University of Zawia Journal of Engineering Sciences and Technology (UZJEST)
Volume 3, Issue 1, (2025), pp 1-11, ISSN 3006-2764

Evaluating Student Performance Based on Deep Learning Predictions

Ahmed Salamh[1], Ali Dowa[1]

[1]Department of Computer, Faculty of Education, University of Zawia, Zawia, Libya

*Corresponding author email: a.salamh@zu.edu.ly

Received: 24-02-2025 | Accepted: 07-05-2025 | Available online: 15-06-2025 | DOI:10.26629/uzjest.2025.01

ABSTRACT

Artificial intelligence has been widely used and attracted many researchers in the field of education, offering valuable data-driven insights that enhance educational decision-making. Traditional assessment methods have limited ability to manage the complex and interconnected factors that influence academic success. Therefore, this paper utilizes a synthetic student performance dataset that includes features such as study time, sleep hours, socioeconomic background, and class attendance, all of which have a direct or indirect impact on academic success. The goal of this paper is to predict student success using machine learning models, addressing the real-world challenge of predicting academic outcomes based on these variables. In this paper, we implemented evaluation metrics, including the loss curve, cross-validation, and callback functions such as early stopping to avoid overfitting and check the bias and the robustness of the model. Therefore, the model achieved significant results with an accuracy of 97.62% and an average accuracy of 95% in cross-validation, indicating its strong predictive capabilities. Moreover, the model's performance demonstrates that it can be implemented for institutions to understand the factors that contribute to student success and predict future academic outcomes. By providing accurate predictions, this method enables institutions to implement correct interventions for at-risk students, introducing a more personalised and effective learning environment. Ultimately, the findings suggest that machine learning can enhance institutional practices, improve student success rates, and evolve educational strategies for more effective outcomes. Additionally, the study emphasises the importance of feature selection and addressing dataset biases to ensure the model's fairness and generalisation capabilities.

Keywords: Student performance, Deep learning, Artificial intelligence.

تقييم أداء الطلاب بناءً على تنبؤات التعلم العميق

أحمد سلامة[1]، علي دوة[1]

[1]قسم الحاسوب، كلية التربية، جامعة الزاوية، الزاوية، ليبيا

ملخـــــــص البحـــــــث

لقد تم استخدام الذكاء الاصطناعي على نطاق واسع وجذب العديد من الباحثين في مجال التعليم، حيث قدم رؤى قيمة تعتمد على البيانات والتي تعزز عملية اتخاذ القرار التعليمي. إن طرائق التقييم التقليدية لديها قدرة محدودة على إدارة العوامل

المعقدة والمترابطة التي تؤثر على النجاح الأكاديمي. لذلك، تستخدم هذه الورقة مجموعة بيانات أداء الطلاب التي تتضمن ميزات مثل وقت الدراسة وساعات النوم والخلفية الاجتماعية والاقتصادية وحضور الفصول الدراسية، وكلها لها تأثير مباشر أو غير مباشر على النجاح الأكاديمي. الهدف من هذه الورقة هو التنبؤ بنجاح الطلاب باستخدام نماذج التعلم الآلي، ومعالجة التحدي الحقيقي المتمثل في التنبؤ بالنتائج الأكاديمية بناءً على هذه المتغيرات. في هذه الورقة، قمنا بتنفيذ مقاييس تقيم والتي تتضمن منحنى الخسارة **loss curve** والتحقق المتقاطع **cross validation** ووظائف الإرجاع **callback function** مثل التوقف المبكر لتجنب فرط التخصيص والتحقق من التحيز ومتانة النموذج. لذلك، حقق النموذج نتائج مهمة بدقة 97.62% ومتوسط دقة 95% في التحقق المتبادل، مما يشير إلى قدراته التنبؤية القوية. وعلاوة على ذلك، يوضح أداء النموذج أنه يمكن تنفيذه للمؤسسات لفهم العوامل التي تساهم في نجاح الطلاب والتنبؤ بالنتائج الأكاديمية المستقبلية. من خلال توفير تنبؤات دقيقة، تمكن هذه الطريقة المؤسسات من تنفيذ التدخلات الصحيحة للطلاب المعرضين للخطر، وتقديم بيئة تعليمية أكثر تخصيصًا وفعالية. في النهاية، تشير النتائج إلى أن التعلم الآلي يمكن أن يعزز الممارسات المؤسسية، ويحسن معدلات نجاح الطلاب، ويطور استراتيجيات تعليمية لتحقيق نتائج أكثر فعالية. بالإضافة إلى ذلك، تؤكد الدراسة على أهمية اختيار الميزات ومعالجة تحيزات مجموعة البيانات لضمان عدالة النموذج وقدراته على التعميم.

الكلمات الدالة: أداء الطلاب، التعلم العميق، الذكاء الاصطناعي.

1. Introduction

The correct evaluation of students' learning outcomes is one of the main issues facing every nation's educational system. Education is essential in shaping life, and the integration of artificial intelligence into traditional teaching methods is transforming higher education institutions to improve academics overall [1]. The performance of students is always of utmost importance to educational institutions, and there are many research studies performed to evaluate the performance of students in universities [2]. Student performance modelling in educational data mining (EDM) is both challenging and popular due to the complexity and multitude of factors that influence performance in non-linear ways [3]. These factors can include academic behaviours, social interactions, psychological traits, and more [4]. Analysing performance, providing high-quality education, formulating strategies for evaluating the students' performance, and identifying future needs are some challenges faced by most universities today [5]. Grades have long been considered a measure of a student's academic performance; however, various factors, including a student's socio-economic background, dietary habits, sleep patterns, and attendance, also impact their performance. Performance evaluations are typically administered with manual technologies through assessments, examinations, and grades, yet this method of evaluation is greatly lacking in areas such as predicting a student's performance and identifying students who may be at risk during their formative years [6]. On the other hand, machine learning has transformed possibilities as chances of automating the processes of evaluating students alongside improving prediction measures, which would in turn help educators and administrators alike with their decision-making processes.

Managing educational institutions is particularly challenging for managers because of the complexity of data structures, the multitude of data sources, and the large volume of data. This makes it difficult to streamline processes and ensure effective decision-making [7]. The traditional educational model typically assesses students' performance through basic, state-mandated measurements, such as grades on single tests, assignments, and final exams. Moreover, the performance in the real world is shaped by many dynamic, interrelated factors that may not always be reflected in traditional measures. Some of these factors might be how long a student studies for, how many hours of sleep, their socioeconomic background or how often they actually attend classes. Machine learning models can produce more comprehensive and nuanced predictions of student performance by taking into account all of these factors.

Recent advancements in technology, particularly in artificial intelligence and data mining, have indeed transformed student learning behaviours [8]. These technologies enable students to learn more efficiently with greater satisfaction. Machine learning is a branch of artificial intelligence which concerned with creating algorithms capable of learning from data and making predictions or decisions autonomously from programming instructions [9,10]. Machine learning algorithms applied to student performance prediction can utilize a range of input features to understand the student's progress and to maximize academic achievement. Being able to anticipate how actions may impact future results, such as diagnosing potential problems early, providing feedback for improvement, and personalizing learning, are just a few examples of how this may benefit education. Using machine learning to evaluate student performance is definitely an outstanding advantage, as it is capable of processing big data with a multitude of features. The major difference with the traditional approach is that machine learning can find hidden correlations and interactions between variables that may be very complex. A student's performance may depend on both the study time and the amount of quality sleep they get, and the machine learning model is the one that will define the relationship between these variables. Moreover, machine learning also allows for the models to continually develop as new data is collected. The addition of more students and the addition of new variables will make the model training possible so that the predictions can be refined and the model's capacity to adjust to variable conditions in the educational environment can increase. This is especially of great importance for long-term projects or when the student community is changing.

Integrating deep learning into the evaluation of student performance in any institute or university offers significant advancement. However, it requires overcoming various practical challenges. Therefore, addressing issues such as data privacy, resource limitations, and bias is important for the successful implementation of these technologies. Moreover, it is essential to integrate deep learning in a way that complements and enhances the university's role, ensuring that AI serves as a tool to support students at risk. With careful planning, support, and ongoing adaptation, deep learning can revolutionise the way student performance is assessed and improve educational outcomes.

Predicting student performance remains a complex task due to the challenges in accurately assessing individual talents and efforts to enhance academic performance [11]. Researchers face difficulties analysing student performance outcomes because educational databases often consist of vast amounts of data, making it hard to train models with small sample sizes. The large educational datasets introduce their challenges, such as complexity, noise, and scalability. On the other hand, small sample sizes can be even more problematic, especially when training deep learning models. However, we can still make progress by carefully selecting and tuning models, using techniques like cross-validation based on dataset needs in order to overcome the limitations of small data. In the end, the key lies in balancing the model complexity and data availability.

2. Related Works

There are many studies about predicting the outcome of students using machine learning algorithms such as Decision Trees (DT), Naive Bayes (NB), Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Sequential Minimal Optimization (SMO), and Neural Networks [12]. Table 1, gives a summary of related works. In an online learning environment, the group researcher aimed to use data mining (DM) methods to detect students at risk. They tested four algorithms, including KNN, DT, NB, and Neural Networks, and KNN gave the highest accuracy of 87%. The study concentrates on the performance of students in two subjects OOP and SPL implementations of traditional machine learning algorithms, achieving major good results with SVM, which was able to achieve an accuracy of about 95% in the academic data. Even when the model is trained and tested without academic data, it still performs well, reaching an accuracy of 93% [13]. The researchers integrated neutrosophic theory, which accounts for truth, falsity, and indeterminacy, with deep learning, which excels at understanding complex data patterns. Accuracy in student data classification (attendance, grades) was very high, with a percentage of 95.00, showing promise for identifying high-risk students and follow-on interventions [14]. Some papers applied machine learning approaches to predict student performance from features such as gender, family income, board results, and attendance. Computational methods used in the test: C4.5, SMO, NB, 1-Nearest Neighbour, and Multi-Layer Perceptron (MLP). From this set, SMO gave the best average test accuracy at 66% rather than other models [15]. The GRSO algorithm-optimised hybrid deep learning Convolutional Recurrent Network (CRN) scheme improves the classification performance. Using GRSO on CNN and RNN hyperparameters tuning, the model has a sensitivity of 94%, accuracy (off the chart) 93% [11]. The simulation results indicate that the GRSO-CRN model has better performance. As evident, the SVM-SMOTE approach leads to better results, and XG-Boost performs best in this context.

The AI/ECOS+XG-Boost hybridization produced accuracy and an F1-score 94.17, 94.13 [16]. The discovered result shows the ability to integrate machine learning techniques with metaheuristic algorithms that provide accurate and fast prediction for student performance classification rules for educational administrators to use for an improvement on how the education system works.

Table 1. Represents the summary of the related works.

Methods	**Accuracy%**
DT [13]	87.00
Neural Networks[12]	87.00
KNN[12]	87.00
LR [13]	93.00
SVM [13]	95.00
Neutrosophic Deep Learning Model [14]	95.00
SMO[15]	66.00
(RNN) with GRSO[11]	93.00
Ecosystem + XG-Boost[16]	94.17
SVM_ best prediction[1]	96.00

Depending on their goals and theories, researchers focus on different features. That is, one paper trying to predict student performance would weight features like past grades, study habits and extra-curricular activities more heavily than a paper trying to predict the mental health of students, which would weight social behaviors, stress levels, or family background as more important features. Depending on the research aim, the domain or field of study, and the type of machine learning model used, the approach used will vary. However, each method might handle different characteristics within the data; therefore, researchers have to select features cautiously so that they reflect appropriately on the model and task they intend to solve. In machine learning, the algorithms may vary in their effectiveness depending on the type and quality of the dataset. Deep neural networks can perform exceptionally well in processing complex and diverse data. On the other hand, support vector machines are effective in clear cases, while decision trees and random forests offer greater flexibility in dealing with unclear datasets while maintaining interpretability. Working with the appropriate algorithm depends on the research goals. For binary classification, such as determining if a student will perform well, algorithms like SVMs or decision trees are effective. For more comprehensive predictions of overall student performance, methods like neural networks may be more appropriate. Ultimately, understanding the context and goals is basic for applying AI models successfully in education.

3. Research method

This paper uses machine learning for the prediction of student performance (a binary classification of grades where >50 is Pass and ≤50 is Fail for students). This paper proposed a model for student performance prediction and offers a more trustworthy performance evaluation of the model via cross-validation [4], by splitting data into **n** essentially identical folds, training the model on k−5 folds and testing on the rest to test if the model has learnt more and tested on all data. In this paper, we implement effective techniques to mitigate bias and overfitting and optimise model performance. Moreover, the use of cross-validation ensures robust evaluation of the model's generalisation ability, while the analysis of the loss curve suggests that the model is fitting the data appropriately. Therefore, these measures indicate satisfactory results; they give correct evidence to indicate that the approach is well prepared. Additionally, the callback function to monitor training and fine-tuning of the model parameters is a basic part that we use to ensure sustained performance improvements, ultimately leading to robust and reliable results.

3.1 Students Dataset

The main goal of this methodology is to develop and test a predictive model on the students' performance dataset of 1,388 instances [17]. These samples consist of different student attributes like study habits, sleep patterns, socioeconomic factors and class attendance known as features in order to create a prediction for a binary pass or fail outcome of a student. It consists of two main steps. First, training of the model and second, evaluation of the performance by different metrics: accuracy, precision, recall, F1 score, ROC curve and AUC.

Educators and institutions could use machine learning models trained on this dataset as a resource to assess what factors drive student success, as well as for predicting academic outcomes. Instituting interventions based on performance metrics and features that have the biggest impact on grades allows institutions to more precisely support students in need of support, thus improving educational outcomes. Table 2, displays dataset information, including mismatched, missing data, mean and standard deviation. The data shows no missing or mismatched values for the variables. On average, students study for 4.56 hours, sleep for 8.05 hours, and have an attendance rate of 58.5%. Their socioeconomic score averages 0.55, and their grade percentage is 40.7%. The standard deviations for these variables are 1.9 for study hours, 1.37 for sleep hours, 11.7% for attendance, 0.26 for the socioeconomic score, and 9.46% for grades, indicating varying levels of consistency across the factors.

Table 2. The summary of dataset information, including mismatched, missing data, mean and standard deviation

Features	**Mismatched**	**Missing**	**Mean**	**Std. Deviation**
Study Hours	0	0	4.56	1.9
Sleep Hours	0	0	8.05	1.37
Attendance (%)	0	0	58.5	11.7
Socioeconomic Score	0	0	0.55	0.26
Grades (%)	0	0	40.7	9.46

3.2 Dataset Pre-processing

More than features on student performance are present in the dataset. The last column (students' grades%) is the target variable, and other columns are the features, including this habit-study hours, sleep hours, present class attendance, etc. A target of student grades (continuous) is turned into a binary classification target. If the grade of the student is less than 50, fail (label = 0); else pass (label = 1). Table 3, shows the difference between the two labels in the Grades feature, with only 272 students having grades greater than or equal to 50 (label = 1) compared to 1,116 students with grades less than 50 (label = 0). In contrast, the Socioeconomic Score feature is more balanced, with 762 students having a score greater than or equal to 50 (label = 1) and 626 students with a score less than 50 (label = 0).

Table 3. Summary of Features pre-processing

Feature	**Bigger than or equal to 50 (label = 1)**	**Less than 50 (label = 0)**
Grades	272	1,116
Socioeconomic Score	762	626

3.3 Performance evaluation

To compute the performance of the proposed model using machine-learning techniques, the most commonly used Metrics in this work appear in equations (1), (2), (3), (4) and are defined as follows: TP (True Positives) is the number of correctly predicted positive instances. TN (True Negatives) is the number of correctly predicted negative instances, FP (False Positives) is the number of incorrect positive predictions, and FN (False Negatives) is the number of incorrect negative predictions. Area under the curve (AUC) and the Receiver Operating Characteristic curve (ROC) are considered for further analysis of model performance.

1. Accuracy (Acc) indeed serves as a good metric to assess the performance of a classification algorithm. It is computed as the ratio of the number of correct predictions to the total number of predictions. Accuracy can be mathematically expressed as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

2. Recall shows the samples positively classified in the total number of positive samples.

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

3. Precision gives the proportion of positive samples classified correctly in the total number of positive predicted samples.

$$PPV = \frac{TP}{FP + TP} \quad (3)$$

4. F-score represents the harmonic mean of precision and recall, and high values of F-score show high classification performance.

$$\text{F1\_score} = \frac{2TP}{2TP + FP + FN} \quad (4)$$

3.4 Proposed Machine Learning Model

Deep learning is similar to teaching a computer to think like an analyst [18]. Instead of just following hard-coded instructions, it learns from data to uncover patterns and make decisions on new, unseen data. Machine learning can be used in different areas in the field of education to evaluate student performance based on collected data [19- 22]. Different models can be utilised to predict student performance.

We present a modified model based on the ANN and CNN that is used for churn prediction[23] which is designed for different fields related to customer retention and attrition in various industries. The CNN model expands on this ANN by adding a convolution layer at the input stage with 30 filters (3×3), and three fully connected layers (20×20×20) are connected to the output of the convolution layer. The ANN model showed that the deep learning-based models gave much better results than the other existing models in this field, with 97.62% accuracy of the CNN model in terms of success in the classification and prediction. This combined with other evidence that helps us to suggest a new model.

In this paper, the model designs for feature input consist of several layers for classification. This model includes a convolutional layer, which applies 25 filters of size 3x3 to the input features as shown in Figure 1. The convolutional layer is followed by batch normalization and a ReLU activation. These steps help in extracting meaningful features from the input, stabilizing the training, and introducing non-linearity. After the convolutional layer, the model transitions to fully connected layers for further processing and classification. A fully connected layer (fc2) with 20 neurons, followed by a ReLU activation. Another fully connected layer (fc3) with 5 neurons processes the features further, with a ReLU activation applied afterwards. The output layer, which has many neurons equal to the unique labels in the training data, generates raw outputs for classification. The softmax layer converts these outputs into class probabilities, and the final classification layer assigns the input to the appropriate class based on the probabilities.

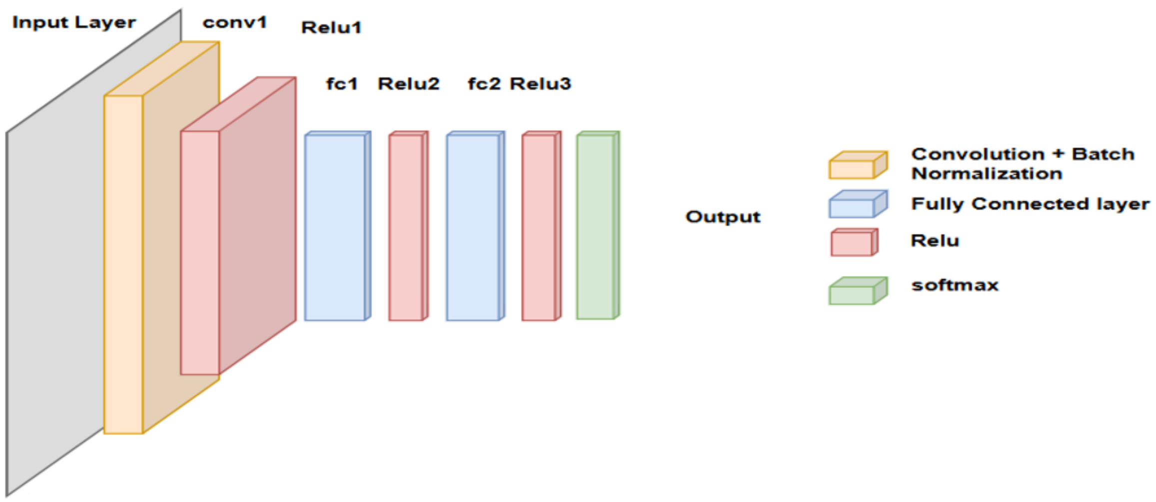


Figure 1. CNN layers of the proposed deep learning model for student performance prediction

4. Training model

The model was trained using Google Colab with an NVIDIA Tesla T4 GPU featuring 16GB of memory, a batch size of 128 and Adam optimizer for training. The model will be saved as the best version in the case that it achieves a higher validation accuracy than the previous. The model is then trained on the training data for 200 epochs and using the checkpoint callback function to save the best version of the model based on validation accuracy to ensure that the best results are saved. This approach helps save time and avoid overfitting.

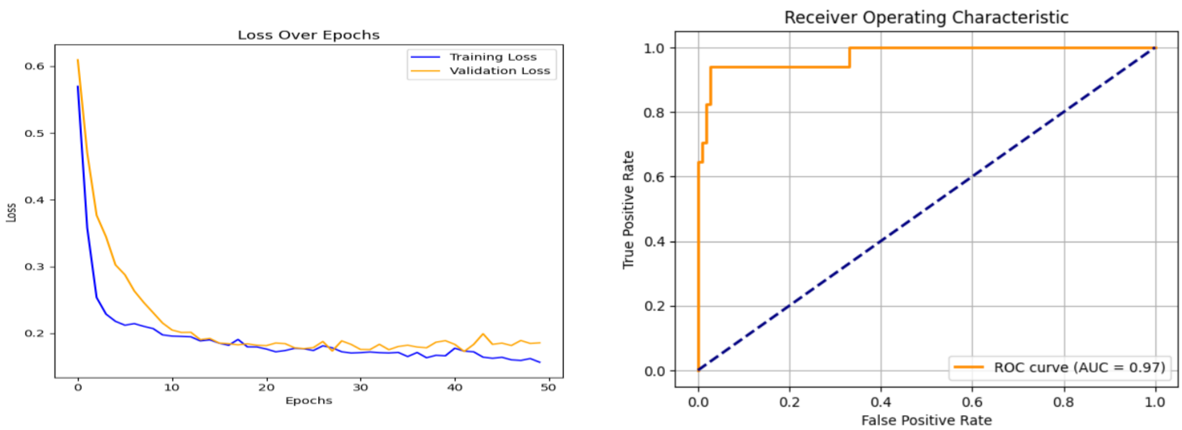
The dataset is further split into training and test sets (to guarantee a valid evaluation of the model) to be exploited by the two main split functions for datasets. First, divide the data set into an 80%-20% training set and test set. This is to mimic the synthetic situation where the model is trained with one subset of the dataset and tested with another, unseen subset of the dataset. The second time, cross-validation is used to improve the robustness of the model performance estimation. The dataset is split into 5 roughly equal folds, and the model is trained and tested 5 times using folds of differing data. A single fold is used for testing, and the others in which used for training. This is to reduce the risk of overfitting, bias and confirm stable model performance on different parts of the data. This is especially useful for small datasets, such as the 1,388 samples of this experiment, and allows to use of the maximum of available data while minimising the bias from a single random train-test split.

The model was trained on the dataset-based features like study hours, sleep hours, attendance, socio-economic status, which finally predict a binary target variable (Pass/Fail). It is a very interpretable and understandable model that fits perfectly for prediction like this. The results are presented in Table 4.

Table 4. Represents the results of the best version of the model

Accuracy%	**Precision%**	**Recall%**	**F1 Score%**	**AUC%**
97.62	96.00	94.00	95.00	97.00

Accuracy of 97.62% measures the overall correctness of the model, indicating that about 97.62% of predictions were accurate. Precision of 96.00% measures the model's ability to correctly predict positive instances. Precision means that out of all the positive predictions made, 96.00% were correct. A recall means that the model correctly identified 94.00% of all actual positive cases. An F1 score of 95.00% indicates a good balance between precision and recall. Area under the curve (AUC) of 97.00% represents the model's ability to distinguish between classes, as appears in Figure 2, which is a good fit.



a b

Figure 2. Depicted student performance evaluation. a) Training loss, validation. b) Represents the Receiver Operating Characteristic curve

Cross-validation is a technique used to assess the generalizability of a model by dividing the dataset into several folds (in this case, 5). The model is trained on each subset and tested on the remaining data. Figure 3 shows that the average accuracy across 5 folds is 95.00%, which means the model performs consistently.

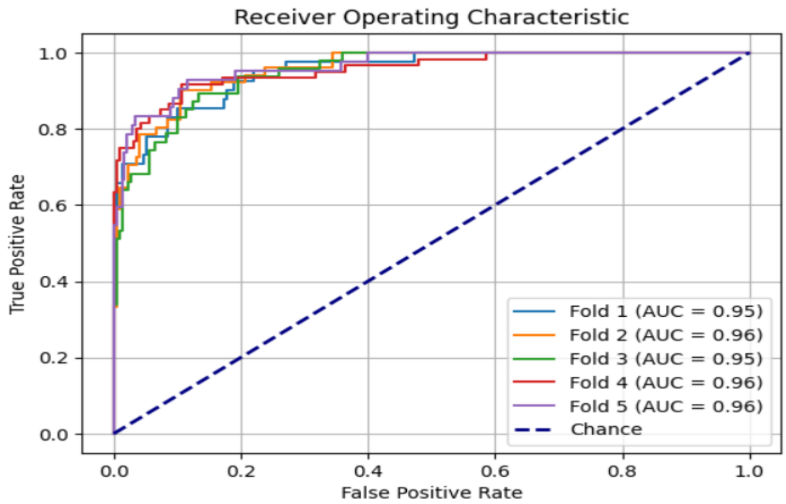


Figure 3. The True Positive Rate against the False Positive Rate for 5-fold cross validation

5. Results and Discussion

The model predicts student performance with the number of predictor variables: study hours, sleep hours, attendance percentage, and socioeconomic status. Within all the evaluation metrics, the model performs pretty well, as shown by the notably high accuracy of 97.62% and AUC of 97.00%. The model can have a good recall of 94.00%, its precision of 96.00%, and an F1 is a strong indicator to tell that this model performs well. The additional cross-validation step reinforces the model's reliability, making it well-suited for deployment in real-world applications. The number of students with grades greater than or equal to 50 (label 1) is much smaller (272 vs. 1,116), the model might tend to predict the majority class (label 0, less than 50) to improve overall accuracy. This can lead to a higher number of false negatives, thus lowering recall for the minority class (those with grades >=50). Grades have a notable difference between the two classes, with many more students having grades less than 50. This is the main reason that recall is lower because the model is biased towards predicting label 0.

Table 5, gives the accuracy comparisons that this paper's model is highest in general contexts (97.62%) compared with other approaches. The SVM (used in an OOP course) accuracy is 95.00%, SVM_ best prediction gave 96.00%, Decision Trees (DT) and Logistic Regression (LR) got both 93.3% in the same course. The Neutrosophic Deep Learning Model also reached 95.00% accuracy in general applications. The average accuracy from the model in this paper was 95.00%, matching the performance of the Neutrosophic Deep Learning Model; nevertheless, the best result was the highest among all compared algorithms.

Table 5. The comparison of the proposed model and the state-of-the-art models

Methods	Accuracy%
SVM [13]	95.00
DT [13]	93.3
LR [13]	93.3
Neutrosophic Deep Learning Model [14]	95.00
(RNN) with GRSO[11]	93.00
Ecosystem + XG-Boost [16]	94.17
SVM_ best prediction [1]	96.00
This paper (the best prediction)	**97.62**
This paper (average)	**95.00**

Further improvements could be made by tuning the model to increase recall without sacrificing precision, thereby achieving a more balanced performance across all metrics. However, it's important to address challenges such as potential bias, data quality, and model interpretability to ensure fair and accurate predictions. By expanding features, conducting longitudinal analysis, and integrating the model within educational tools can further enhance its impact.

6. Conclusions

This paper proposes a deep learning model that is developed to learn from important features in order to predict student performance. The model showed strong performance using several evaluation techniques, including accuracy, indicating robustness and generalisability of model performance. The proposed model outperforms almost all models in terms of accuracy and achieves the best prediction accuracy of all other models, with the best accuracy of 97.62% and average accuracy of 95.00%. This helps the early identification of students at risk, enabling the creation of personalised learning interventions in order to provide correct recommendations for students and the optimisation of outcomes. In addition to these benefits, the findings of this paper could feed into university-wide strategic efforts; therefore, this allows the university to improve educational programs, prioritize interventions and create impactful engagement systems. The results of the deep learning model are evidence that supports the possibility of using data to confirm and improve not only student performance but also the university's comprehensive plans for achieving educational success in the future. Moreover, based on the dataset analysis, this paper provides an excellent foundation for how to create a dataset for evaluating student performance in any sort of institution.

Completing feature sets, performing additional longitudinal studies, and collaborating the model with educational media are some of the changes that can be applied to achieve this goal and consequently improve educational outcomes and aid in the development of the students.

REFERENCES

[1] E. Ahmed, "Student Performance Prediction Using Machine Learning Algorithms," *Appl. Comput. Intell. Soft Comput.*, vol. 2024, no. 1, p. 4067721, 2024.

[2] Z. Iqbal, J. Qadir, A. N. Mian, and F. Kamiran, "Machine Learning Based Student Grade Prediction : A Case Study," pp. 1–22.

[3] G. Zhang, "Construction of a Student Learning Behavior Prediction Model Based on Decision Tree Algorithm," in *Proceedings of the International Conference on Modeling, Natural Language Processing and Machine Learning*, 2024, pp. 322–327.

[4] A. Khan and S. K. Ghosh, "Student performance analysis and prediction in classroom learning: A review of educational data mining studies," *Educ. Inf. Technol.*, vol. 26, no. 1, pp. 205–240, 2021.

[5] P. Using and M. Learning, "education sciences A Systematic Literature Review of Student ' Performance Prediction Using Machine Learning Techniques," 2021.

[6] A. Al-Gahmi, "On the Portability and Robustness of Early Student Performance Predictions," 2024.

[7] I. Taufik, S. F. Fabrianne, A. Triayudi, I. D. Sholihati, C. Gupta, and A. Saha, "Student Performance Prediction Model based on Supervised Machine Learning Algorithms Student Performance Prediction Model based on Supervised Machine Learning Algorithms," doi: 10.1088/1757-899X/928/3/032019.

[8] N. H. Haron, R. Mahmood, N. M. Amin, A. Ahmad, and S. R. Jantan, "An Artificial Intelligence Approach to Monitor and Predict Student Academic Performance," *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 44, no. 1, pp. 105–119, 2025.

[9] A. B. S. Salamh and H. I. Akyüz, "A Novel Feature Extraction Descriptor for Face Recognition," *Eng. Technol. Appl. Sci. Res.*, vol. 12, no. 1, pp. 8033–8038, Feb. 2022, doi: 10.48084/etasr.4624.

[10] A. B. S. Salamh and H. Akyüz, "New Deep Learning Model for Face Recognition and Registration in Distance Learning," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 12, pp. 29–41, Jun. 2022, doi: 10.3991/ijet.v17i12.30377.

[11] S. Nayani and S. R. P, "Combination of deep learning models for student's performance prediction with a development of entropy weighted rough set feature mining," *Cybern. Syst.*, vol. 56, no. 2, pp. 170–212, 2025.

[12] A. S. Hashim, W. A. Awadh, and A. K. Hamoud, "Student performance prediction model based on supervised

machine learning algorithms," in *IOP conference series: materials science and engineering*, 2020, vol. 928, no. 3, p. 32019.
[13] A. Ahmed *et al.*, "Students' performance prediction employing Decision Tree," *CTU J. Innov. Sustain. Dev.*, vol. 16, no. Special issue: ISDS, pp. 42–51, 2024.
[14] A. M. Shitaya, M. E. S. Wahed, A. Ismail, M. Y. Shams, and A. A. Salama, "Predicting student behavior using a neutrosophic deep learning model," *Neutrosophic Sets Syst.*, vol. 76, pp. 288–310, 2025.
[15] A. Acharya and D. Sinha, "Early prediction of students performance using machine learning techniques," *Int. J. Comput. Appl.*, vol. 107, no. 1, pp. 37–43, 2014.
[16] B. Cheng, Y. Liu, and Y. Jia, "Evaluation of students' performance during the academic period using the XG-Boost Classifier-Enhanced AEO hybrid model," *Expert Syst. Appl.*, vol. 238, p. 122136, 2024.
[17] U. Zia, "Predict Student Performance [Data set]," 2024, doi: Kaggle. https://doi.org/10.34740/KAGGLE/DSV/10303802.
[18] A. Acharya, "Early Prediction of Students Performance using Machine Learning Techniques," vol. 107, no. 1, pp. 37–43, 2014.
[19] M. Silva, R. Rupasingha, and B. Kumara, "Identifying complex causal patterns in students' performance using machine learning," *Technol. Pedagog. Educ.*, vol. 33, no. 1, pp. 103–119, 2024.
[20] N. A. Kuadey, C. Ankora, F. Tahiru, L. Bensah, C. C. M. Agbesi, and S. O. Bolatimi, "Using machine learning algorithms to examine the impact of technostress creators on student learning burnout and perceived academic performance," *Int. J. Inf. Technol.*, vol. 16, no. 4, pp. 2467–2482, 2024.
[21] D. Castillo-Merino and E. Serradell-López, "An analysis of the determinants of students' performance in e-learning," *Comput. Human Behav.*, vol. 30, pp. 476–484, 2014.
[22] D. Khairy, N. Alharbi, M. A. Amasha, M. F. Areed, S. Alkhalaf, and R. A. Abougalala, "Prediction of student exam performance using data mining classification algorithms," *Educ. Inf. Technol.*, pp. 1–25, 2024.
[23] A. Info, "Customer Churn Prediction Using Ordinary Artificial Neural Network and Convolutional Neural Network Algorithms : A Comparative Performance Assessment," vol. 36, no. 2, pp. 720–733, 2023, doi: 10.35378/gujs.992738.