International Conference on Recent Trends in Data Science and its Applications
DOI: rp-9788770040723.009

Student Performance Analysis Using Deep Learning Technologies

Dr. Parkavi K
School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
parkavi.k@vit.ac.in#

Nivedha M
School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
nivedhamohan01@gmail.com#

Sandhya K
Department of Computer Science and Engineering
Jeppiaar Engineering College
Chennai, India
sandhyakumaravel2312@gmail.com#

***Abstract*— Learning is a process, in this course of long run it is necessary to measure it's performance on the major people discharging it. Because there is so much data available, using Deep Learning to anticipate what will happen in the future is becoming more and more common. The ability of deep learning to forecast students' academic performance is suggested by this study. At the end of the school day, students either pass or fail. When that happens, it's too late to save the students. To keep kids from failing, it is necessary to predict SAP in advance. When that happens, it's too late to save the students. To stop pupils from failing, it is necessary to anticipate their academic performance in advance. The study of data mining, machine learning, and statistics as they relate to information produced in educational contexts is known as education data mining (EDM). It is all about enhancing learning outcomes by examining data gathered while we lecture. The main goal of this paper is to give an appropriate method of measuring student performance using FNN, DBN, ICGAN-DSVM, LSTM, and BLSTM. This process involves collection of datasets, that is gathered from students' academic dataset which can be further enhanced by ICGAN that increases the data volume. It is discussed how the data were gathered, prepared, and developed. The models are built, trained and applied on the dataset. These results are compared to each other and are classified situationally to the requirement.**

Keywords*— Education Data Mining, FNN, DBN, ICGAN-DSVM, LSTM, BLSTM

I. Introduction

There lies a huge difference between learning, testing and deploying it in a real world. Performance of a student is based on these factors. These factors can be measured with various technologies such as machine learning, deep learning, etc. In the age of the information revolution, mathematics is one of the fundamental cornerstones for numerous subjects, as well as the backbone of scientific endeavor. A difficult activity that might assist students and teachers in monitoring student performance development is measuring student performance. Several strategies have been researched and compared to produce the optimum prediction model in an effort to improve the results of prediction. Learning is affected by various factors such as quantity of information, emotions, mistakes, novelty of the brain, learning styles, social learning, and teaching. Cognitive overload is the term used by brain scientists to describe the state in which a person's brain is overloaded with new information. Cognitive overload is caused by having too much new information, which eventually hinders learning. How might cognitive overload be lessened? There are primarily two methods. The first is the quantitative approach, in which you merely present less novel information. Before introducing new knowledge, you give students time to comprehend the majority of what they have learned. The other approach is qualitative, where you modify how you offer material to make it less intimidating. Research has shown that our emotions have an impact on every aspect of our lives, including how we remember information and how we interpret it. Kids, who are less mature, are especially at risk for this.

Students study a lot throughout the year and are tensed of proving it in an examination. Performance is measured based on testing, but this cannot prove an individual's content learned and applied. Considering various intrinsic values apart from the above ones, such as data being gathered on knowledge gained, skill growth, values clarification, and performance level. For all educational institutions, raising educational standards and student performance is of vital importance. There has been an increase in interest from higher education providers and institutions to adopt machine learning techniques to understand students' behaviors, learning patterns, and drop-out trends in education. However, educational analytics in learning has never been explored as is the case in the health sciences and chemistry.

Deep learning is chosen as a key for the following grounds. For more focused performance that is required to be similar to human activities, deep structures learning, also known as hierarchical learning, or Deep Learning (DL), first appeared in 2006. The capabilities of DL are strong in the areas of prediction, classification, identification, and detection. DL is more popular than other machine learning algorithms due to its ability to process both structured and unstructured data as well as manage large amounts of data. When it comes to prediction, DL has been used to a variety of issues, including the Intelligent Transport System and several economic and educational areas. Learning and higher education challenges have been addressed with DL in a number of studies in the realm of education. DL is effective for classifying, detecting, and identifying a lot of data in addition to prediction. Classifying, identifying, or detecting are examples of four DL tasks that can be combined and used interchangeably in a single study.

DOI: rp-9788770040723.009

Numerous problems need to be solved in education, particularly higher education. Starting with the most basic model, the Feedforward Neural Network, ICGAN-DSVM is best supported for small datasets, five classes of machine learning algorithms, and the Deep Belief Network (DBN) framework for deep learning were executed and compared.

II. RELATED WORK

Aya Nabil et al. [1] says that one of the significant research questions in the discipline of educational data mining is how to forecast the academic performance of students at a preliminary phase of a **course** (EDM). Subjects like "Data Structures" and "Programming" in undergraduate programs have significant failure and dropout rates because students struggle with a variety of issues. As a consequence, EDM is utilized to evaluate student data collected from different school environments in order to forecast the academic performance of students, which enables them to perform well in next courses. The primary objective of this study is to investigate the effectiveness of deep learning in the area of EDM, particularly in forecasting the academic performance of students and identifying students who are at danger of failing. Using the grades obtained in the courses written in the previous academic year, the K-nearest neighbor, decision tree, deep neural network (DNN), support vector classifier, logistic regression, gradient boosting and random forest were utilized in this paper to create models that would predict academic performance of students in future courses. Additionally, they compared ADASYN, SMOTE, SMOTE-ENN and ROS among other resampling techniques to address the issue of an unbalanced dataset. According to the experimental findings, the presented DNN model outperforms support vector classifiers, K-nearest neighbors, logistic regression and decision trees in predicting performance of students in the course of data structure and identifying students at threat of failing early in a semester. The result is obtained with an accuracy of 89%.

Mohammad Hafiz MohdYusof, et al. [2] proposed that the ideas of precision medicine have been adopted by precision education. It uses machine learning, algorithms, and techniques of data manipulation to make predictions, and will eventually be used to create customized school intervention plans. Numerous applications and themes in precision education research have been studied in this work. After that, a model was created using one of its techniques, deep learning in Malaysia. The purpose of the research is to forecast the performance of students in English. When the proposed deep learning model is tested for folder classification using three test datasets, it performs with 93% accuracy.

PhaukSokkhey, Takeo Okazaki [4] proposed that evaluating performance of students is a difficult undertaking that can assist both students and teachers in monitoring student performance development. One of the fundamental foundations of many topics, mathematics serves as the foundation for every scientific endeavor in the period of the tech revolution. Several strategies have been researched and compared to produce the optimum prediction system in an effort to enhance the results of prediction. In this article, they offered a comparative examination of machine learning (ML) algorithms, statistical analytic methods and a deep learning architecture for forecasting mathematics performance of students. The Deep Belief Network (DBN), a deep learning architecture, five classes of machine learning algorithms, and the statistical approach structural equation modelling (SEM) were used. The same properties from two datasets of varying sizes were employed. Random Forest (RF) was discovered to perform better than other models in predicting the performance of students across the three different datasets.

Kwok Tai Chui et al. [5] say that it's been established that supported learning has been extremely important in raising the quality of education. Tutoring in the classroom and at home offers pupils individualized support and constructive criticism of their learning. Predicting the performance of students, which shows their knowledge of the subjects, is of great interest. To provide a solid basis for upcoming courses and a profession, it is especially important for students to manage their core knowledge. This study proposes an enhanced conditional generative adversarial network-based deep support vector machine (ICGAN-DSVM) technique to forecast the performance of students in learning environments including classroom and home tutoring. Due to the nature of the student academic record, sample sizes are often small. Due to the academic dataset's limited sample size, ICGAN-DSVM offers two advantages: ICGAN increases the volume of data, while DSVM improves the accuracy of prediction through deep learning. The suggested ICGAN-DSVM produces specificity, sensitivity, and area under the receiver operating characteristic curve (AUC) of 0.968, 0.971, and 0.954, respectively, according to results from cross-validation of 10-folds. Additionally, the results indicate that including both home and school tutoring in the model may enhance performance over using just home tutoring or just school tutoring alone. A comparison between ICGAN & DSVM and the conventional conditional generative adversarial network has been done to demonstrate the need for ICGAN and DSVM. Additionally, the suggested kernel design using heuristic-based multiple kernel learning (MKL) is contrasted with conventional kernels including radial basis function (RBF), linear, sigmoid and polynomial. Following the presentation of the forecast of performance of students with and without GAN, a comparison with DSVM and regular SVM is made. In terms of the performance metrics specificity, sensitivity, and AUC, the suggested ICGAN-DSVM performs 8 to 29% better than comparable studies.

Yueh-huiVanesa Chiang, et al. [6] proposed that to forecast student performance in introductory computer programming classes, this paper utilized deep learning algorithms with Moodle logs. Particularly, this study would like to use prediction results to identify potential low-performing students who may need assistance from teachers. The results suggested that deep learning models are promising to predict student performance and identify low-performing students in the researched context. What the prediction results provided by the models can inform teachers in learning settings was also further discussed in this paper.

DOI: rp-9788770040723.009

DiaaUliyan [7] says that predictive analytics' ability to assist institutions judge performance of students is becoming more widely known. Big data analytics may offer information that supports academic performance and completion rates, like student demographic facts. For instance, analytics of learning is a crucial part of big data at universities and may give strategic decision-makers the chance to analyze learning activities across time. This study involved a two-year retrospective review of student learning data from the University of Ha'il. Bidirectional long short-term model (BLSTM), a deep learning approach, was used to analyze students for whom retention would have been at risk. The algorithm offers a variety of variables that may be used to gauge a new student's performance, which helps with dropout and early retention prediction. Furthermore, every student's label was predicted separately using the conditional random Field (CRF) approach for sequence labelling. The predictive model's experimental findings suggest that utilizing BLSTM and CRF approaches, student retention may be predicted with a high degree of accuracy.

III. METHODOLOGY

In this paper, we are discussing the different possible ways in the prediction of performance of students using deep learning technologies. The deep learning technologies that are discussed here are Feed Forward Neural network, Deep Belief Network (DBN), Improved Conditional Generative Adversarial Network based Deep Support Vector Machine (ICGAN-DSVM), Long Short-Term Memory (LSTM), Bidirectional Long Short-term model (BLSTM)and Conditional Random Fields (CRF). But before we apply the model, we need to do some steps which support the prediction. They are Collecting the data, Pre-processing the data, Method to validate the model and finding out the evaluation measures. Let us see each of these steps in detail.

A. Collecting the Data

In this step we need to collect the student data. To determine the performance of students with higher accuracy, we might require three types of data. They are Demographic data (Age, Gender, Area of location, Financial Status), Academic data (GPA and CGPA of students, Academic tests, Earlier test scores), and Behavioural data(Activity submission log, Library history).

B. Pre-Processing the Data

This process converts the data which is collected in raw format into the required format which is useful in acquiring the required result of predicting the performance.

a. Cleaning the data: In this step, the values which are missing are removed.

b. Data Discretization: In this step, attribute values of continuous data are converted into finite set of intervals. For example, let the student's data has different categories based on marks like Excellent, Very Good, Good, Poor, Fail. We can convert it into two categories called Pass and Fail in this step.

c. Encoding the feature: As the algorithm doesn't understand categorical data, we convert all the categorical data into numerical data. We can use label encoding for this.

d. Handling Imbalanced Dataset Problem: Suppose we have two class labels. They are Pass and Fail. Pass have higher number of occurrence than Fail. So Pass becomes the majority class and Fail becomes the minority class which affects the performance of the algorithm. So we need to use a resampling method to solve this problem. The resampling methods are Over sampling (It increases the count of minority classes. This method is used for small sized datasets), Under sampling (It reduces the count of majority classes. This method is used for large sized datasets), Hybrid (It is a combination of both Over sampling and Under sampling).

e. Feature Scaling: To improve the learning process of the method, the dataset's features need to be scaled so that it has a small range. The most common scaling technique that is used is standard scaler. We need to find out the standard deviation and mean for every feature to standardize the feature.

$$X_{Scaled} = \frac{x-\mu}{\sigma} X_{Scaled} = \frac{x-\mu}{\sigma} \qquad (1)$$

C. Validating the Model

a. Method of Random Hold-out: We divide the dataset into 80% and 20%. 80% is used in training and the remaining 20% is used in testing.

b. K-fold Cross-Validation: For the dataset with a small size, we can use K-Fold Cross-Validation. This method divides the dataset into equal sizes containing K subsets. All the subsets except one will be used for training the model. One subset will be used for testing.

D. Deep Learning Models

a. Feed Forward Neural Network: The simple one among other types is Feed-Forward Neural Network. Here all the neurons are fully connected and does not form a cycle. The neurons are in inputs, output, and hidden layers. Multiple hidden nodes are traversed by the data since it always travels forward but not backward. The Feed Forward Neural network would be executed using the train and test data. TensorFlow may be used as the backend, and Keras as the frontend. Between the hidden layers, the ReLU activation function is utilized, while at the output layer, the softmax activation function is utilized. [2][3]

b. Deep Belief Network (DBN): A subset of deep learning is deep belief networks. It's a multi-layer belief network made up of Restricted Boltzmann Machines (RBMs) that are built on top of one another to create a DBN. A greedy layer-wise learning process is followed by a three-layer DBN that includes the pre-training. Particularly, while one layer is added to the network at a time using the contrastive divergence (CD) method, only the top layer of the network is trained as an RBM. Once each RBM has been trained, the weights were clamped along with the introduction of a new layer to repeat the procedure. In a DBN, there are two phases. Unsupervised pre-training, which is the first stage, teaches features using only input values with no labels. The

DOI: rp-9788770040723.009

second level of tuning involves employing a label using supervised fine-tuning and an error back propagation method. In order to maximise the benefits of the DBN model, we must build numerous models with varied hidden layer counts, hidden layer node counts, and hyper parameters.[4]

c. Improved Conditional Generative Adversarial Network based Deep Support Vector Machine (ICGAN-DSVM): Since educational datasets often contain modest amounts of data, the majority of machine learning algorithms exhibit a shallow learning approach. ICGAN mimics fresh training datasets to solve the issue of low data volume, whereas DSVM extend the SVM algorithm from shallow learning approach to deep learning approach. As a fundamental distinction from conventional deep neural networks, DSVM excels with limited datasets. In comparison to earlier efforts, the ICGAN-DSVM method increases specificity, sensitivity, and AUC by around 8 to 29%. More performance of student data for training are produced using ICGAN. The prediction model for performance of students is handled by DSVM. To expand the dataset's data amount, GAN is selected. In the training stage, the generator and discriminator strive to reach the Nash equilibrium. In reality, educational dataset which are of small size is usual. In general, it is challenging to collect continuous data at the beginning, while learners are still young. The random noise vector (the generator's input) is unhindered in the original GAN, which might result in disastrous theory corruption. Conditional variables are added to the generator and discriminator to overcome this restriction. The current architectures of CGAN, InfoGAN, and ACGAN are combined to create improved CGAN (ICGAN). ICGAN generates data with less bias. The concepts of (i) adding a conditional variable to the discriminator, (ii) including another network with the discriminator, and (iii) giving a label to each produced sample are all incorporated within ICGAN. The problems of Classification are typically solved by the Support Vector Machine algorithm. We can divide the n-dimensional space into classes using the best line we can draw using SVM, making it simple to classify fresh data points. The best line is called the best decision boundary or hyperplane. Utilizing DSVM architecture, the prediction model to predict the performance of students is put into practice. In general, it has an output layer of SVM and a number of hidden SVM layers. In comparison to many other deep learning architectures like deep neural networks, DSVM has a number of advantages, including the ability to handle the issue of very large input vectors and small training datasets, more flexibility in the design of kernel functions, and strong regularisation power in the output layer SVM to prevent over-fitting. [5].

d. Long Short-Term Memory (LSTM): Long Short-Term Memory (LSTM) is a form of artificial neural network used in deep learning and artificial intelligence. Unlike conventional feedforward neural networks, LSTM has feedback connections. In addition to single data points, a recurrent neural network (RNN) of this type may analyze entire data sequences. Both "long-term memory" and "short-term memory" are compared to a standard RNN. The name of LSTM is obtained due to this reason. This network model can be made using four layers. They are Masking, LSTM, Dropout and Dense layers. Cross-entropy method is used when training the model. Nadam is set as an optimizer. Masking layer was aimed at filtering the missing values caused by data from different semesters and courses with incompletely identical learning activities. The filtered data served as input to the LSTM layer. In each cell, its cell state, hidden state, and input data from 10 time points was processed consecutively one by one through the input gate, forget gate and output gate. The output of the last time point was the LSTM layer output and was sent to the Dropout layer for preventing overfitting. Last, the output was sent to the Dense layer, using softmax as activation function to calculate the possibilities of three classes (i.e., low-performing, average-performing, and high-performing). The highest possibility was the result of categorizing. [6].

e. Bidirectional Long Short-term model (BLSTM) and Conditional Random Fields (CRF): BLSTM includes a variety of elements that may be used to judge the performance of a new student, which helps with early dropout and retention prediction. Furthermore, the condition Random Field (CRF) approach to sequence labelling was used to forecast each student's label separately. One of the most effective artificial recurrent neural networks (RNNs) for time series data categorization and prediction is BLSTM. Using the sequence labelling approach, it may be utilized to process whole data sequences and keep them in compressed form. Additionally, BLSTM excels in a variety of sequence labelling tasks, including handwriting and speech recognition, and with long-term data dependencies. This benefit led us to develop the prediction problem from a sequence perspective. The BLSTM method must be trained over the student features in both directions, specifically forward and backward with hidden states, before concatenating the output from both directions. This follows the stage in which we attempt to describe students using their features. BLSTM has shown its effectiveness in a number of related tasks. Each unit in the BLSTM structure is made up of an LSTM unit. The primary responsibility of BLSTM is to extract and encode specific student attributes. To forecast each label of a student on their own, CRF is utilized for sequence labelling. It may then study the relationships between sets of features and labels using a state score and model the interactions between neighbouring labels using a transition score [7].

E. Evaluation Measures

Some of the evaluation measures are Precision, F1-score, Accuracy, Classification error, Recall, Sensitivity, and Specificity. The most common evaluation measure is accuracy. When the dataset contains the same number of instances for each class, accuracy can be employed. Here, TP - True Positive, TN - True Negative, FP - False Positive, FN - False Negative, N_n – Number of negative samples, N_p – Number of positive samples.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP+FN}\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

$$\text{Precision} = \frac{TP}{TP+FP}\text{Precision} = \frac{TP}{TP+FP} \tag{4}$$

$$F1 - \text{Score} = 2 * \left(\frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}\right) \tag{5}$$

$$\text{Classification Error} = \frac{FP+FN}{TP+TN+FP+FN} \tag{6}$$

$$\text{Specificity} = \frac{TN}{N_n}\text{Specificity} = \frac{TN}{N_n} \tag{7}$$

$$\text{Sensitivity} = \frac{TP}{N_p}\text{Sensitivity} = \frac{TP}{N_p} \tag{8}$$

IV. Results

a. Feed-Forward Neural Network

This model gives accuracy of 93% in 3 testing folds, 88% in 5 testing folds, and 77% accuracy in 10 testing folds [2]. This model gives an accuracy of around 35% when it is tested for accuracy [3]. This model sometimes predicts the Pass students as Fail and Fail students as Pass. So, this model is far from perfect.

b. Deep Belief Network (DBN)

This model generates an accuracy of 75% and 88% [4]. So this model also doesn't meet up the required prediction accuracy.

c. Improved Conditional Generative Adversarial Network based Deep Support Vector Machine (ICGAN-DSVM)

The suggested ICGAN-DSVM produces specificity, sensitivity, and AUC of 0.968, 0.971, and 0.954, respectively, according to results from 10-fold cross-validation.

d. *Long Short-Term Memory (LSTM)*

The accuracy was around 60% for the given LSTM model.

e. *Bidirectional Long Short-Term Model (BLSTM) and Conditional Random Fields (CRF)*

For the given method the precision was found to be 89.1%, and recall is 88.5%. Accuracy is found around 90%.

From these results, we can understand that the two methods work well in predicting student performance. They are

Deep Support Vector Machine with improved Conditional Generative Adversarial Network (ICGAN-DSVM)

Bidirectional Long Short-term model (BLSTM) and Conditional Random Fields (CRF)

We can also note that **BLSTM and CRF** give the highest accuracy of around 90%. This is due its long-term memory and its ability to train the student in both directions.

V. Conclusion

In this paper, we provided with a detailed method to conduct Student Performance analysis. Here we concentrated on Deep Learning technologies. The initial stages of the analysis include Data Collection, Pre-Processing the data, and Validating the Model. During the Pre-Processing stage, we are performing data cleaning, data discretization, encoding the feature, handling imbalanced dataset problems and feature scaling. The validation of the model is done using the method of Random Hold-out and K-fold Cross-Validation. This paper aims to produce multiple classifiers to investigate the model that can best describe a student's performance. The deep learning models considered here are Feed Forward Neural Network, Deep Belief Network (DBN), Improved Conditional Generative Adversarial Network based Deep Support Vector Machine (ICGAN-DSVM), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term model (BLSTM) and Conditional Random Fields (CRF). By using different evaluation measures, we came to the conclusion that Improved Conditional Generative Adversarial Network based Deep Support Vector Machine (ICGAN-DSVM) and Bidirectional Long Short-Term model (BLSTM) and Conditional Random Fields (CRF) provide us with the best result. This may use to enhance the student's performance by analyzing their results in an earlier stage and to reduce the dropout rate which is at the primary level (Classes 1 to 5) increased to 1.5 percent in the academic year 2021–2022 from 0.8 percent in the previous year, according to the data. Dropout rates have increased at the upper primary level (Classes 6-8), rising to 3% in 2021–22 from 1.9% in 2020–21.

References

[1] A.Nabil, M. Seyam,and A.Abou-Elfetouh,"Prediction of students' academic performance based on courses' grades using deep neural networks," IEEE Access, vol. 9, pp.140731-140746, 2021.

[2] M.H.M. Yusof,and I.Abd Khalid,"Precision education reviews: a case study on predicting student's performance using feed forward neural network," In 2021 International Conference of Technology, Science and Administration (ICTSA), IEEE, pp. 1-4, March, 2021

[3] A.Yunita, H.B. Santoso,and Z.A.Hasibuan,"Deep Learning for Predicting Students' Academic Performance," In 2019Fourth International Conference on Informatics and Computing (ICIC), IEEE, pp. 1-6, October2019.

[4] P. Sokkhey,and T.Okazaki,"Comparative study of prediction models on high school student performance in mathematics," In 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), IEEE, pp. 1-4, June2019.

[5] K.T.Chui, R.W.Liu, M. Zhao,and P.O.De Pablos,"Predicting students' performance with school and family tutoring using generative adversarial network-based deep support vector machine," IEEE Access, vol. 8, pp.86745-86752, 2020.

[6] Y.H.V.Chiang, Y.R. Lin,and N.S.Chen,"Using deep learning models to predict student performance in introductory computer programming courses," In 2022 International Conference on Advanced Learning Technologies (ICALT), IEEE, pp. 180-182, July2022.

[7] D.Uliyan, A.S.Aljaloud, A.Alkhalil, H.S.Al Amer, M.A.E.A. Mohamed, and A.F.M.Alogali,"Deep learning model to predict students retention using BLSTM and CRF," IEEE Access, vol. 9, pp.135550-135558, 2021.

[8] L.A.BuschettoMacarini, C.Cechinel, M.F.Batista Machado, V.FariaCulmant Ramos,and R.Munoz,"Predicting students success in blended learning - evaluating different interactions inside learning management systems," Applied Sciences, vol. 9, no. 24, p.5523, 2019.

[9] H.A.Mengash,"Using data mining techniques to predict student performance to support decision making in university admission systems," IEEE Access, vol. 8, pp.55462-55470, 2020.

DOI: rp-9788770040723.009

[10] B.Sekeroglu, K. Dimililer,and K.Tuncal, “Student performance prediction and classification using machine learning algorithms,” In Proceedings of the 2019 8th International Conference on Educational and Information Technology, pp. 7-11, March2019.

[11] K. Parkavi,and P.Vivekanandan, “A novel cluster-based energy efficient protocol for wireless networks,” In IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM-2012), IEEE, pp. 645-650, March2012.

[12] J.S. RC, and K. Parkavi, “Investigations on Bio-Inspired Algorithm for Network Intrusion Detection–A Review”.

[13] F.Ofori, E. Maina,and R.Gitonga, “Using machine learning algorithms to predict studentsâ€™ performance and improve learning outcome: a literature based review,” Journal of Information and Technology, vol. 4,no1, 2020.

[14] F.Kouser, A.F. Meghji,and N.A.Mahoto, “Early Detection of Failure Risks from Students' Data,” In 2020 International Conference on Emerging Trends in Smart Technologies (ICETST), IEEE, pp. 1-6, March, 2020.

[15] D.Baneres, M.E. Rodríguez-Gonzalez,and M.Serra, “An early feedback prediction system for learners at-risk within a first-year higher education course,” IEEE Transactions on Learning Technologies, vol. 12, no. 2, pp.249-263, 2019.

[16] D.Das, A.K.Shakir, M.Rabbani, S.Golam, M.Rahman, S.M.Shaharum, S.Khatun, N.B.Fadilah, K.M.Qaiduzzaman, M. Islam,and M.Arman, “A comparative analysis of four classification algorithms for university students performance detection,”In InECCE2019, Springer, Singapore, pp. 415-424, 2020.

[17] M.R.I.Rifat, A. Al Imran,and A.S.M.Badrudduza, “EduNet: a deep neural network approach for predicting CGPA of undergraduate students,” In 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), IEEE, pp. 1-6May, 2019.

[18] Rajesh, M., &Sitharthan, R. (2022). Introduction to the special section on cyber-physical system for autonomous process control in industry 5.0. Computers and Electrical Engineering, 104, 108481.

[19] A.Hernández-Blanco, B.Herrera-Flores, D. Tomás,and B.Navarro-Colorado, “A systematic review of deep learning approaches to educational data mining,” Complexity, 2019.

[20] F.Chen, and Y.Cui, “Utilizing Student Time Series Behaviour in Learning Management Systems for Early Prediction of Course Performance,” Journal of Learning Analytics, vol. 7, no. 2, pp.1-17, 2020.

[21] M.Tight, “Student retention and engagement in higher education,” Journal of further and Higher Education, vol. 44, no. 5, pp.689-704, 2020.

[22] Z.Cui, R.Ke, Z. Pu,and Y.Wang, “Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values,” Transportation Research Part C: Emerging Technologies, vol. 118, p.102674, 2020.

[23] Sitharthan, R., Vimal, S., Verma, A., Karthikeyan, M., Dhanabalan, S. S., Prabaharan, N., ...&Eswaran, T. (2023). Smart microgrid with the internet of things for adequate energy management and analysis. Computers and Electrical Engineering, 106, 108556.

[24] N.S. Sapare,and S.M., Beelagi, “Comparison study of regression models for the prediction of post-graduation admissions using machine learning techniques,” In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), IEEE, pp. 822-828, January, 2021.

[25] O.Embarak, “Apply machine learning algorithms to predict at-risk students to admission period,” In 2020 Seventh International Conference on Information Technology Trends (ITT), IEEE, pp. 190-195, November2020.

[26] E. Alyahyan,and D.Düştegör, “Predicting academic success in higher education: literature review and best practices,” International Journal of Educational Technology in Higher Education, vol. 17, no. 1, pp.1-21, 2020.

[27] D.Buenaño-Fernández, D. Gil,and S.Luján-Mora, “Application of machine learning in predicting performance for computer engineering students: A case study,” Sustainability, vol. 11, no. 10, p.2833, 2019.

[28] R. Ghorbani,and R., Ghousi, “Comparing different resampling methods in predicting students’ performance using machine learning techniques,” IEEE Access, vol. 8, pp.67899-67911, 2020.

[29] I.Khan, A.Al Sadiri, A.R. Ahmad, and N.Jabeur, “Tracking student performance in introductory programming by means of machine learning,” In 2019 4th mec international conference on big data and smart city (icbdsc), IEEE, pp. 1-6, January2019.

[30] K. Al Mayahi,and M.Al-Bahri, “Machine learning based predicting student academic success,” In 2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), IEEE, pp. 264-268, October, 2020.

[31] M.Akour, H. Alsghaier,and O.Al Qasem,“The effectiveness of using deep learning algorithms in predicting students achievements,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 19, no. 1, pp.387-393, 2020.

[32] J.Sultana, M. Usha,and M.A.H.Farquad,“An efficient deep learning method to predict students performance,” Higher Education Quality Assurance and Enhancement, 2018.

[33] U.Pujianto, W.A. Prasetyo,and A.R.Taufani,“Students Academic Performance Prediction with k-Nearest Neighbor and C4. 5 on SMOTE-balanced data,” In 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), IEEE, pp. 348-353, December 2020.

[34] J. Gajwani,and P.Chakraborty,“Students’ performance prediction using feature selection and supervised machine learning algorithms,” In International Conference on Innovative Computing and Communications, Springer, Singapore, pp. 347-354, 2021.