Early Prediction of Student Performance Using Deep Neural Networks

P. Sankara Rao [*1], V. Sai Snehitha [*2], P. N. D. Yamini [*3], P. Jagadeesh[*4], T. Aakash[*5]

[*]Department of Electronics and Communication Engineering, Raghu Engineering College (A), Visakhapatnam, Andhra Pradesh, 531162, India

Corresponding author: sankar.rec2016@gmail.com

Article Info
Page Number: 2130-2143
Publication Issue:
Vol. 71 No. 4 (2022)

Article History
Article Received: 25 March 2022
Revised: 30 April 2022
Accepted: 15 June 2022
Publication: 19 August 2022

Abstract

In contemporary educational systems, the ability to predict student performance is deteriorating gradually. Predicting a student's performance in advance can help students and teachers track a student's development. In today's world, a continuous evaluation strategy has been adopted by a great number of institutions. These kinds of systems are beneficial to students because they improve their overall academic performance. The encouragement of regular students is the goal of the continuous evaluation process. In recent years, neural networks have been implemented successfully in a variety of data mining applications, frequently outperforming traditional classifiers. In this paper, we used deep Neural Networks to predict student performance using Learning Management System (LMS) data within the context of Educational Data Mining. The various training features are derived from LMS data collected throughout the duration of each course and include usage statistics such as class test, assignment, attendance, and mid-exam marks in order to determine which of these factors are associated with the students' overall performance. The proposed method extracts informative data weighted appropriately. In this technique for generating neural networks, numerous updated hidden layers are utilized. These layers are determined based on feed forwarding and back propagation data from previous cases. Precision, recall, F1 score, accuracy and root mean squared error are used to evaluate the proposed method performance. On the basis of these results, we can conclude that deep neural networks outperform all existing state-of-the-art methods and could be used to accurately predict future student performance.

Keywords: Performance prediction, Data mining, CNN, deep neural networks.

1. Introduction

The process of extracting data from large databases is known as data mining. It reveals hidden data from various data sources pertaining to various fields. Data mining has applications in weather forecasting, oil research, business and medicine, marketing, and enterprise data management [1]. Educational data mining is a subfield of data mining that focuses on extracting and analyzing educational data sources' knowledge. The extraction of knowledge from educational environments through the use of EDM data involves the use of data mining, statistics, and machine learning. It is in high demand and gaining increased attention due to the expansion of educational data accessible through e-learning systems and the evolution of more conventional forms of education. Massive quantities of unprocessed data are combed through to discover information that can be used to advance and appreciate learning processes [2]. It is extremely concerned about the emergence of new techniques for identifying the unique types of data that can only be found in academic environments. In contrast to EDM, which provides solutions to additional problems such as "predict the students who are more likely to pass," conventional database records can be examined to answer questions such as "find the students who failed the exams”.

One of the most important applications of EDM in educational institutions is the improvement of student models in order to predict student characteristics or performances in advance. This is one of the most crucial EDM applications. Consequently, a large number of researchers began investigating various data mining techniques in an effort to aid educators or instructors in evaluating and enhancing the organization of their respective courses [3, 6].

Our current educational systems are becoming less accurate at predicting the academic performance of students. It is possible to maintain or improve the quality of education if it is possible to predict student performance in advance. This enables educators to maintain or enhance the quality of education by predicting student subject interests, student level activities, and aiding in the improvement of student performance in schools, universities, and other educational institutions. Identifying dropout points is another possibility with this method [4]. Several organizations are currently utilizing a continuous evaluation system that combines EDM with machine learning techniques. With the aid of these strategies, it is possible to improve student performance. Priority number one for any system of continuous evaluation must always be the well-being of regular students. The deployment of machine learning algorithms is accompanied by the development of pre-processing pipelines and data transformations. They make contributions to data representation to facilitate active machine learning and to highlight the deficiencies of existing learning algorithms [5].There were some survey results on some applications of Deep Learning that can be used in various fields such as image processing, natural language processing, and object detection [6].

2. Deep Learning

Conventional Neural Network (CNN): The ability of CNN's algorithm to recognize a wide variety of full behaviour’s has led to the algorithm's wide spread adoption across all subfields of the image recognition industry [7]. Consequently, its application was expanded to include the education and learning prediction procedure. In this sense, CNN is similar to neural networks in that it consists of multiple neurons connected in a hierarchical structure, and the layered structure is completed through training. CNN is utilized for the purpose of analyzing student behavior by means of the extraction of new features at specific time points; these features take into account characteristics that are derived from the student's educational background [8].

Recurrent Neural Network (RNN): It takes into account one of the neural network algorithms and performs exceptionally well with data sequences. This algorithm's ability to remember the previous state so that it can be utilized in either the current state or the next state [9] is one of its best characteristics. In addition to the hidden layer, a dynamic input and output layer also exists. There is input and output cases within the hidden layer, which are represented by the output weight moving from one node to another. As a result of connection and feedback paths in hidden layers, these algorithms are suitable for prediction, which is an advantage of their use in the training process.

Long Short-Term Memory (LSTM): Its model is characterized as a variant of the recurrent neural network (RNN). This algorithm derives its value from the creation of a self-loop within the hidden layers; in addition to generating a short path during each iteration, it also generates paths automatically while the system is running. It is comparable to DNN, but there is a difference in how the weight that influences the sort path in neural networks is updated. This type utilizes historical data to extract useful information (features) for more accurate prediction of student behavior [10].

Deep Neural Network (DNN): It is a neural network that conceals information in multiple layers. This model's performance is enhanced when dealing with complex data and nonlinear functions. The back propagation algorithm guides the training of this type of deep learning so that it can adapt to any changes made to the hidden layers while it is being trained. Because DNN performs well with scalable data in complex data-using prediction models, it is deemed suitable for education-related applications of deep learning prediction [11, 12].As depicted in Figure 1, deep learning is used to comprehend the data's behavior and make accurate predictions, so this information will be useful in the near future. Machine learning now has the potential to comprehend data and its behavior in complex systems as a result of advancements in data science and modern technology, such as big data and high-performance computers. The concept of machine learning refers to the ability of computers to acquire knowledge through the use of a variety of algorithmic processes, rather than through the use of a predetermined programme or a constrained set of instructions [13].

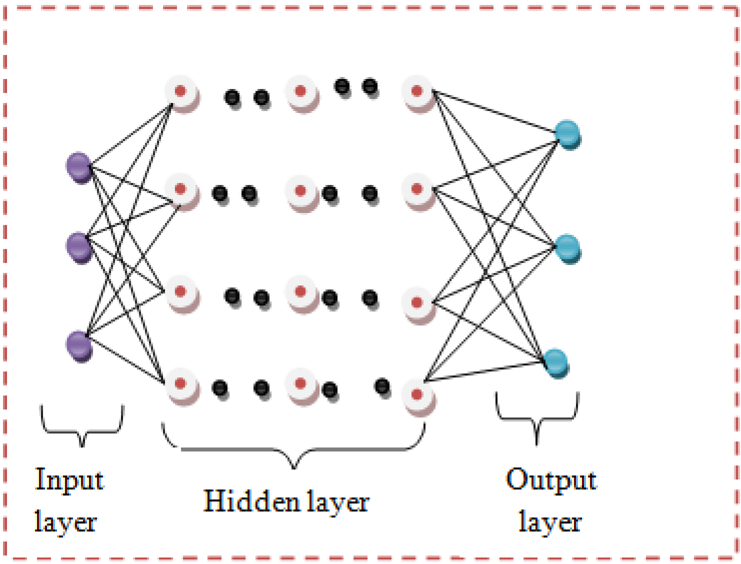

Figure 1. Deep neural networks architecture

Deep learning will automatically generate features in order to simulate the outcomes that are desired [14]. There are many hidden layers that contribute to decision-making, and one of the ways they do this is by using feedback from one layer to the layer below it to achieve a better result [13]. In order to maximize the efficiency of computers, DL allows them to perform complex calculations by combining simpler ones. Because it is difficult for a computer to comprehend complex data, such as a collection of data from the literature or a series of complex data, we employ deep learning algorithms instead of conventional learning methods [15].

3. Proposed Method

The research makes use of the data that was gathered during the training process for the proposed method in order to be able to predict the performance of the students on the final exam based on the marks they received on the class tests, assignments, and midterm exams for the respective courses. This is accomplished by making use of the data that was collected during the training process for the proposed method. After real data has been collected from the university/institution, it will be subjected to data pre-processing to remove redundant attributes, noise, etc. The data will then be separated into three distinct collections: the first dataset will be used for training, the second dataset will be used to test the proposed method, and the third dataset will be used to validate the proposed method. Deep learning models are constructed with the help of neural networks. A neural network receives inputs, which are then processed in the hidden layers using weights that are modified during the training procedure. The model will then produce a prediction for you. Adjustments are made to the weights to search for patterns in order to generate more precise forecasts.

Implementation steps: The following steps are necessary for the design of the proposed system.

- Loading the dataset
- Specifying the training and test data;
- Design and implement the network
- Separate data into training, testing, and validation datasets for network training
- Develop the network
- Passing arbitrary or test data through the network to evaluate it.

Figure 2 represents the general flowchart of the proposed work using DNN.

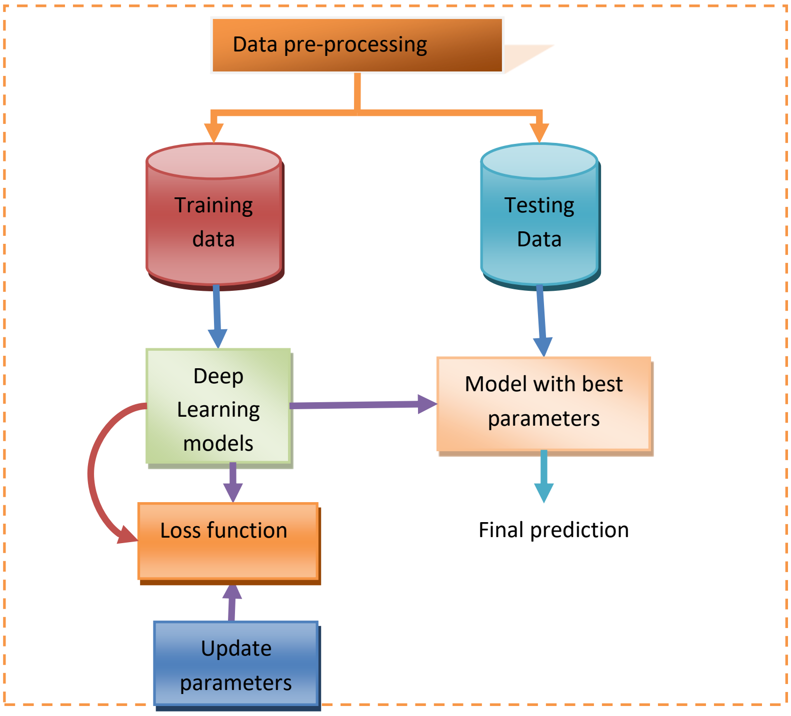

Figure 2. The proposed system's general flowchart.

Data loading: load the dataset using excel sheet, which contains main attributes such as the class tests, assignments, attendance, midterm exams and final score.

- ‘SpreadsheetImportOptions’ allows you to specify the manner in which MATLAB imports tabular data from spreadsheet files. The object's properties regulate the data import procedure, including the handling of errors and missing data.
- ‘readtable’ command used to make a table from a file.
- Separate the training data from the target data.

Training dataset: use ‘train network’ function to train the given data.

- ‘Trainlm’ optimizes network weight and bias using Levenberg-Marquardt. Trainlm is the fastest back propagation algorithm in the toolbox and the first-choice supervised algorithm, despite needing more memory.
- ‘Fitnet’ returns a neural network function fitting hidden Sizes.

Test Network: use ‘test network’ function to test the network.

- Predict function Perform predictive analytics by training a neural network or performing regression analysis on your data.

4. Evaluation Metrics

The confusion matrix is a type of matrix that is employed in the process of evaluating the efficacy of various classification models with regard to a specific collection of test data.

Where, TP= a significant number of these positive instances were correctly classified.
FP= number of incorrectly classified instances that constitute a positive instance
FN= a number of errors in identifying that an occurrence is negative
TN= number of instances correctly classified as negative

Table 1. ConfusionMatrix

	Actual values	
Predicted values	TP	FP
	FN	TN

The proposed model's performance was measured using a set of standard evaluation metrics, including precisions, recall, accuracy, and F-score. From table 1 calculate precision, recall, F1-score and accuracy.

Precision: The term "precision" refers to the percentage of students whose marks were correctly predicted out of the total number of students and their marks.
$$Precision(P) = \frac{TP}{TP+FP} \quad (1)$$

Recall: The term "recall" refers to the percentage of the data set's records that were correctly predicted.

$$Recall(R) = \frac{TP}{TP+FN} \quad (2)$$

F1-Score: The F-score represents the harmonic mean of both the recall and the precision ratings.

$$\mathrm{F1} - \mathrm{Score} = \frac{2*\mathrm{P}*\mathrm{R}}{\mathrm{P}+\mathrm{R}} \quad (3)$$

Accuracy: The term "accuracy" refers to the percentage of students whose marks were correctly identified relative to the total number of marks contained in the data set.

$$\mathrm{Accuracy} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \quad (4)$$

5. Experimental Results And Discussion

Experiments are conducted on MATLAB2021b.

Loading data: The data set was divided into training and testing with a 70:30 split. Experiments were carried out on our own dataset, which included class tests, assignments, attendance, and midterm exams as features. The findings demonstrated that the model, when utilizing all of the components, generated the optimal result.

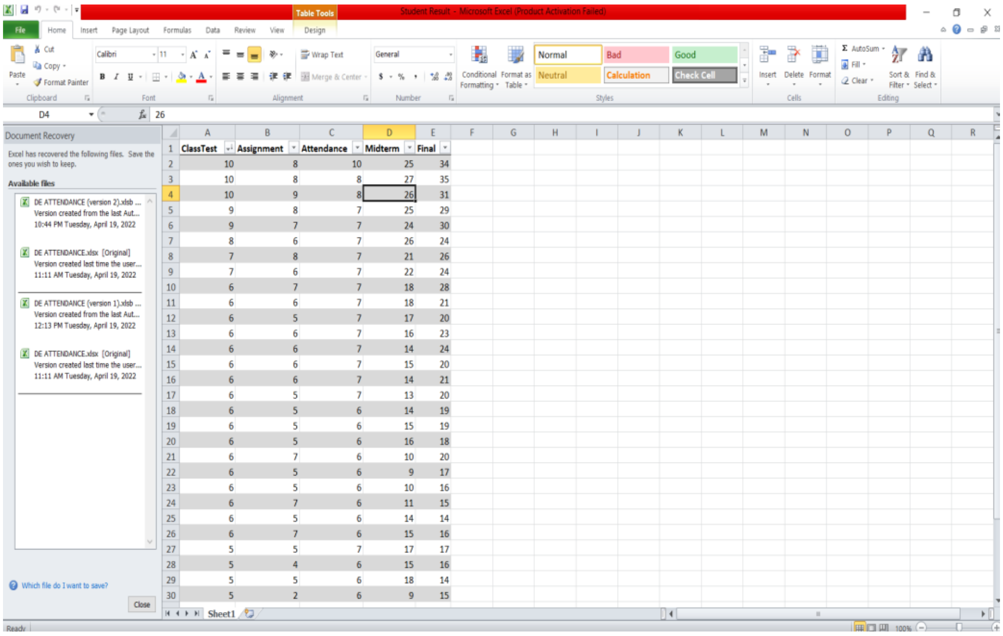

Figure 3. Student sample data (class test, Assignment, attendance, midterm marks)

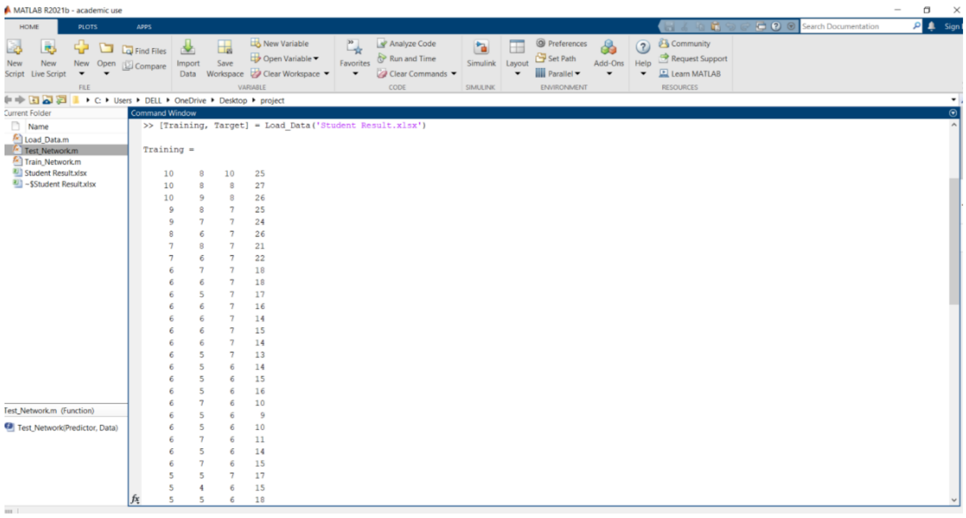

Figure 4. Load training data

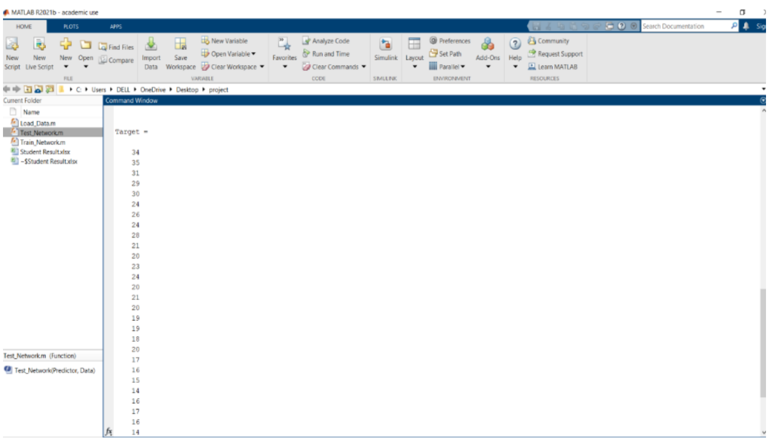

Figure 5. Load Target Data

Training data: The Levenberg-Marquardt algorithm is specifically designed to work with loss functions that take the form of a sum of squared errors. Levenberg-Marquardt algorithm is quickest one, but it typically requires a lot of memory.

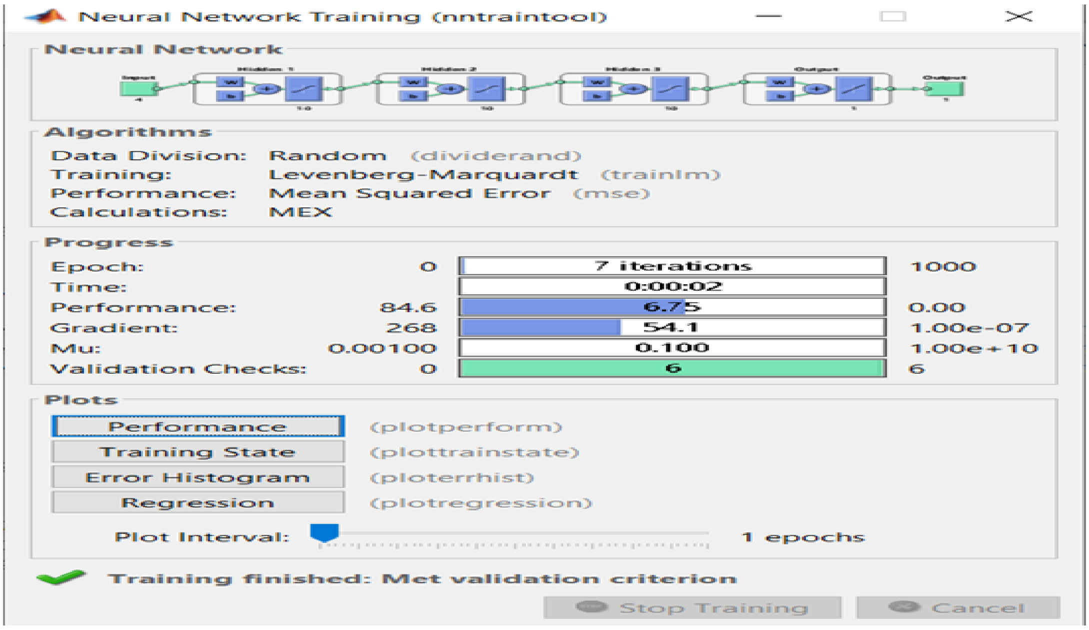

Figure 6. Training process

Epochs: The number of epochs is a hyper parameter that specifies how many times the learning algorithm will traverse the entire training dataset. Each sample in the training dataset has had the opportunity to update the internal model parameters during one epoch.

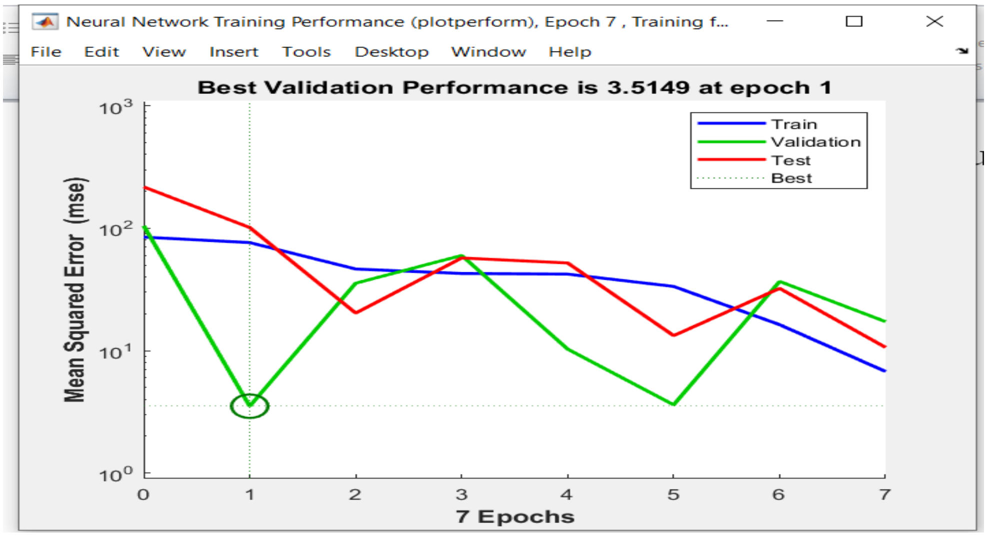

Figure 7. Regression in the training window

Gradient descent plot: A gradient measures the amount by which the output of a function changes when the inputs are altered slightly. Gradient descent is typically the slowest training algorithm, but also the one requiring the least memory.

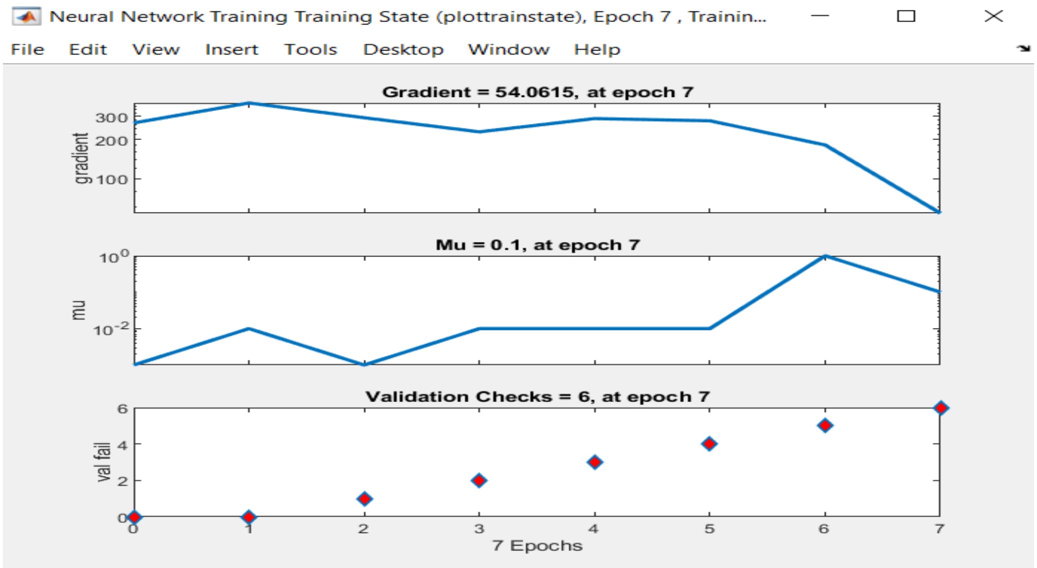

Figure 8. Gradient descent

MSE: The mean squared error (MSE) measures the average of the squares of an estimator's errors.

$$Mean\ square\ error = \frac{1}{N}\sum_{k=1}^{N}(x_k - \widehat{x_k})^2$$

(5)

MSE is utilized to determine how accurately estimates or forecasts reflect the current state of the market. The smaller the MSE, the more closely the forecast matches actual values. This is a measure that is used to evaluate models, and a lower value indicates a better fit for the data being analyzed by the regression model.

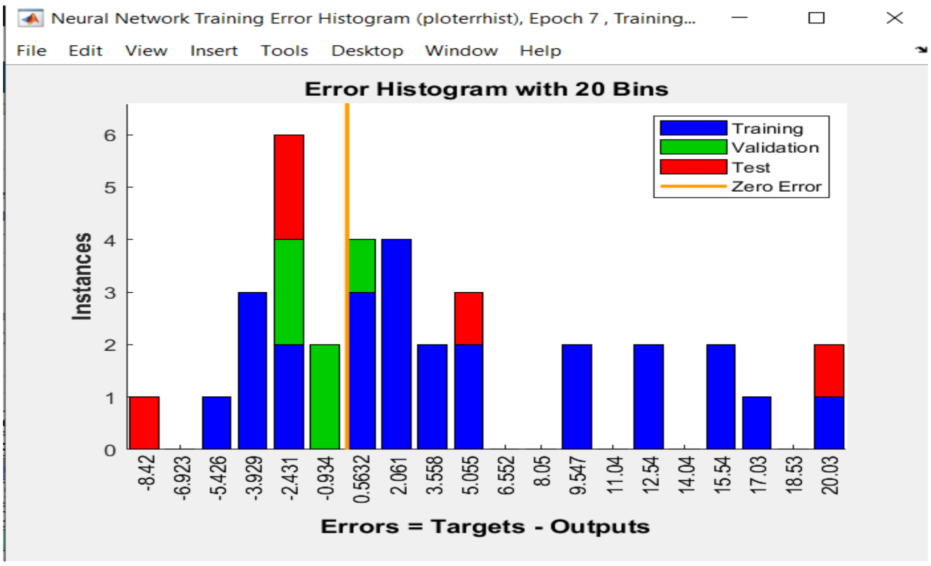

Figure 9. Mean Squared Error window.

Error Histogram: Figure 10 presents the error histogram that was compiled throughout the entire training process. The error histogram illustrates how errors between target values and predicted values can be visualized after training a feed forward neural network.

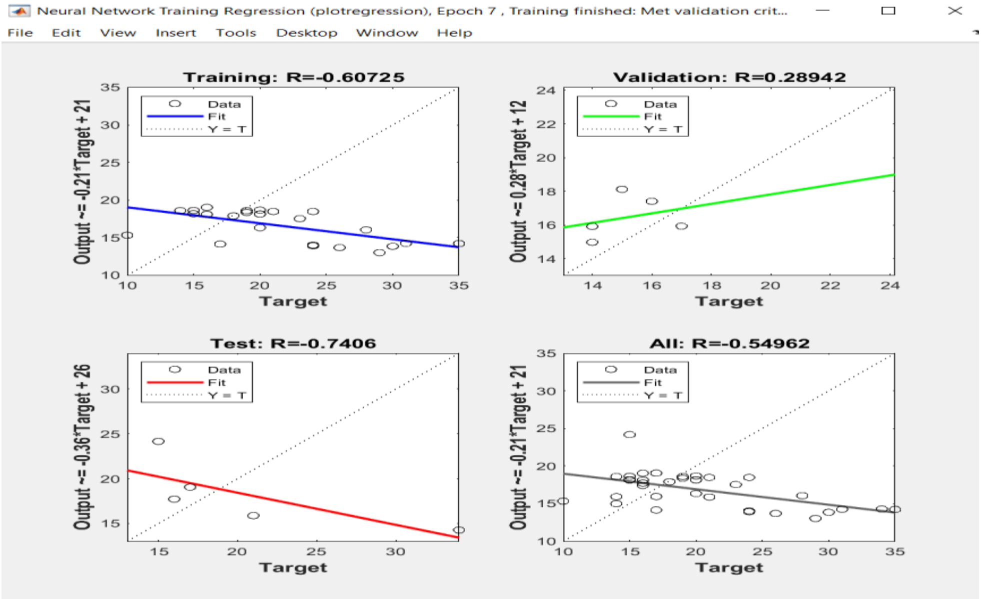

Figure 10. Training errors using a histogram

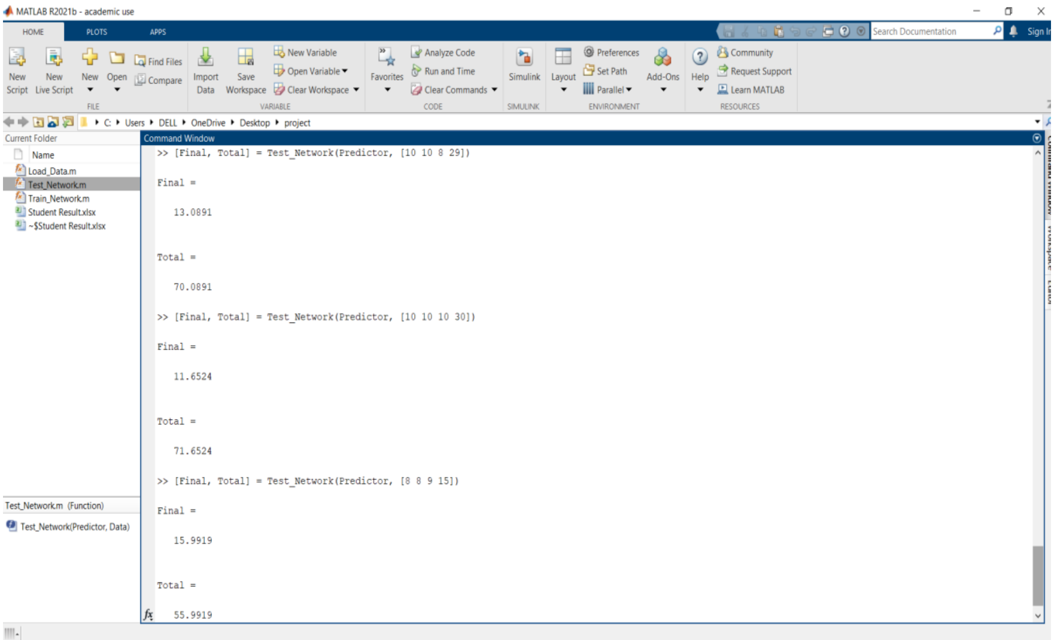

Figure 11. Testing window (Sample)

Testing Results:

Test1 result:

[Final, Total] = Test_Network(Predictor,[7 8 9 25])

Final = 13.5287

Total =62.5287

Test2 result:
[Final, Total] = Test_Network(Predictor,[5 4 3 10])
Final = 24.0767
Total = 46.0767

Test-3 result:
[Final, Total] = Test_Network(Predictor,[7 6 3 29])
Final = 14.1139
Total = 59.1139

Test-4 result:
[Final, Total] = Test_Network(Predictor,[10 10 10 30])
Final = 11.6524
Total = 71.6524

Test-5 result:
[Final, Total] = Test_Network(Predictor,[1 2 3 15])
Final = 20.6720
Total = 41.6720

Comparison between the proposed method and existing methods:

Table 2 shows the proposed model performance in terms of accuracy% to evaluate student performance and assessing levels as compare with other existing techniques. Our method achieved best result.

Table 2. Comparison between the proposed method and existing methods

S. No	**Method**	**Accuracy**
1	G. Gray et al.[16]	69.00%
2	S. T. Jishan et al. [17]	75.00%
3	Lubna Mahmoud Abu Zohair [18]	79.00%
4	Rodolfo C. Raga Jr et al. [19]	80.95%
5	**Proposed method**	**84.6%**

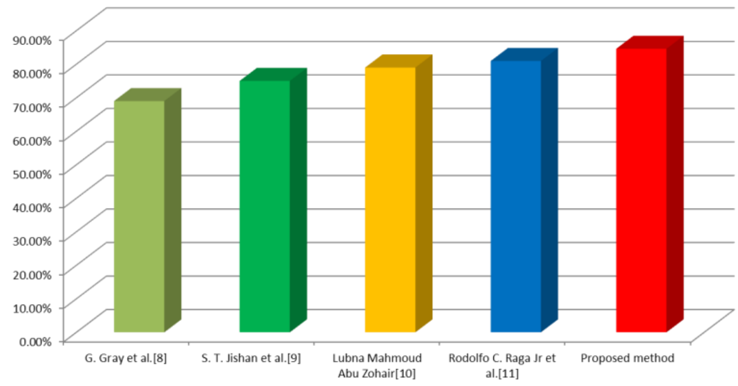

Figure 12. Accuracy % comparisons with other existing techniques

Table 3 shows the proposed model performance while utilizing the features that were available. It has an accuracy score of 84.6%, a precision score of 88.2 %, a recall score of 86.4 %, and an F-score of 87.2 %.

Table 3. Comparison of the proposed method with other existing methods [25]

Method	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F-Score (%)**
Recurrent Neural Network (RNN)	72.55	76	62	82
Convolutional Neural Network (CNN)	76.57	72	71	79
Long Short - Term Memory (LSTM)	74	89	74	84
Bidirectional Long Short-Term Memory (B-LSTM)	82.7	78	82	86
Proposed DNN method	**84.6**	**88.2**	**86.4**	**87.2**

6. Conclusion and Future Work

The accurate student academic performance prediction model is in high demand at educational institutions across the globe in the present day. Predicting the performance of students is primarily useful for the purpose of assisting teachers and students in improving their respective learning and teaching processes. The most difficult task, however, is frequently resolving the data quality issues that arise in student performance prediction models. The purpose of this research was to present a model for predicting student performance that was based on the deep learning technique. Experimentation revealed that a DNN can perform better with fewer data points if the dataset is well understood and the model's quality is improved. The proposed model achieved an accuracy rate of 84.60%. With larger dataset records and features, a Deep Neural Network can achieve greater accuracy and outperform other machine learning algorithms. This model is accurate and can assist in predicting a student's performance and identifying those with a greater likelihood of failing in advance so that remediation can be provided. In the future, the suggested model is going to be validated using a large dataset that contains a greater number of attributes.

References

[1] J. Han, J. Pei and M. Kamber, "Data Mining Concepts and Techniques." The Morgan Kaufmann series in data management systems, 3rd edition, 2011.
[2] O. Scheuer and B.M McLaren, "Educational data mining." In Encyclopaedia of the sciences of learning, 2012, pp. 1075–1079.
[3] C Romero and S. Ventura, "Educational data mining: a review of the state of the Art." IEEE Transactions on Systems, Man, and Cybernetics, 2010, Vol. 40, no. 6, pp. 601–618.
[4] O.D. Oyerinde and P.A. Chia, "Predicting Students' Academic Performances-A Learning Analytics Approach using Multiple Linear Regression." International Journal of Computer Applications, 2017, Vol. 15, no. 4, pp. 37-44.
[5] J. Schmidhuber. "Deep learning in neural networks: an overview." Neural and evolutionary computing, 2015; 61:85–117.
[6] J. Sultana, M. Usha rani and M.A.H Farquad, "An Extensive Survey on Some Deep Learning Applications." Emerging Research in Data Engineering Systems and Computer Communications. 2019; Vol. 1054, pp.511-519.
[7] M. Xu, "Image processing system based on FPGA and convolutional neural network." Microprocessors and Microsystems, 2020.
[8] J. Zong, C. Cui , Y. Ma , L. Yao , M. Chen and Y. Yin, "Behavior- driven student performance prediction with tri branch convolutional neural network." In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, 2020.
[9] Y. He, R.Chen , X. Li , C. Hao , S. Liu , G. Zhang and et al., "Online at-risk student identification using RNN-GRU joint neural networks." Information, 2020, Vol. 11, no.10, pp.1-12.
[10] H. Liu, Y. Zhu, T. Zang , Y. Xu, J. Yu, and F. Tang, "Jointly modeling heterogeneous student behaviors and interactions among multiple prediction tasks." Machine Learning, 2021,
[11] T. Samann , P. Schlicht and F. Huger, "Strategy to increase the safety of a DNN-based perception for HAD systems." Computer vision and pattern recognition, 2020,
[12] Y. Lin, H. Zhao, Y. Tu , S. Mao , and Z. Dou, "Threats of adversarial attacks in DNN-based modulation recognition." In Proceedings of IEEE Conference on Computer Communication, 2020.

[13] D. Prado and M. Lopez "Advances in Financial Machine Learning." John Wiley & Sons, Hoboken, NJ, USA. 2018.
[14] G. Sulong and A. Mohammedali, "Recognition of human activities from still image using novel classifier." Journal of Theoretical & Applied Information Technology, 2015, Vol.71, no. 1, pp.115–121.
[15] B.T. Atiyha , S. Aljabbar , A. Ali , A. Jaber, "Improved cost estimation for unit commitment using back propagation algorithm." Malaysian Journal of Fundamental and Applied Sciences, 2019, Vol. 15, no. 2, pp. 243–248.
[16] G. Gray, C. McGuiness, and P. Owende, "An application of classification models to predict learner progression in tertiary education." International Advance Computing Conference, 2014
[17] S.T. Jishan, R.I. Rashu, N. Haque, and R.M. Rahman, "Improving accuracy of students final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique." Decision Analytics, 2015, Vol. 2 , no. 1, pp. 1-25.
[18] L.M.A. Zohair, "Prediction of Student's performance by modelling small dataset size." International journal of educational technology in higher education, 2019.
[19] C. Rodolfo. Jr. Raga, D. Raga Jennifer. "Early Prediction of Student Performance in Blended Learning Courses using Deep Neural Networks." International symposium on educational technology, 2019.