Evaluation of Student Performance in Educational Data Mining Using Hybrid Deep Learning Technique

Jaikumar M. Patil [1], Dr. Sunil R. Gupta [2]

[1]Shri Sant Gajanan Maharaj College of Engineering, Shegaon, Maharashtra, India
[2]Prof. Ram Meghe Institute of Technology & Research, Badnera, Maharashtra, India
Corresponding Author: jaimpatil1011@gmail.com

Abstract: Educational failure is prevalent. The surge in the number of students who quit school has multiple root factors. The inability to succeed academically is a primary factor in why students drop out of school. Since many students struggle to adjust to their new school, this affects performance. Our study aims to identify all elements affecting undergraduate academic achievement. Thus, this initiative aims to help students identify the factors that lead to their successes so they can take steps to change their results. Students, course instructors, and others can improve the environment after identifying and assessing its main components. We used a Recurrent Neural Network and Long Short-Term Memory classification technique to forecast student academic success early. This method is compared to numerous machine learning classifiers and a deep learning classification model. Using study findings from numerous trials, we examined the classification performance of several standard machine learning techniques, such as support vector machine, random forest, J48, artificial neural network, and naive bayes, as well as deep learning models, such as RNN. RNN-LSTM sigmoid, Tan – h, and ReLU function are used to predict student performance and enhance teaching. The results are compared to deep learning and machine learning methods. RNN-LSTM (ReLU) has the highest accuracy rate of 97%, as per experiments. Our technique has great classification accuracy on different datasets or real-time complex huge datasets of students with multivalued variables.

Keywords: Educational Data Mining, RNN-LSTM, Student Performance, Deep Learning, Machine Learning.

DOI Number: 10.48047/nq.2022.20.19.NQ99163 NeuroQuantology2022; 20(19): 1877-1892

1. Introduction

The desire to obtain the proper education is developing as a result of the world becoming more competitive on a daily basis. The most significant difficulty currently facing the education sector is assisting universities and other educational institutions in developing teaching methodologies that are more effective, pertinent, and accurate. Data Mining (DM) [1, 2] is regarded as the most perfect system for delivering more data to the educator, pupils in the high school graduating class, the administrator, and other staff members because it is capable of obtaining a vast amount of information [3, 4, 5, 6]. This is due to the fact that it is regarded as the most effective option. One of the elements of the educational system that is seen as being of utmost importance is higher education. The organizational structure of higher education places a strong emphasis on the ongoing evaluation of teaching. Advanced educational methodologies may be able to fill knowledge gaps with the growth of data mining [7]. The application of data mining techniques can enhance the procedures' thoroughness, effectiveness, and speed. These improvements have the potential to improve the effectiveness of the inclusion program, lower the number of pupils who drop out of school, increase the rate of progression, improve the level of uniformity, increase the percentage of students who switch classes, increase the percentage of teachers who adopt new teaching strategies, increase the number of students who learn, and then lower the amount of money spent on education. To achieve the aforementioned grade transformation, we require a data mining technique that can impart the necessary skills and information for improved higher education design [8, 9, 10, 11].

In the research area known as "Educational Data Mining" (EDM) [12, 13, 14], methods including data mining, machine learning, and statistics are used to data generated in academic contexts (such as universities and colleges). The goal of education data mining is to use data mining techniques to address problems that develop in the field of education [15, 16]. In order to enhance and more precisely forecast academic accomplishment, EDM is regularly used with a wide range of Machine Learning (ML) approaches in the education industry [17, 18, 19, 20]. Machine learning techniques were used in the initial study to provide predictions about student performance using binary categories, such as passing or failing. However, judging students' performance and refining instruction only on whether or not they received a passing grade does not provide a more realistic picture of their level of achievement. Their method does not analyze the whole interplay of the predictor elements present in the student data, which is another significant flaw in their methodology. Traditional machine learning classifiers are unable to generate accurate predictions of student achievement based on academic data. Academic institutions have recently used supervised learning techniques to help them automatically extract high-level features from raw data. RNN-LSTM, also known as Long Short-Term Memory with Recurrent Neural Network, is used in this study to predict student success based on real-time big heterogeneous datasets. A comparison between Deep Learning classifiers and traditional machine learning techniques is also carried out using the suggested RNN-LSTM.

The study is divided into multiple parts; part II provides a detailed explanation of all machine learning and deep learning approaches. Part III describes the relevant work of many researchers. The proposed study approach is thoroughly covered in part IV. Part V describes in detail the experimental findings and discussion, and Part VI concludes up the system.

2. Background

Several machine learning and hybrid deep learning methods have improved teaching and predicted student achievement [21, 22, 23, 24, 25, 26]. This section describes all techniques.

2.1 Support Vector Machine

This approach solves categorization and regression problems. Support vector machines are increasingly useful for massive data classification. SVM finds the best decision boundary to divide an n-dimensional space into categories. This will allow long-term data categorization. Hyperplane might mean optimum choice boundary. It can also classify linear data. A non-linear mapping expands training data into more dimensions. Each student is unique; thus we consider them as multidimensional items. This additional dimension searches for the linear-optimal hyper-plane, or "decision border," that divides students into two groups. Hyperplanes can always divide data from two groups (H1 and H2). Edges (Big and Small) and support vectors help the SVM find this hyperplane (also known as "important" processing data sets). Despite its gradual training stage, this method provides great accuracy, especially for relatively tiny quantities of support vectors that are unaffected by item dimension, according to data analysts. Despite its efficacy, the support vector machine excels in categorizing a restricted training sample set with many variables. A non-linear method can be used to compare each student's parameters to others to predict their accomplishment group. This approach avoids overfitting better than others.

2.2 Random Forest

Supervised learning includes Random Forest, classification and regression problems can be solved with it. Ensemble techniques combine many classification methods to solve difficult problems and increase the model's performance. It estimates the average of several Decision trees (DT) applied to subsets of the dataset. This improves dataset predictions. The random forest model uses all tree forecasts rather than a single DT [27]. The majority of forecasts' votes determine the model's prediction. More trees reduce high precision and overfitting mistakes.

2.3 Artificial Neural Network

An Artificial Neural Network has weighted linkages connecting input and output units (ANN). Adjusting connection weights helps the ANN estimate the correct target label for specific input data scenarios. Back-propagation is a popular method for training artificial neural networks. The artificial neural network is used when the class label and dataset attributes are unknown due to its many benefits, such as its strong tolerance to inaccurate information and its ability to categorize sequences it has not been trained on. The ANN can classify sequences it hasn't been trained on. ANNs are adaptive systems because they can modify their structure in response to the information that goes through their network while learning, whether it comes from inside or outside. Numerical quantity-based non-linear modeling uses neural networks. Prediction and handwriting recognition are two applications of artificial neural networks. Neural networks have layers. Neuronal layers serve similar activities. These three categories describe layers. Input neurons receive user program input. Output neurons form the output layer. Hidden layers are between input and output. The hidden layer neurons only have network connections, not user program connections. Hidden layers are optional. This structure needs input/output layers.

2.4 Naive Bayes

The Naive Bayes classifier is widely regarded as the simplest basic type of Bayesian network. Given the objective feature state, it is assumed in this method that there is no relationship between the attributes of any two features. Every item x in the collection is given one of the attribute values, a1–ai. The objective function for each value from the finite set V = (v_1,..., v_j) that has been provided is denoted by f(x). The formula shown below is used by the naive Bayes model [28].

$$V_{max} = \underset{V_j \in V}{Max} \; P(v_j)\, \Pi_i\, P(a_i \backslash v_j) \qquad (1)$$

The model's target, v, can be found in the training dataset by computing P (a_i |v_j) and P (v_j).

2.5 J48

J48 is a free Java implementation of C4.5 DT. J48 extends ID3. J48 also handles missing data, DT pruning, and rule derivation. J48 uses predictive ML to forecast a new instance's value based on numerous characteristic values. The internal nodes of a DT indicate distinctive properties, the branches connecting them identify possible values in the data obtained, and the final nodes indicate the dependent variable's inherent value (classification) [29]. Before employing Data Mining on the database, we set up our study. UCI Ml repository, real-time data sets, and Kaggle data were collected first. Pre-processing eliminates superfluous and imprecise data after data collection. Feature extraction and selection reduce the worst and most repetitive features and choose the most significant properties for model construction. The final step trains and tests the algorithm to better forecast student academic performance.

2.6 Recurrent Neural Network

Recurrent Neural Networks have linkages that form a directed acyclic graph along a sequence. Time series data is its main use. Recurrent neural network uses sequential data. Recurrent networks (RNNs) perform the same task for every sequence element. RNNs have a "memory" that remembers past calculations. Figure 1. illustrates Framework of recurrent neural network.

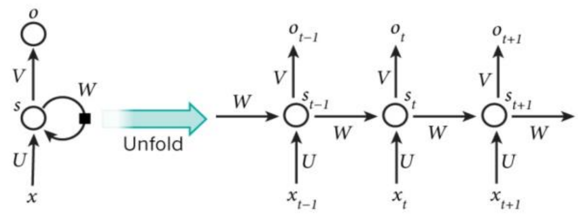

Figure 1. Framework of recurrent neural network

The forward pass of the recurrent neural network is similar to an MLP with one hidden layer, except that activation is transferred to the hidden layer from both the existing external input and the hidden layer activations 1 stage backwards in time. Consider the following equation for concealed unit input.

$$a_h^t = \sum_{i=1}^{I} W_{ih}\, X_i^t + \sum_{h=1}^{H} W_{hh}\, b_h^{t-1} \qquad (2)$$

$$b_h^t = \theta\,(a_h^t) \qquad (3)$$

The output unit's recurrent neural network back propagation is regular. Starting at t = T and applying the following functions, reducing t at each step, yields the whole delta value series.

$$a_k^t = \sum_{h=1}^{H} W_{hk}\, b_h^t \qquad (4)$$

δj T+1= 0, ∀j, because no error is obtained from over the termination of the sequence.

$$\delta_h^t = \theta^j(a_h^t) \sum_{k=1}^{K} \delta_k^t W_{hk} + \sum_{h^{l}=1}^{H} \delta_{h^{l}}^{t+1}\, W_{hh^{l}} \qquad (5)$$

$$\delta_h^t = \frac{\partial O}{\partial a_j^t} \qquad (6)$$

2.7 Long Short-Term Memory

RNNs are a type of long short-term memory. Recurrent neural networks receive information from prior phases. Hochreiter & Schmidhuber established Long Short-Term Memory. It addressed long-term recurrent neural network reliance, where the RNN can forecast words from current data but not from long-term memory. RNN's potency decreases with gap length. Data can be stored in long-term memory. It analyses, forecasts, and classifies time-series data.

LSTM Structure:

The Long Short Term Memory chain structure consists of four NNs (cells) and several memory building elements.

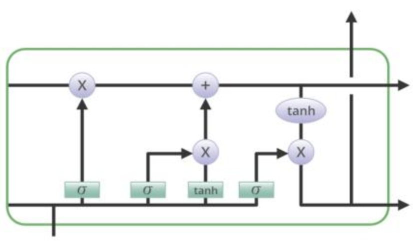

Figure 2. Structure of LSTM

Gates and cells modify and preserve memories. Following are the t Three gates:

- **Forget Gate:** It eliminates cell state data. x t (current input) and h t-1 (previous cell output) are mixed with weight matrix before bias is introduced to the gate. The activation function outputs the outcome in binary form. If a cell state's return is 0, the data is destroyed; if it's 1, it's preserved for later use.

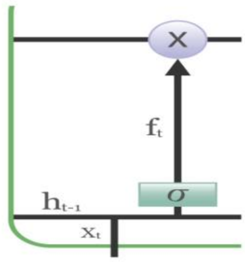

Figure 3. Forget Gate

- **Input Gate:** It adds appropriate data to the cell state. The sigmoid activation function controls and filters data using the inputs h t-1 and x t. The tanh technique provides a vector with every value between h t-1 and x t, spanning from -1 to +1. Multiplying the vector values and controlled values extracts valuable information.

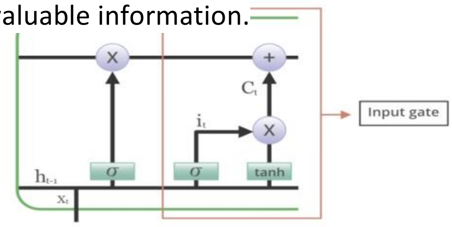

Figure 4. Input Gate

- **Output Gate:** It retrieves and displays cell state info. The cell is vectorized using tan h. The sigmoid activation function controls the data after filtering by the numbers to be retained using h t-1 and x t. Finally, the vector and controlled values are multiplied and delivered as outputs and inputs to the next cell.

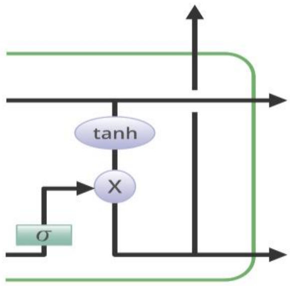

Figure 5. Output Gate

3. Related Work

Because people have different learning styles, Worapat Paireekreng et al. [2] recommend providing the best teaching resources and information. Identifying a pupil's favourite learning technique may be difficult due to limited student data and no learner profile. Educational concept surveys must be completed to determine the learner's preferred learning technique. The learner's academic history and socio-demographic aspects should also be considered. Thus, learning preferences need a better algorithm. This study aims to identify new students' learning styles. An ensemble classifier for learning styles was created. The research showed that ensemble classification strategies outperformed other classification methods. In 2019, Nabila Khodeir et al. [13] reported that EDM analyzes educational data using ML, DM, and statistics. Learner simulation is an EDM program that suggests educational system resources. This study analyzes educational systems. Affective state modeling, academic achievement prediction, and learning style modeling are all explored. Student profiling, categorization, and collaborative analysis are addressed.

According to Sagar deep Roy et al. [28] EDM research uses DM, ML, and statistical tools to understand student learning, predict achievement, and improve learning. This paper summarizes EDM's effects on student achievement evaluation. In 2019, Balqis Al Breiki et al. [10] sought to improve student performance projections to give education system with data required to enhance educational objectives as soon as possible. This research employs regression and ML algorithms to construct learning methods that accurately predict student GPA. According to S. Senthil et al. [34] many DM models can predict students' educational outcomes. This research compares NB, ANN, LR, SVM, Instance-based, DT, and Rule-based classification methods using a dataset from the UCI-ML Repository.

In 2020, Ndiatenda Ndou et al. [30] published a strategy for predicting student achievement per year until graduation for South African higher education students. The latest research applies a variety of classification methods to a synthesized data set produced by a Bayesian classifier to demonstrate that these classification methods can also be used to anticipate student achievement to improve student achievement and avoid the negative effects of students who struggle to finish their education or drop out. Anoop Kumar M et al. [31] analyze educational data to construct models for improving academic standards and institutional efficacy. This paper collects and distributes relevant literature to computer instructors and professional organizations. It finds well-supported studies to improve education and reenergize the school's vulnerable students. These findings illuminate strategies for improving instructional procedures, forecasting student progress, comparing DM accuracy, and developing open-source tools.

In 2018, Edona Doko et al. [32] defined the classic flipped classroom as home-based video tutorials. Nurul Hidayat et al. [33] focus on data-mining student participation in an e-learning system. It used APTIKOM Consortium's second distant learning program's learning process database. This research discovers knowledge structures and reorganizes the online course utilizing frequent patterns and categorization. This study will assist academics simulate the data pre-preparation process using Moodle log data. In 2018, Wei Zhang et al. [34] reported that online learning systems save massive volumes of learner behavioral and academic data. The study opens with BDE, EDM, and virtual learning platform concepts. EDM is then demonstrated. Finally, the main DM approaches are grouped by function and explained in connection to online learning. This article recommends web based EDM study and use.

Vinayak Hegde et al. [35] developed a questionnaire and test approach to identify academically slow learners. Instead of merely looking at internal grades and tests, the suggested technique selects pupils dependent upon a deeper understanding of them. Entropy value proves the method selects better pupils for investigation. Frequent faculty coaching helps students strengthen their weakest ideas and benefit the university. Steven Lehr et al. [36] employ EDM to estimate student retention. The study taught data feature generation, technique selection, and verification using academic data. According to Rashi Bansal et al. [37] in 2017, DM is particularly effective when assessing students' behavior in an online learning environment. DM can directly deconstruct and uncover data, which is tough and laborious without technology. Almost all sectors use a variety of DM tactics and technology to deliver company information and enhance the decision-making procedure. This paper's key aim is to focus on the areas of EDM wherever e-learning can be implemented.

According to Moustafa M. Kurdi et al. [38], educational datasets have too much data to predict and assist student behavior in 2018. Lebanon does not track student progress and development. Educational data is evaluated to determine students' mis behavior and choose relevant remedies and treatment paths. Predicting student success is made easier by adopting DM approaches to boost student performance. It could help students, professors, and large universities succeed academically. This article uses anticipatory computation to discover pupils' most important information (behavior). The strategy could boost student performance and inform students, teachers, and academic institutions. In 2018, Viviana Parraga et al. [39] conducted a Systematic Mapping Study (SMS) on educational information extraction to discover factors affecting higher education academic attainment. 20 key research showed that data mining can predict student achievement and lower dropout rates. In 2020, Aberbach Hicham et al. [40] reported that DM can evaluate and retrieve meaningful data from massive data streams in several application domains. In 2017, Rita Tavares et al. [41] created a digital scientific education resource for elementary schools using EDM. The theoretical model predicts how the instructional technique would affect students' self-regulated learning and scientific competency. Thus, student behavior toward available aid, formative feedback, and suggestions and student research of learning processes will be examined.

Karan Sukhija et al. [42] developed the inter-disciplinary research on EDM, which uses several methods to analyze data from various educational sources. Hanife Goker et al. [43] create a student data warehouse that may be utilized with DM approaches to improve an alert system that predicts students' academic performance for students and parents and identifies the important elements affecting that performance. Mudasir Ashraf et al. [44] find and analyze educational data components that significantly improve student accomplishment. The system uses analysis of variance and Structural Equation Modeling for mining (SEM).

4. Research Methodology

In our proposed study, student academic achievement is predicted using both synthetic datasets and real-time student data utilizing a hybrid deep learning approach based on an RNN-LSTM classification algorithm. For a further explanation of the suggested design based on the RNN-LSTM classifier, see figure 6.

Figure 6 depicts the suggested system architecture with the RNN-LSTM classifier. We started by compiling data from a range of sources, including several web applications, information about actual students, and a few synthetic data sets from various sources. The obtained data may contain values that are irrelevant, excessive, duplicated, or incomplete. To eliminate such material, data pre-processing uses a variety of techniques, such as data filtering. Many data filtration techniques have been used to reduce redundant and unnecessary data, including trend analysis, categorization, grouping, visualization, association, and regression.

4.1 Database Layer: The research technique that has been presented makes use of the Kaggle dataset, which includes information such as student scores as well as demographic, interpersonal, and academic characteristics. Sometime the data has read from local repository instead of collect from various resources. The entire dataset contains 33 attributes including 3 grades which is achieved by student in each unit test. The dataset contains some personal information attributes,

4.2 Data Pre-processing Layer:

Data Filtration: It addresses noisy data, missing information, etc. Different strategies have been adopted when some data in the information is incomplete, such as filling in the gaps or disregarding the tuples. Data may contain null values that are incomprehensible to machines. This noisy data may result from poor data collecting, incorrect data input, etc. Regression, clustering, and the binning approach are used to address it

4.3 Data Normalization: Although data mining is a methodology used to handle enormous amounts of data, data reduction is important. Analyses in these situations grew more difficult when working with large amounts of data. We employ data reduction approach to get rid of this. It attempts to drop the cost of data storage and processing while increasing storage efficiency. Data cube aggregation, attribute subset selection, numerosity reduction, and dimensionality reduction are just a few of the different data reduction techniques employed. For the purpose of building the data cube, an aggregation process is performed to the data. The extremely relevant attributes were employed in the attribute subset selection procedure, while the other features were entirely ignored. Regression models, for instance, can be stored as models of data rather than as entire datasets due to the Numerosity Reduction technique.

4.4 Feature Extraction and Selection Layer:

This method of dimensionality reduction splits the initial information into recognizable categories based on the associations between the pieces of information. The notion that these massive datasets contain a significant amount of parameters and that the interpretation of those variables requires a great deal of computer power is one of the qualities that set them apart from other datasets.

4.5 Arff Feature: The. arff features are associated with the Weka based machine learning classification algorithms. When system deals with machine learning classification algorithms such as SVM, ANN, NB etc. it generates normalized feature vector after pre-processing and generate the arff file which is basically used by Weka classifiers.

4.6 Autoencoder Feature: The autoencoder features are extracted by deep learning based RNN classifier. A sort of artificial neural network called an autoencoder is often used for feature extraction and dimension reduction. It is made up of two parts: a decoder and an encoder. The encoder converts the input s into a low-dimensional vector from the input s. The decoder reconstructs the input using the low-dimensional vector. The low-dimensional vector may be regarded as a latent representation of the input if the autoencoder's reconstructed input resembles the original input.

4.7 Relational Features: The relation features are used to assigned the class able to entire dataset. The dataset contains G1, G2 and G3 are the numeric class labels to entire dataset. The sum rule techniques have used for calculate the average and based n that assigned the class label to whole training and testing dataset.

4.8 Dependency Features: The dependency features are extracted by the classification algorithm during the execution. The SVM Library has been used to extract conventional features.

Arff data in Machine Learning classifiers. SVM generates the matrix for the entire train and test dataset after the cross-validation. Each machine learning classifier utilized the dependency features that cultivate the overall accuracy.

4.9 Classification Layer

In this phase, we used conventional ML methods to predict student's academic achievement, including support vector machines, NB, J48, RF and ANN. Several classification algorithms. These classification models are utilized to efficiently mine a huge synthetic dataset. The classification accuracy of many classic machine learning methods, including Support Vector Machine, Random Forest, J48, Artificial Neural Network, and Nave Bayes etc.

4.11 Machine Learning Classifiers

The below are the classification algorithms we used as machine learning classifies.

- Naive Bayes
- J48
- Artificial Neural Network
- AdaBoost
- Random Forest
- Support Vector Machine

4.12 Hybrid Deep Learning Classifiers

RNN-LSTM: In deep learning methods, including deep neural network, RNN, is observed using experimental results. The suggested technique for forecasting student performance utilizing RNN-LSTM sigmoid, Tan-h, and ReLU function is carried out, and the outcomes are compared with other machine learning and deep learning methods. The experimental results show that RNN-LSTM (ReLU) performs better than other classification methods, with an accuracy rate of 95.5%, which is higher. When a large, complex, real-time student dataset with several value attributes is employed, the suggested approach delivers good classification accuracy.

4.13 Recommendation Layer

One the complete of classification model using both machine learning and deep learning algorithm it provides recommendation to each candidate based on achieved class labels. The recommendation is based on what are the possible improvement or career opportunities for particular candidates.

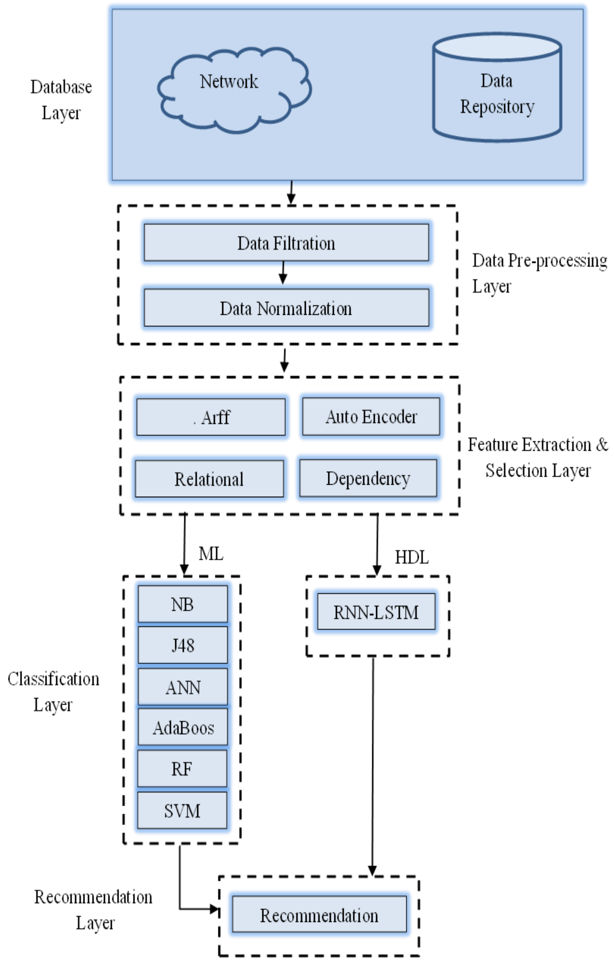

Figure 6. Proposed System Architecture

Algorithm Design

Module Training

Input: Training_DB[ ] as train dataset, set of activation function AF[].

Output: Trained module in .arff file for whole divided dataset

Step 1: Initialize the both methods Train_DB[], AF[], Iteration as epoch_size

Step 2: Extract_Feature_DataSet ← Extract_Feature(Train_DB[])

Step 3: Select_Feature_DataSet [] ← optimization (Extract_Feature_DataSet)

Step 4: Training.pkl ← Build_Classifier(Select_Feature[])

Step 5: Return Training.arff

Module Testing

Input: Test_DB [] as test instance set, Train Background Knowledge Training.arff, threshold Th

Output: Op_Map <forecasted_class_label, Sim_weight> optimized instance recommends by classification model

Step 1: Read every test records with the help of following equation

$$\text{test\_Feature (m)} = \sum_{m=0}^{n} (.\ feature\_Set[A[i] \ldots\ldots A[n] \leftarrow Test\_DB)$$

Step 2: Retrieve selected features from whole testing record $testFeature(m)$ by using following function.

Extract_Feat_set_x[t…n] = $\sum_{x=1}^{n}(t)\ test\_Feature(m)$

The feature vector is the collection of extracted hybrid attributes from given input

Step 3: Extract every train instance from trained components by using following function

$$\text{train\_Feature (m)} = \sum_{m=1}^{n} (.\ feature\_Set[A[i] \ldots\ldots A[n] \leftarrow Train.arff)$$

Step 4: Input the testing instances or record set to test classification model as $testFeature(m)$ by using following equation

Extract_Feat_Set_x[t……n] =$\sum_{x=1}^{n}(t)\ test\_Feature(m)$

The whole class labels' feature vectors are included in the Extract_Feat_Set_x[t].

Step 5: Validate individually all test instance with every training features

Calculate_weight = Calculate_Sim (Feature_Set_x || $\sum_{i=1}^{n} Feature\_Set\_y[y]$ *)*

Step 6: Return calculate_weight

5. Experimental Result and Discussion

In this experiment, a suggested RNN-LSTM method and a hybrid machine learning technique are used to predict the students' performance. We used a real-time data set with records for 480 students that contains information on their age, ethnicity, place of birth, phases, degree, section ID, number of absences per student, first semester class, second semester class, and other factors. Pre-processing and normalization are used to improve the performance of the dataset. After the dataset has undergone the necessary pre-processing, it is divided into two halves. The first is a test dataset, the second a train dataset. The three sections of the data set are divided in a 3:1 ratio.

Following the development of the hybrid deep learning method employing educational data mining, a comparison with machine learning classifiers such as support vector machines, naive bayes, random forests, J48, ANN, and the suggested RNN-LSTM is conducted.

The classification accuracy of several ML methods and the suggested RNN-LSTM utilizing EDM in the Weka framework are shown in Table 1 below.

Table 1. Performance of various machine learning algorithms and proposed RNN-LSTM using EDM

Algorithm	Accuracy	Precision	Recall	F-score
SVM	0.93	0.94	0.95	0.95
ANN	0.92	0.91	0.89	0.90
Naïve Bayes	0.89	0.88	0.93	0.91
Random Forest	0.87	0.86	0.92	0.89
J48	0.90	0.92	0.85	0.88
RNN	0.97	0.95	0.97	0.97

Consider the following figure 7 which depicts the prediction of student's performance using machine learning techniques like support vector machine, naïve bayes, artificial neural network, J48, Random Forest and proposed RNN-LSTM

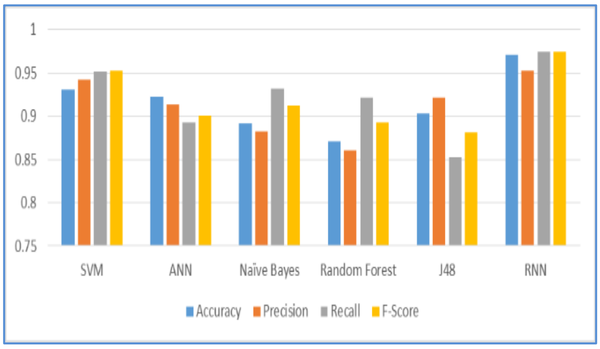

Figure 7. Performance of various machine learning algorithm and proposed RNN-LSTM using EDM

In our proposed methodology of predicting student academic performance using RNN-LSTM, three experimentations are performed to obtain, accuracy, precision, recall and f-score with various cross validation which are explained as follows

5.1 Experimentation Using RNN-LSTM (Sigmoid)

In this experiment of RNN-LSTM (sigmoid) model, accuracy, precision, recall and f-score with various cross validation are obtained. Consider the following figure 8 which depicts the validation of model with 5-, 10- and 15-fold cross validation using RNN-LSTM (sigmoid) classifier.

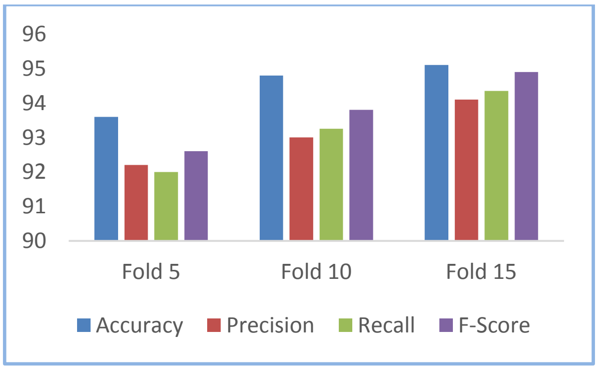

Figure 8. Validation of model with 5-, 10- and 15-fold cross validation using RNN-LSTM (Sigmoid) classifier

As per experimental findings, 15-fold cross-validation has obtained the better average classification accuracy of 95.10%.

5.3 Experimentation Using Recurrent Neural Network (Tan h) model

In this experiment of RNN-LSTM (Tan h) model, accuracy, precision, recall and f-score with various cross validation are obtained. Consider the following figure 9 which depicts the validation of model with 5, 10, 15-fold cross validation using RNN-LSTM (tan h) classifier.

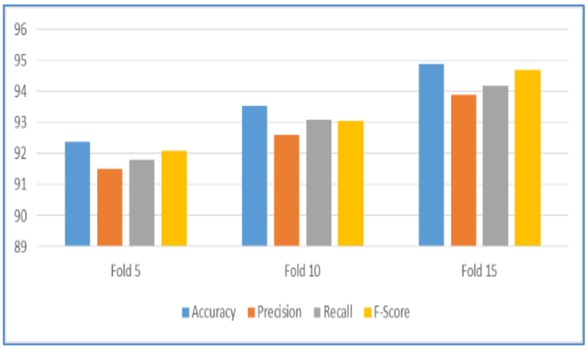

Figure 9. Validation of model with 5-, 10- and 15-fold cross validation using RNN-LSTM (Tan h) classifier

As per experimental findings, 15-fold cross-validation has obtained the better average classification accuracy of 94.9%.

5.4 Experimentation Using Recurrent Neural Network (ReLU)

In this experiment of RNN-LSTM (ReLU) model, accuracy, precision, recall and f-score with various cross validation are obtained. Figure 10 illustrates the validation of model using 5-fold, 10-fold and 15-fold cross validation respectively using RNN-LSTM (ReLU).

Consider the following figure 10 which depicts the validation of model with 5-fold cross validation using RNN-LSTM (ReLU) classifier.

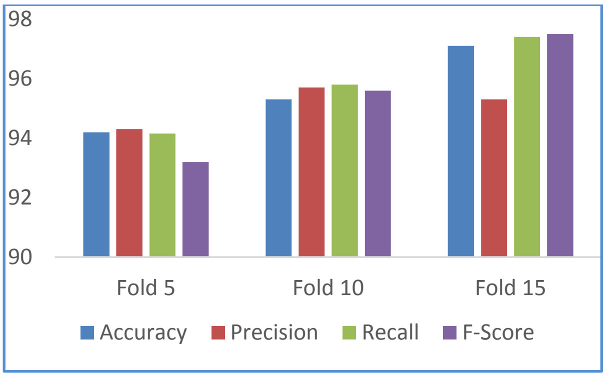

Figure 10: Validation of model with 5-, 10- and 15-fold cross validation using RNN-LSTM (ReLU) classifier

As per experimental findings, 15-fold cross-validation has obtained the better average classification accuracy of 97.1%.

In order to forecast student's performance, we used a minimum of 3 hidden layers. As per empirical findings, we conclude that RNN-LSTM with ReLu gives better detection accuracy than that of the RNN-LSTM (Tan h) and RNN-LSTM (Sigmoid) function.

Our proposed recommendation system is divided into two phases. In phase 1, recommendation is done based on student class whereas in phase 2, recommendation is based on student id. If class 1 is selected, higher technical education with aptitude knowledge is required.

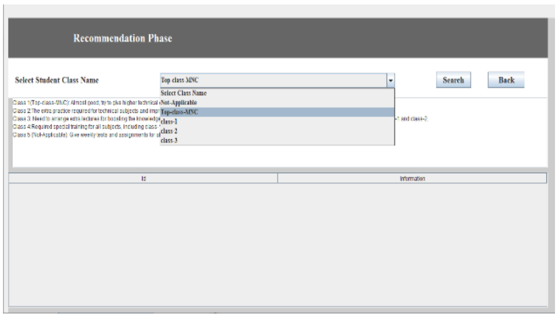

Figure 11. Selection of Student Class in Recommendation Phase

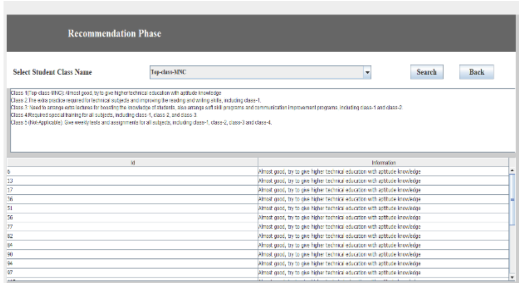

Figure 12. Selection of Student Class 1 in Recommendation Phase

If class 2 is selected, extra practice is required for technical subjects and there is a need of improving reading and writing skills.

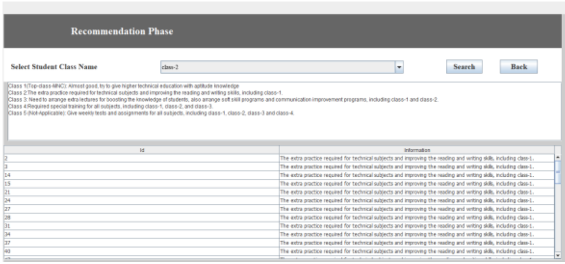

Figure 13. Selection of Student Class 2 in Recommendation Phase

If class 3 is selected, arrangement of extra lectures is needed for boosting the knowledge of students and also arranges soft skill programs and communication improvement programs.

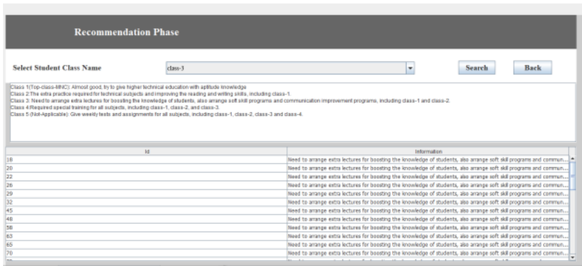

Figure 14. Selection of Student Class 3 in Recommendation Phase

If class 4 is selected, special training is required for all subjects including class 1, 2 and 3.

NeuroQuantology|December 2022 | Volume 20 | Issue 19 | Page 1877-1892| doi: 10.48047/nq.2022.20.19.NQ99163

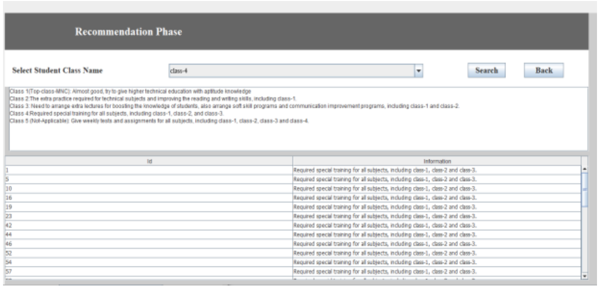

Figure 15. Selection of Student Class 4 in Recommendation Phase

If class 5 is selected, weekly tests and assignments are required including class 1, 2, 3 and 4.

If student id is selected, for example student id = 5, class 4 is selected where special training is required for all subjects including class 1, 2 and 3.

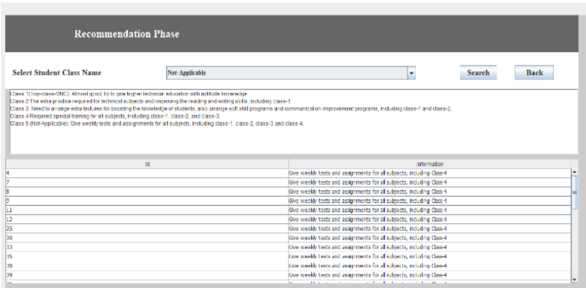

Figure 16. Selection of Student Class 5 in Recommendation Phase

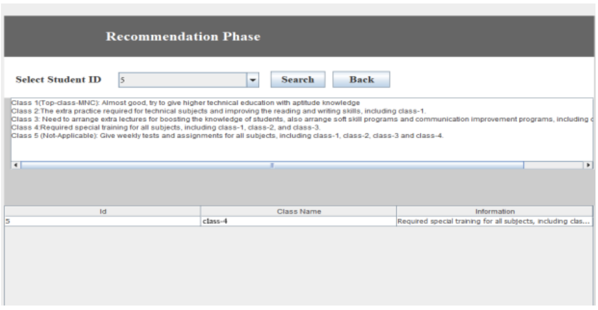

Figure 18. Selection of Student ID in Recommendation Phase

6. Conclusion and Future Scope

We evaluated the performance of numerous well-known Machine Learning methods, including the support vector machine, random forest, J48, artificial neural network, naive bayes, and proposed RNN-LSTM, using experimental results. The proposed technique for forecasting student performance utilizing Tan-h, RNN-LSTM sigmoid, and ReLU function is carried out, and the outcomes are compared with various machine learning and deep learning methods. The experimental results show that RNN-LSTM (ReLU) performs better than other classification methods, with an accuracy rate of 97 %, which is higher. When employed with a heterogeneous dataset or a real-time challenging big dataset of student data with multiple valued attributes, our proposed approach achieves good classification accuracy. Our suggested methodology can help students choose the best program choices and better manage their education. It can also help teachers assess student achievement to improve their teaching techniques. Our suggested model may review student data in educational environments to solve educational issues and thus improve educational processes. It is directed at the many people who are involved in the educational systems, including students, researchers, administrators, and teachers. It raises and elevates the standard of learning as well as the processes for learning and comprehension.

The suggested system makes use of both a huge real-time student dataset and the Kaggle dataset. Future evaluation of our model using a huge data set may be possible by adding more attributes. The suggested methodology can be used in the intelligent tutorial system to predict students' academic progress and then use that data to provide the most accurate evaluation of the students. Results were obtained with the initial set of data and confirmed. Some of the samples were excluded because they had incomplete or irrelevant responses. Given that the sample rate was reduced, there may be a drop in accuracy as a result. In the future, the clustering stage itself will be improved to include more features. Even yet, we are unable to pinpoint the specific reasons why a student or class performed poorly. We can consider psychological issues, aspects of social behavior, aspects of parental and family caregiving, and more. These elements would increase accuracy and be helpful in illuminating why some pupils underperform, enabling us to help them and point them in the direction of the "right" way to proceed.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, Jaikumar M. Patil & Dr. Sunil R. Gupta; methodology, Jaikumar M. Patil & Dr. Sunil R. Gupta; Writing-draft preparation, Jaikumar M. Patil.

References

[1] Sheena Angra and Sachin Ahuja, "Implementation of Data Mining Algorithms on Students Data using Rapid Miner", 2017, International Conference on Big Data Analytics and Computational Intelligence (ICBDAC). IEEE, *IEEE,* 2017.

[2] Worapat Paireekreng and Takorn Prexawanprasut, "An Integrated Model for Learning Style Classification in University Students Using Data Mining Techniques", *12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), IEEE*, 2015, DOI: 10.1109/ECTICon.2015.7206951

[3] Fergie Joanda Kaunang and Reymon Rotikan, "Students' Academic Performance Prediction using Data Mining", *Third International Conference on Informatics and Computing (ICIC)*. IEEE, 2018.

[4] Lu W. "Research on Educational Management System and Operation Mechanism Based on Data Mining. In 2020 International Conference on Computers, Information Processing and Advanced Education (CIPAE) (pp. 39-42). IEEE. Manas Chaturvedi. "Data Mining and its Application in EDM Domain", 2017, International Conference on Intelligent Computing and Control Systems ICICCS, IEEE.

[5] Naeimeh Delavari, "Application of Enhanced Analysis Model for Data Mining Processes in Higher Educational System", *6th international conference on information technology based higher education and training (pp. F4B-1), IEEE,* 2005.

[6] Miss. Sharayu N. Bonde and Dr. D. K. Kirange, "Survey on Evaluation of Student's Performance in Educational Data Mining", 2018, *Second International Conference on Inventive Communication and Computational Technologies (ICICCT) (pp. 209-213), IEE,*2018.

[7] Gabriella Casalino, Giovanna Castellano, Andrea Mannavola and Gennaro Vessio. "Educational Stream Data Analysis: A Case Study", *IEEE 20th Mediterranean Electrotechnical Conference (MELECON). IEEE*, 2020.

[8] Akansha Mishra, Rashi Bansal and Dr. Shailendra Narayan Singh. "Educational Data Mining and Learning Analysis", *7th International Conference on Cloud Computing, Data Science & Engineering-Confluence*. *IEEE*, 2017.

[9] Camilo Vieira, Alejandra J. Magana and Mireille Boutin, "Using Computational Methods to Analyze Educational Data", *Frontiers in Education Conference (FIE) (pp. 1-4), IEEE*, 2019.

[10] Balqis Al Breiki, Nazar Zaki and Elfadil A. Mohamed, "Using Educational Data Mining Techniques to Predict Student Performance", *International Conference on Electrical and Computing Technologies and Applications (ICECTA) (pp. 1-5), IEEE,* 2019.

[11] Dr. R. Raju, Mrs. N. Kalaiselvi, Aathiqa Sulthana M, Divya I and Selvarani A, "Educational Data Mining: A Comprehensive Study", *International Conference on System, Computation, Automation and Networking (ICSCAN)*. IEEE, 2020.

[12] Bo Guo, Rui Zhang, Guang Xu, Chuangming Shi and Li Yang, "Predicting Students Performance in Educational Data Mining", International *Symposium on Educational Technology, IEEE* ,2015.

[13] Nabila Khodeir, "Student Modeling Using Educational Data Mining Techniques", *6th International Conference on Advanced Control Circuits and Systems (ACCS) & 2019*

5th International Conference on New Paradigms in Electronics & information Technology (PEIT), IEEE, 2019.

[14] Jie Jiang and Tong Chen, "Research on the Value of Smarter Education in the Era of Big Data", *5th IEEE International Conference on Big Data Analytics (ICBDA)* (pp. 42-45*), IEEE*, 2020.

[15] Chun-li Wang, "Research on the Core Technology of Education Big Data Based on Data Mining", *6th International Conference on Big Data Analytics (ICBDA)* (pp. 5-8), *IEEE*, 2021.

[16] Sushil Shrestha and Manish Pokharel, "Machine Learning algorithm in educational data", *Artificial Intelligence for Transforming Business and Society (AITB)*. Vol. 1. *IEEE*, 2019.

[17] Muhammad Sammy Ahmad, Ahmed H. Asad and Ammar Mohammed, "A Machine Learning Based Approach for Student Performance Evaluation in Educational Data Mining", *International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*. *IEEE*, 2021.

[18] Alla Abd El-Rady, "An Ontological Model to Predict Dropout Students Using Machine Learning Techniques", *3rd International Conference on Computer Applications & Information Security (ICCAIS)*. *IEEE*, 2020.

[19] Muhib Al-kmali, Hamzah Mugahed, Wadii Boulila, Mohammed Al-Sarem and Anmar Abuhamdah. "A Machine-Learning based Approach to Support Academic Decision-Making at Higher Educational Institutions", I*nternational symposium on networks, computers and communications (ISNCC)*. *IEEE*, 2020.

[20] J. de O. Santos Kelly, Angelo G. Menezes, Andre B. de Carvalho and Carlos A E. Montesco, "Supervised Learning in the Context of Educational Data Mining to Avoid University Students Dropout", 19th *International Conference on Advanced Learning Technologies (ICALT), IEEE,* 2019.

[21] Chitra Jalota and Rashmi Agrawal, "Analysis of Educational Data Mining using Classification", *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (Com-IT-Con), IEEE,* 2019.

[22] Nongnuch Ketui, Warawut Wisomka and Kanitha Homjun. "Using Classification Data Mining Techniques for Students Performance Prediction", *2019, 4th International Conference on Digital Arts, Media and Technology and 2nd ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering, IEEE,* 2019.

[23] Keisuke Abe, "Data Mining and Machine Learning Applications for Educational Big Data in the University", *Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech) (pp. 350-355), IEEE,* 2019.

[24] Abdul Aleem and Manoj Madhava Gore, "Educational Data Mining Methods: A Survey", *9th International Conference on Communication Systems and Network Technologies (CSNT) (pp. 182-188), IEEE,* 2020.

[25] Moloud Abdar, Mariam Zomorodi-Moghadam and Xujuan Zhou. "An Ensemble-based Decision Tree Approach for Educational Data Mining", *5th International Conference on Behavioral, Economic, and Socio-Cultural Computing (BESC)*. *IEEE*, 2018.

[26] Ms. Abhilasha Dangi and Dr. Sumit Srivastava, "Educational data Classification using Selective Naïve Bayes for Quota categorization", *International Conference on MOOC, Innovation and Technology in Education (MITE) (pp. 118-121, IEEE,* 2014.

[27] Latifaestrelita Indi Pramesti Aji and Andi Sunyoto, "An Implementation of C4.5 Classification Algorithm to Analyze Student's Performance", *3rd International Conference on Information and Communication Technology (ICOIACT), IEEE*, 2020.

[28] Sagardeep Roy and Anchal Garg, "Analyzing Performance of Students by Using Data Mining Techniques," 2017, *4th Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON), IEEE,* 2017.

[29] Dr. S. Senthil and Wong Mu Lin. "Applying Classification Techniques to Predict Students' Academic Results", *IEEE International Conference on Current Trends in Advanced Computing (ICCTAC). IEEE,* 2017.

[30] Ndiatenda Ndou, Ritesh Ajoodha and Ashwini Jadhav, "Educational Data-mining to Determine Student Success at Higher Education Institutions", *2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC), IEEE,* 2020.

[31] Anoopkumar M and Dr. A. M. J. Md. Zubair Rahman, "A Review on Data Mining Techniques and Factors Used in Educational Data Mining to Predict Student Amelioration", *International Conference on Data Mining and Advanced Computing (SAPIENCE), IEEE,* 2016.

[32] Edona Doko and Lejla Abazi Bexheti, "A Systematic Mapping Study of Educational Technologies based on Educational Data Mining and Learning Analytics", *7th Mediterranean Conference on Embedded Computing (MECO) (pp. 1-4). IEEE*, 2018.

[33] Nurul Hidayat, Retantyo Wardoyo and Azhari SN, "Educational Data Mining (EDM) as a Model for Students' Evaluation in Learning Environment", *Third International Conference on Informatics and Computing (ICIC)* (pp. 1-4), *IEEE*, 2020.

[34] Wei Zhang and Shiming Qin, "A Brief Analysis of the Key Technologies and Applications of Educational Data Mining on Online Learning Platform", *3rd International Conference on Big Data Analysis (ICBDA) (pp. 83-86), IEEE*, 2018.

[35] Vinayak Hegde and Sushma Rao H S, "A Framework to Analyze Performance of Student's in Programming Language Using Educational Data Mining", *International Conference on Computational Intelligence and Computing Research (ICCIC) (pp. 1-4), IEEE*, 2017.

[36] Steven Lehr, Hong Liu, Sean Klinglesmith and Alex Konyha, "Use Educational Data Mining to Predict Undergraduate Retention", *16th International Conference on Advanced Learning Technologies (ICALT) (pp. 428-430, IEEE*, 2016.

[37] Rashi Bansal, Akansha Mishra and Dr. Shailendra Narayana Singh, "Mining of Educational Data for Analysing Students Overall Performance", *7th International Conference on Cloud Computing, Data Science & Engineering-Confluence (pp. 495-497), IEEE,* 2017.

[38] Moustafa M. Kurdi, Hatim Al-Khafagi and Imad Elzein, "Mining Educational Data to Analyze Students' Behavior and Performance", *JCCO Joint International Conference on ICT in Education and Training, International Conference on Computing in Arabic, and International Conference on Geocomputing (JCCO: TICET-ICCA-GECO) (pp. 1-5), IEEE,* 2018.

[39] Viviana Parraga & Juan Zaldumbide. "Systematic Mapping Study of Literature on Educational Data Mining to Determine Factors That Affect School Performance", *International Conference on Information Systems and Computer Science (INCISCOS) (pp. 239-245), IEEE,* 2018.

[40] Aberbach Hicham, Adil Jeghal, Abdelouahed Sabri and Hamid Tairi, "A Survey on Educational Data Mining [2014-2019]", *International Conference on Intelligent Systems and Computer Vision (ISCV) (pp. 1-6), IEEE,* 2020.

[41] Rita Tavares, Rui Vieira and Luís Pedro, "A preliminary proposal of a conceptual Educational Data Mining framework for Science Education Scientific competences development and self-regulated learning", *International Symposium on Computers in Education (SIIE) (pp. 1-6), IEEE*, 2017.

[42] Karan Sukhija, Dr. Manish Jindal and Dr. Naveen Aggarwal. "The Recent State of Educational Data Mining: A Survey and Future Visions", *IEEE 3rd International Conference on MOOCs, Innovation and*

Technology in Education (MITE). *IEEE*, 2015.

[43] Hanife Goker, Halil Ibrahim Bulbul and Erdal Irmak, "The Estimation of Students' Academic Success by Data Mining Methods", *12th International Conference on Machine Learning and Applications (Vol. 2, pp. 535-539), IEEE,*2013.

[44] Mudasir Ashraf, Dr. Majid Zaman and Dr. Muheet Ahmed, "Performance Analysis and Different Subject Combinations: An Empirical and Analytical Discourse of Educational Data Mining", *8th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 287-292), IEEE,* 2018.