IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, June 2025
Impact Factor: 7.67

Optimizing Educational Outcomes: A Deep Neural Network Approach to Predicting Student Success

Mohamed Samiulla Khan[1], Dand Hiren Jayantilal [2], Anjum Zameer Bhatt[3]
[1]Research Scholar, Department of CSE, Shri JJT University, Jhunjhunu, Rajasthan, India
[2]Research Guide, Department of CSE, Shri JJT University, Jhunjhunu, Rajasthan, India
[3]Research Co-Guide, Department of CSE, Shri JJT University, Jhunjhunu, Rajasthan, India
Corresponding Author: Mohamed Samiulla Khan: Email: mdsamiulla@gmail.com

Abstract: *In recent years, Deep Learning and Educational Data Mining have attracted significant attention. This paper presents a Deep Neural Network (DNN) model aimed at predicting students' academic performance from an early stage. This prediction enables educational institutions to identify students at risk of failing, allowing them to implement targeted interventions. The proposed DNN model was evaluated against existing machine learning algorithms using a subset of the OULAD dataset. Additionally, the impact of various activation functions on the model's performance was thoroughly analyzed to refine the results. The results show that the DNN model achieved an accuracy of 88.00% with the ReLU activation function, surpassing the performance of other machine learning approaches.*

Keywords: *Deep Learning*

I. INTRODUCTION

In the digital age, education has become increasingly data-driven, with large datasets offering valuable insights into student performance and learning behaviors. Educational Data Mining (EDM) and Deep Learning (DL) have emerged as powerful tools to harness this data and improve educational outcomes, particularly by identifying students at risk of underperforming.

The rise of online learning platforms has led to the collection of vast amounts of educational data, including interaction logs, assignment scores, and forum participation. While traditional methods often fall short in analyzing such complex, high-dimensional data, Deep Learning especially Deep Neural Networks (DNNs) offers a more effective approach by learning hierarchical patterns directly from raw data.

This study focuses on evaluating the impact of different activation functions within DNN models for predicting student performance. Among the functions tested, ReLU (Rectified Linear Unit) delivered the best results, with the DNN_ReLU model achieving an accuracy of 88.00%, outperforming all traditional machine learning models.

These findings highlight the potential of deep learning for early prediction of academic performance. By adopting such models, educational institutions can proactively identify struggling students and provide timely support, leading to improved outcomes and reduced dropout rates.

II. LITERATURE REVIEW

The application of machine learning and deep learning in education has gained substantial momentum in recent years, with a growing focus on predicting student performance to enable early interventions. Several studies have explored diverse approaches, datasets, and models, highlighting the evolving nature of Educational Data Mining (EDM).

Ahmed et al. proposed a deep learning framework using the OULAD dataset, achieving an accuracy of 81.29% and emphasizing the importance of data preprocessing and deep neural networks in virtual learning environments [1]. Similarly, Meka and Veeranjaneyulu reviewed various machine learning techniques, underscoring the significance of feature selection and student attributes in improving model accuracy [5][11].

Sokkhey and Okazaki conducted a comparative study involving traditional statistical methods, machine learning algorithms, and Deep Belief Networks. Their results indicated that the Random Forest algorithm delivered superior

ISSN 2581-9429 IJARSCT

performance in predicting mathematics achievement among high school students [8]. Chowdhury et al. demonstrated the effectiveness of ensemble methods such as Random Forest and Gradient Boosting [3].

Other research efforts have explored advanced model architectures. Reddy and Kumar introduced a hybrid approach combining decision trees and genetic algorithms, which improved accuracy through optimized feature selection and model tuning [7]. Thulasidoss et al. utilized TensorFlow-based deep learning models for predicting students' future development, showcasing adaptability across learning environments [9].

Vihavainen et al. applied deep learning to introductory programming courses, revealing the capability of such models to identify at-risk students early [10]. Krishna et al. emphasized the critical role of attribute integration and predictive modeling in enhancing educational outcomes [6].

Overall, deep learning models particularly Deep Neural Networks (DNNs) have shown strong potential in student performance prediction due to their ability to automatically learn complex patterns from high-dimensional data. The evaluation of various activation functions within DNNs further supports their adaptability and effectiveness in educational settings.

III. METHODOLOGY

This section outlines the comprehensive methodology used to develop and evaluate the Deep Neural Network (DNN) model for predicting student performance. The methodology encompasses the selection and preprocessing of data, the design and configuration of the DNN architecture, and the criteria for evaluating the model's performance. This systematic approach ensures that the model is robust, accurate, and generalizable to different educational contexts.

Dataset Description

The dataset utilized in this study originates from the Open University Learning Analytics Dataset (OULAD), which is widely recognized in educational research for its detailed records of student activity, demographics, and academic outcomes. The dataset was accessed from an Excel file, and it includes various fields that are critical to the analysis.

Key fields from the dataset include:

- **code_module**: Identifies the module code that the student is enrolled in, which helps in tracking academic progress within specific courses.
- **code_presentation**: Denotes the specific presentation of the module, typically corresponding to the academic term or year.
- **id_student**: A unique identifier assigned to each student, ensuring the ability to track individual progress across modules.
- **gender**: Categorical data representing the student's gender, important for analyzing potential gender-based performance trends.
- **region**: The geographical region where the student resides, which could influence academic performance due to regional factors.
- **highest_education**: The highest level of education attained by the student prior to the current enrollment, providing insight into their academic background.
- **imd_band**: The Index of Multiple Deprivations (IMD) band, indicating the socio-economic status of the student's area, which can be a critical factor in educational outcomes.
- **age_band**: The age range of the student, categorized into bands such as '0-35', '35-55', etc., allowing for the analysis of performance across different age groups.
- **num_of_prev_attempts**: The number of previous attempts made by the student in the same module, which can highlight persistence or recurring challenges.
- **studied_credits**: The total number of credits the student is studying, indicating their academic load.
- **disability**: A binary indicator of whether the student has declared a disability, potentially impacting their learning experience.

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 6, June 2025 Impact Factor: 7.67

- **final_result**: The target variable, representing the student's final outcome in the module, categorized as 'Pass' or 'Fail'.
- **date_registration**: The date on which the student registered for the module, which might be relevant in assessing engagement.
- **date_unregistration**: The date the student deregistered from the module, if applicable, potentially indicating early withdrawal.These features were selected based on their relevance to predicting student performance, with final_result being the primary target variable for classification.

Preprocessing Steps

The following preprocessing steps were applied to the student performance data to extract and engineer relevant features, enabling the development of predictive models:

IV. FEATURE EXTRACTION FROM STUDENT DATA

- Relative Score: Weighted sum of a student's assessment scores, reflecting performance relative to course expectations.
- Late Score: Count of assessments submitted late, indicating time management.
- Raw Score: Unweighted total of assessment scores earned.
- Sum Clicks: Total interactions with the LMS, showing overall engagement.
- Interactive Clicks: Clicks on interactive LMS features (e.g., forums), measuring active participation.
- Content Clicks: Clicks on course materials (e.g., readings, videos), indicating content engagement.
- Assessment Completion: Ratio of completed assessments, showing progress and commitment.

Feature Selection

To reduce the dimensionality of the feature space and enhance model performance, a feature selection process was applied using the SelectKBest method with mutual information as the scoring function. The top 10 most relevant features were selected for each dataset, ensuring that the models trained on the most informative variables.

Model Evaluation

A range of machine learning models were evaluated on the preprocessed datasets:

- **Logistic Regression**: A baseline linear model commonly used for binary classification tasks.
- **Support Vector Machine (SVM)**: A model that finds the optimal hyperplane for classification tasks.
- **Random Forest**: An ensemble method that uses multiple decision trees to improve accuracy and control overfitting.
- **K-Nearest Neighbors (KNN)**: A non-parametric method that classifies based on the majority vote of the nearest neighbors.
- **Naive Bayes**: A probabilistic classifier based on Bayes' theorem, suitable for categorical data.
- **Decision Tree**: A model that splits the data into branches based on feature values to make predictions.
- **XGBoost**: An advanced ensemble method that builds models sequentially to correct the errors of previous models.
- **Deep Neural Networks (DNN)**: Multiple DNN models with different activation functions (ReLU, Tanh, Sigmoid, Leaky ReLU, ELU) were tested to explore their effectiveness in capturing complex patterns in the data.

V. RESULTS

A variety of machine learning models were applied to the preprocessed dataset to evaluate their predictive performance. These included traditional models such as Logistic Regression, Support Vector Machine (SVM), Random Forest, K-

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, June 2025
Impact Factor: 7.67

Nearest Neighbors (KNN), Naive Bayes, Decision Tree, and XGBoost, as well as several Deep Neural Network (DNN) models using different activation functions.

Among all models, the **DNN with the ReLU activation function** achieved the highest accuracy at **88.00%**, demonstrating superior performance in capturing complex patterns in the data. Traditional models such as **SVM** and **Random Forest** performed reasonably well but were outperformed by the ReLU-based DNN. These findings underscore the potential of deep learning models—particularly those using ReLU activations—in delivering high predictive accuracy on educational datasets.

Model Performance on Dataset

Model	Accuracy
Logistic Regression	0.7684
SVM	0.8000
Random Forest	0.7754
KNN	0.7439
Naive Bayes	0.6421
Decision Tree	0.7053
XGBoost	0.7474
DNN_relu	0.8856
DNN_tanh	0.8689
DNN_sigmoid	0.8684
DNN_LeakyReLU	0.8695
DNN_ELU	0.8719

VI. CONCLUSION

The evaluation of various machine learning models on the dataset revealed that deep learning approaches significantly outperformed traditional methods in predicting student performance. Among all models tested, the DNN with ReLU activation achieved the highest accuracy at 88.56%, followed closely by DNN with Tanh and Sigmoid activations, both exceeding 86% accuracy. In contrast, traditional models such as SVM and Random Forest achieved lower accuracies.These results highlight the effectiveness of deep neural networks, particularly with ReLU activation, in capturing complex patterns within the data and provide strong evidence for their use in educational prediction tasks.

REFERENCES

[1]. Ahmed, M., Abdullah, S., Khan, A. A., & Khan, A. A. (2022). A Deep Learning Framework for Predicting the Student's Performance in the Virtual Learning Environment. *2022 International Conference on Advanced Learning Technologies (ICALT)*, Bucharest, Romania, pp. 180-182. doi: 10.1109/ICALT55010.2022.00060.

[2]. Bafna, G. J., Aziz, A. A., & Zaharim, A. (2022). Student Graduation Prediction Model using Deep Learning Convolutional Neural Network (CNN). *2022 International Conference on Advanced Learning Technologies (ICALT)*, Bucharest, Romania, pp. 1-6. doi: 10.1109/ICALT55010.2022.00060.

[3]. Chowdhury, A., Chahar, B. K., Thulasidoss, M. C., Fionalita, S., Vel, S., & Bafna, G. J. (2022). Student Performance Prediction using Ensemble Technique. *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, pp. 252-255. doi: 10.1109/ICACCS54159.2022.9785084.

[4]. Li, Y. (2021). Application of Xgboost Algorithm in Prediction of Students Grades Performance. *2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)*, pp. 93-96. doi: 10.1109/CEI52496.2021.9574453.

[5]. Meka, N. K., & Veeranjaneyulu, N. (2022). A Review on Student Performance Prediction Based on Machine Learning Techniques. *2022 International Conference on Advanced Learning Technologies (ICALT)*, Bucharest, Romania, pp. 1-6. doi: 10.1109/ICALT55010.2022.00060.

[6]. Rahman, M. A., Islam, M. K., & Khatun, M. T. (2023). Early Prediction of Students' Academic Performance Using Artificial Neural Networks: A Case Study in Computer Engineering Department. *2023 International Conference on Artificial Intelligence and Data Science (ICAIDS)*, pp. 120-125.

[7]. Reddy, P. R., & Kumar, V. (2022). Predicting Academic Performance of Students Using Modified Decision Tree Based Genetic Algorithm. *International Journal of Information Technology and Computer Science*, 14(2), 22-30. doi: 10.5815/ijitcs.2022.02.03.

[8]. Sokkhey, P., & Okazaki, T. (2022). Comparative Study of Prediction Models on High School Student Performance in Mathematics. *2022 International Conference on Advanced Learning Technologies (ICALT)*, Bucharest, Romania, pp. 1-6. doi: 10.1109/ICALT55010.2022.00060.

[9]. Thulasidoss, K., Hossin, M. N., & Vassileva, J. D. (2022). Prediction Model for Students' Future Development by Deep Learning and TensorFlow Artificial Intelligence Engine. *2022 International Conference on Advanced Learning Technologies (ICALT)*, Bucharest, Romania, pp. 1-6. doi: 10.1109/ICALT55010.2022.00060.

[10]. Vihavainen, F., Wei, S. L., Shehata, S., & Alizadeh, M. (2022). Using Deep Learning Models to Predict Student Performance in Introductory Computer Programming Courses. *2022 International Conference on Advanced Learning Technologies (ICALT)*, Bucharest, Romania, pp. 1-6. doi: 10.1109/ICALT55010.2022.00060.

[11]. Ying, L. (2021). Application of Xgboost Algorithm in Prediction of Students Grades Performance. *2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)*, pp. 93-96. doi: 10.1109/CEI52496.2021.9574453.

[12]. Zhang, Y., & Wang, H. (2022). Modeling and Estimation Methods for Student Achievement Recognition Based on XGBoost Algorithm. *Journal of Computer and Communications*, 10, 1-10. doi: 10.4236/jcc.2022.105001.

[13]. Zafar, M., & Khan, M. (2024). Predicting Academic Performance of Students Using Modified Decision Tree Based Genetic Algorithm. *2024 Second International Conference on Data Science and Information System (ICDSIS)*. doi: 10.1109/ICDSIS61070.2024.10594426