assessing student achievement would improve the student occupations, and support the instructors and universities to use modern education systems or changes in the existing ones. **V. Sarala et al., 2015** discuss the applications of data mining in educational institution to extract useful information from the huge data sets and providing analytical tool to view and use this information for decision making processes by taking real life examples. **Jasvinder Kumar, 2015** Suggested a Comprehensive research by utilizing educational data mining technique supports to acquire and improve standards for the extension of the learning situation. It presents decision makers a bigger knowledge of student education and the background learning environment. **Jai Ruby & K. David, 2015** offered a comparable study form to predict the efficiency of the educational achievement of the students applying a Multi-Layer Perceptron method. The analysis confirmed that the characteristics distinguished in the research are effectively high influencing representatives in predicting learner achievement. This type of study is essentially concentrated on examining the prediction efficiency of the educational achievement of the learners applying only varying factors by using Multi-Layer Perceptron model and to match it with the prediction efficiency of the academic achievement of the learners by applying a dataset that includes of all

educational, individual and financial attributes of the students. **Jai Ruby & K. David, 2014** Predicting the Achievement of learners in the Higher Education sector by utilizing data Mining Classification Algorithms using the learners information holding only the influencing factors and confirmed that multilayer perceptron algorithm (MLP), gives the best prediction. **A.FEI Gamal, 2013** This study suggested an educational data mining pattern for predicting learner achievement especially in programming courses. The suggested pattern involves three stages; data pre-processing, selection of attributes and extraction of rules. **Thilina Ranbaduge, 2013** shed the light on the application of various data mining methods in the educational data to recognize the extensive dataset on student education which can be employed to assess the student's overall achievements in the electronic learning environment upon systems and learn how these achievements are managed to deliver out diverse education models of the learners. **Annabel Latham et al., 2013** This study suggests a novel approach for profiling student learning methods for a conversational intelligent tutoring system (CITS) which employs a Multilayer Perceptron (MLP) algorithm. The record represents analyses carried with actual students in an active educational environment for profiling two dimensions Felder and Silverman styles. The decisions show that (MLP) can predict education techniques with an efficiency of 84-89%.

T. Chellatamilan et al., 2011 aims to increase the education method evolving, learning content planning and learning objects suggestions based on the immediate data gathered within the e-learning platforms. For instance, Moodle. This paper applies the methods like data classification and clustering systems to predict the education style of the match learners based on their behaviors.

3 Data Mining (DM)

In current years, the vast growth of artificial intelligence (AI) and deep learning presented different method for intelligent prediction and classification. Students' performance in an electronic learning platform is the main interest for educators and supervisors [10]. student's and platform's performance can be predicted through the use of data mining techniques or known as (DM), The definition of (DM) can be viewed as "the non-nugatory method of classifying valid, novel, previously unknown, potentially utilizable and understandable patterns from data in database. DM is generally have been used in commercial fields. Many researchers stated these days, there are a huge focus in applying the applications of data mining techniques in the education sectors [1 1].

There are several recognized (DM) algorithms classification like decision trees, Bayes nets, and MLP which are applied to predict the achievements of the students [12]. The research is largely focus on how student's learning styles can affect the performance of the proposed e-learning platform. By extract a hidden knowledge behind the student's education behaviors inside the platform to predict the performance of the system as well as to make an evaluation on how the prediction can be accurate in the field of Educational Data mining (EDM).

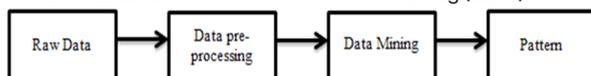


Figure 1. Extracting patterns from data mining process [12]

3.1 Educational data mining (EDM)

EDM concerns with the utilizing of data mining algorithms in the educational fields to understand and overcome many obstacles in the educational sectors especially in an online-learning platforms, in order to develop and upgrade the quality of learning environments for both learners and faculty members [13].

Educational data mining technique (EDM) tries to convert raw data in knowledge discovery through data analysis and processing. EDM, involves 4 steps: data preparation, data mining and data analysis then data evaluation., this is a series of learning knowledge from data produced with an educational background and reusing it to enhance the educational environment.

3.2 EDM process

Educational Data Mining is a growing discipline needs educational environments. For example, e-learning platforms that has capabilities and the needed methods to explore a raw dataset from the student's usage, in order to process these data to create a prediction and perform a better understanding of the learning patterns.

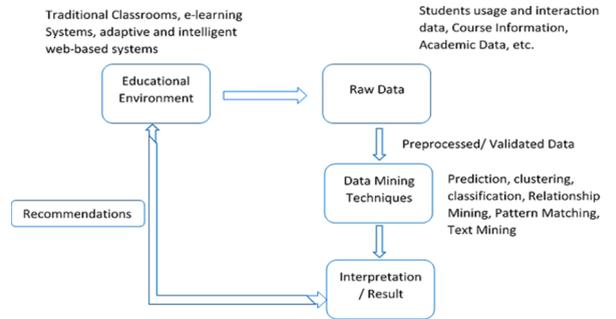


Figure 2. EDM process [11]

In order to create an educational data mining process we need three types of users 1- a students to interact with the proposed platforms, 2- an educator or supervisor to design and plan how the scientific materials can be delivered to the students through there platforms, 3-an educational environment to be used by both students and faculty members to exchange the academic courses.

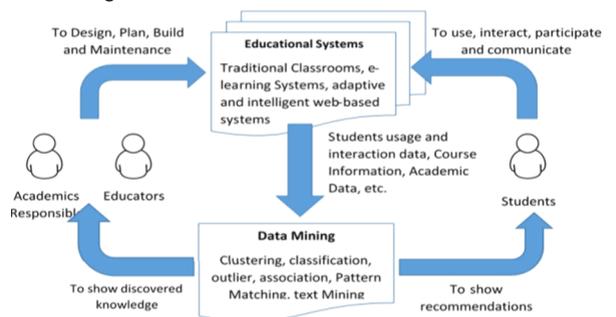


Figure 3. Cycle of EDM process [11]

3.3 Artificial Neural Network (ANN)

Artificial neural network techniques appeared as method employed in educational data mining (EDM). The power of ANN is that it is capable of finding all the similarities among set of different variables, the advantage of available computing power today has led to producing an extra layers of the neural network that can be employed, and deep learning interpretation can be effectively performed [10].

The idea of artificial neural network has inspired by the biology of human brain model, human brain has a great inner computing nodes which referred to as neurons that can performed a complex processes. Therefore, ANNs can perform non-linear, highly complex interactions among set of inputs and outputs, also ANN have the ability to learn from a previous known data, so it can be easily trained by employing supervised learning model in order to create a classification from previously un-classified set of data [8]. ANN has a couple of basic elements named processing elements (neurons), and the associations among them (links). These links have special

parameters associated with it these parameters refer to as (weights). The artificial network of neural has several layers (input layers, hidden layers, and output layers). Hidden layers are considered as the layers that can employ non-linear processing elements. Input values are forwarded through input layers to hidden layers to be processed when they are carried through the hidden layers to either another hidden layer or to the output layers

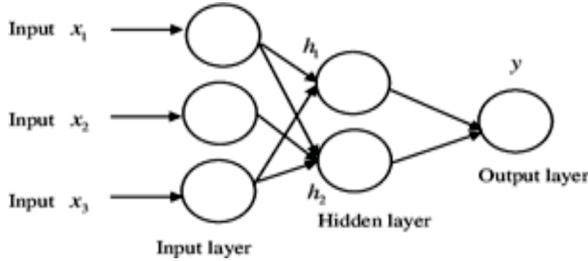


Figure 4. ANN [4]

3.4 Multi-Layer Perceptron (MLP)

MLP – Multi-Layer Perceptron method considered as one of the most common extensively employed method in a general supervised learning artificial neural network algorithm. Multilayer Perceptron is a feed-forward artificial neural network model maps collections of input data in a number of satisfactory outputs. (MLP) has multiple layers of neurons where every layer is entirely related to the next one that can classify many data values in non-linear matter. These layers can learn how to map and direct input data values into a required form of responses. Therefore, they are largely employed for solving many classification and prediction problems [4].

MLP-ANNs have been strongly utilized in the prediction of breast cancer, students' academic performance, modeling student maintenance in science and engineering disciplines, and the prediction of student class choice in electronic higher education institutes [8].

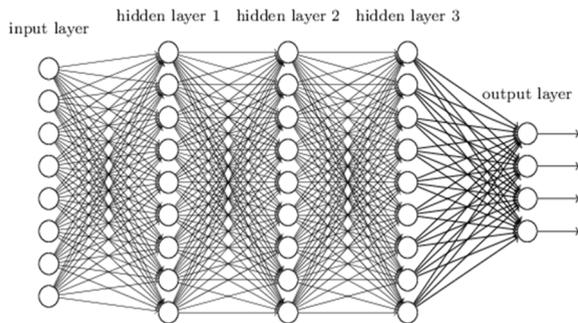


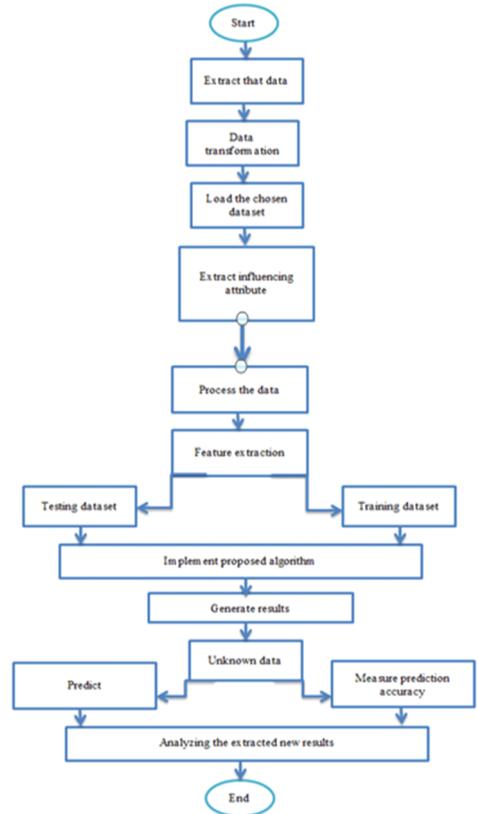
Figure 5. MLP

4. Research methodology

In this paper there are three main phases. First one, focusing on designing and implementing an electronic learning platform. Second one is about collecting the required data. Third one considers MLP algorithm to evaluate the overall performance of the proposed e-learning platform based on students' behavior inside the platform and their grades, to extract knowledge about the most influencing factors on the overall evaluation for future development.

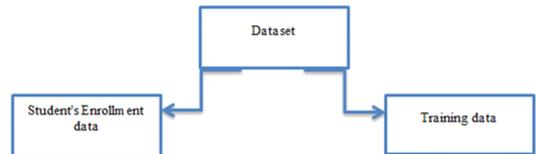
4.1 First Phase

This phase considers the design and implementation of the proposed e-learning platform, the following chart illustrates the steps of designing the platform:



Flowchart 1. Design and implementation of the proposed e-learning platform

4.2 Collecting dataset: in this phase two kind of dataset have been collected.



Flowchart 2. Data Collection

First set: involves collecting data about the courses that will be taken in the platform, there are six main courses (Artificial intelligence, Software engineering, and Database, Networking, Operating system, and Data security). These courses will be chosen selectively by each student, and then there is one more additional course which is English course, for those who wants to strength their English language. The following table shows the student's enrollment data in details:

Table 1. Student's enrollment details

Name	Type	Description
Student ID	Input attribute	Student's profile number
First name	Input attribute	Student's first name
Last name	Input attribute	Student's last name
Gender	Input attribute	Student's gender male/female
Mobile	Input attribute	Student's mobile number
E-mail	Input attribute	Student's or faculty e-mail
Password	Input attribute	Student's encrypted password
University	Input attribute	Student's university that uses this LMS
Country	Input attribute	Student's country
Pre-test	Input attribute	Pre-test score of 100/100
Level 1 quiz	Input attribute	Level one quiz score
Level 2 mid-test	Input attribute	Level two middle test

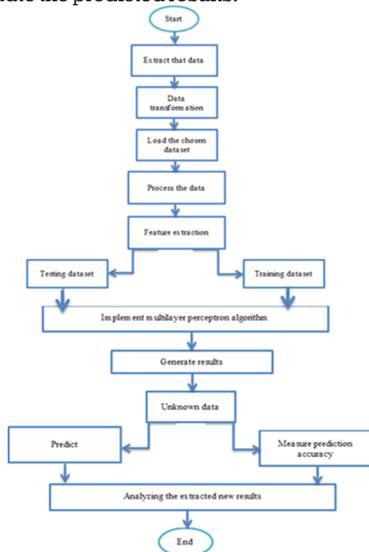
Level 2 final-test	Input attribute	Level three final test
Final score	Input attribute	Final score of the fourth above scores
Pass-fail	Input attribute	Either pass or fail the course
Certification	Input attribute	Either took certification or not
Open-library	Input attribute	Open platform library page
Take-e-course	Input attribute	Take the English extra course
Enter chat	Input attribute	Either enter chat or not
Number of materials downloaded	Input attribute	Number of books downloaded from library page
Number of levels	Input attribute	Number of passed levels
System performance	Input attribute	System self-evaluation

Second set: involves collecting previous data about the student's (demographic information and grades). These data is automatically generated and it is extracted online as a synthetic dataset and then it was transformed to be more suitable for our proposed platform, the data of 1000 students were collected during a three educational years, it is important to train the MLP algorithm in order to test it later, after completing the course by each student, to extract knowledge about the most important factors that influencing the classifier and its accuracy.

4.3 MLP algorithm

Several factors affect the educational achievement of the learners. The research pattern is essentially directed on examining the prediction efficiency of the educational achievement of the learners utilizing uniquely affecting factors with Multi-Layer Perceptron algorithm, utilizing a dataset that includes of all educational and individual factors of the learners, these data is then fitted into a distinct table in order to be scanned by the proposed algorithm. Student profiling is necessary to the progress of 'one course fits all students' during our technological era computer based-learning is transformed into more intelligently educational methods that is suitable for each student.

- Step 1:** Student's dataset is scanned.
- Step 2:** Define the values [m, n, p]
- Where:** m is for training dataset, n is for testing dataset, and p is for testing the accuracy of the classifier
- Step 3:** Calculate the weight of each inserted attribute value.
- Step 4:** Apply the proposed algorithm.
- Step 5:** Evaluate the predicted results.



Flowchart 3. MLP algorithm

5. Experiments and results

5.1 Software Used

The software that is used to train and test MLP algorithm is called WEKA (Waikato Environment for Knowledge Analysis) is a common scale of machine learning software recorded in Java following GNU General Public License, formed at the University of Waikato, New Zealand. The Weka workbench includes a number of visualization accessories and algorithms for information analysis and predictive modeling, mutually with graphical user interfaces for clear access to this functionality.



Figure 6. Weka Software

5.2 Results

Training dataset (MLP accuracy)

In this model we divide the collected data into 75% of the collected students' data (demographic and grades) to train the prediction accuracy of the algorithm, the results shows that the correctly classified data according to system performance field (which is the field that is chosen to be predicted) is approximately 35%, and the incorrectly classified dataset according to the same field is approximately 64%, time taken to build the model is 0.01 second. Figure below shows the results on GUI of WEKA software.

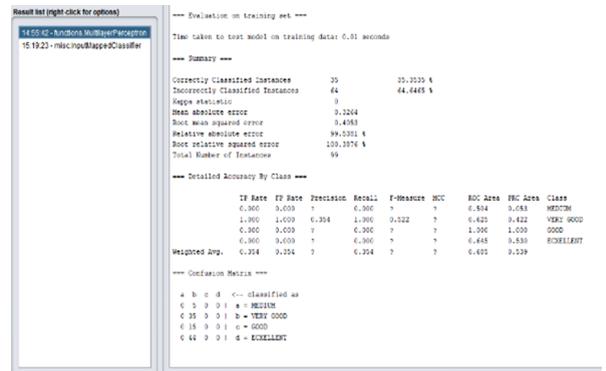


Figure 7. Training dataset

Testing data set (MLP accuracy)

In this model we use 25% of the collected students' data (demographic and grades) for testing the prediction accuracy of the classifier, system performance field was also chosen to build a prediction according to its values, correctly classified data was about 55%, incorrectly classified data was about 44%, and the time taken to build the model is 0.03 seconds. Figure below illustrates the obtained results.

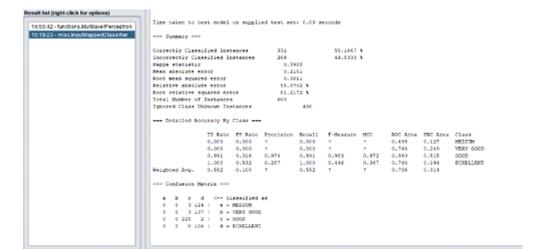


Figure 8. Testing dataset

Final obtained results between training and testing dataset of MLP algorithm shows that after the model has learned from previously known data, it was then tested after hiding the known data in order to extract knowledge and to predict the unknown values.

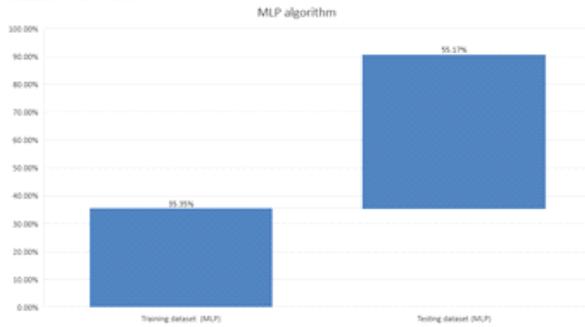


Figure 9. Obtained results

Table 2. Influenced factors in platform evaluation

Attribute	Details
Pre-test	Pre-test score of 100/100
Level 1 quiz	Level one quiz score
Level 2 mid-test	Level two middle test
Level 2 final-test	Level three final test
Final score	Final score of the fourth above scores
Pass-fail	Either pass or fail the course
Certification	Either took certification or not
Open-library	Open platform library page
Take-e-course	Take the English extra course
Enter chat	Either enter chat or not

In addition to measure the accuracy of MLP algorithm prediction, in this paper we are also want to know the most influencing input data values in each student's record the results show that from a 22 columns in the database only 10 columns considers the most influenced columns in each student's course as well as the overall platform performance measure. The table below shows the most influenced columns in the platform database.

6. CONCLUSION

Lately, Educational Data Mining (EDM) has been used as a growing field of data mining and educational methods. EDM employs common mining algorithms to summaries informative data to learn and develop the students' learning process. This paper studies the most influenced factors to evaluated the performance of the proposed e-learning platform that have been designed for higher education studies based on a previous collected student's dataset. Synthetic data of a three educational years with 1000 row each have been generated to train the model, a 75% of the data have been utilized for that purpose, and 25% kept to be used later to test the accuracy of the propped algorithm. A 10 folds cross validation have been generated, and WEKA software is used to train and test the algorithm, the obtained results from training dataset were 35%, and the results after testing the MLP model were 55% which is a promising results. This research shows the characteristics extracted from the dataset are very high influence by utilizing MLP. This paper supports the institutions to better understand the educational situation of each learners in the future and they can focus the student's with weak performance to strength their weakness later and to develop their academic achievement. In a future work this study can be taken by merging two or more new algorithms for more accurate prediction or a different algorithm can be extended for larger classification and prediction utilizing the important influence extracted by EDM.

REFERENCES

1. N. V. K. Rao, N. Mangathayaru, and M. S. Rao, "Evolution and prediction of

radical multi-dimensional e-learning system with cluster based data mining techniques," in 2017 International Conference on Trends in Electronics and Informatics (ICTEI), 2017, pp. 701-707.

2. V. Sarala, Dr. v.v. Jaya Rama krishnaiah, "empirical study of data mining techniques in education system", international journal of advances in computer science and technology (ijacst), vol. 4 no.1, pages: 15-21 2015

3. Jasvinder Kumar, a Comprehensive Study of Educational Data Mining, International Journal of Electrical Electronics & Computer Science Engineering Special Issue - TeLMISR 2015, ISSN: 2348-2273.

4. Jai Ruby & K. David, "Analysis of Influencing Factors in Predicting Students Performance Using MLP - A Comparative Study ", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 2, February 2015.

5. Jai Ruby & K. David, "Predicting the Performance of Students in Higher Education Using Data Mining Classification Algorithms - A Case Study", IJRASET International Journal for Research in Applied Science & Engineering Technology, Volume 2 Issue XI, November 2014

6. A.FEIGamal, "An Educational Data Mining Model for Predicting Student Performance in Programming Course", International Journal of Computer Applications (0975-8887) Volume 70-No.17, May 2013

7. Thilina Ranbaduge, "Use of Data Mining Methodologies in Evaluating Educational Data", International Journal of Scientific and Research Publications, ISSN 2250-3153, Volume 3, Issue 11, November 2013

8. Annabel Latham, Keeley Crockett, David Mclean The Intelligent Systems Group, School of Computing, Mathematics and Digital Technol "Profiling Student Learning Styles with Multilayer Perceptron Neural Networks", The Manchester Metropolitan University 2013

9. T. Chellatamilan, M. Ravichandran, R. M. Suresh, and G. Kulanthaivel, "Effect of mining educational data to improve adaptation of learning in e-learning system," in International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2011), 2011, pp. 922-927.

10. Wilton W.T. Fok; Y.S. He; H.H. Au Yeung; K.Y. Law; KH Cheung; YY. Ai; P. Ho, "Prediction model for students' future development by deep learning and tensorflow artificial intelligence engine - IEEE Conference Publication", International Conference on Information Management (ICIM-2018)

11. Anoopkumar M; A. M. J. Md. Zubair Rahman, "A Review on Data Mining techniques and factors used in Educational Data Mining to predict student amelioration - IEEE Conference Publication," International Conference on Data Mining and Advanced Computing (SAPIENCE-2016)

12. JR K David "An Analysis on Academic Performance of Students using a Hybrid Model for Higher Education", International Journal of Engineering and Technology (IJET) 9 (3), 2175-2182 (2017)

13. S. Q. Wei Zhang, "A brief analysis of the key technologies and applications of educational data mining on online learning platform," IEEE International Conference on Big Data Analysis (ICBDA-2018)