INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

A Graph-Based Neural Network Approach to Student Performance Prediction using Mobile Usage Data

R Ruth Belina[1]*, T Lucia Agnes Beena[2]

1 Research Scholar, Department of Computer Science, Holy Cross College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli - 620002, Tamil Nadu, India
2 Research Supervisor, Department of Computer Science, Holy Cross College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli - 620002, Tamil Nadu, India

OPEN ACCESS

Received: 03/06/2025

Accepted: 14/07/2025

Published: 20/08/2025

Citation: Belina RR, Beena TLA (2025) A Graph-Based Neural Network Approach to Student Performance Prediction using Mobile Usage Data. Indian Journal of Science and Technology 18(31): 2555-2568. https://doi.org/10.17485/IJST/v18i31.1034

* **Corresponding author**.

ruthbelinajr@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To obtain enhanced accuracy and reduced specificity in predicting student academic performance based on mobile phone usage. **Methods:** This paper presents a novel technique called Canberra Normalized Elastic Net Regression-based Stochastic Gradient Multilayer Perceptron Neural Network (CNER-SGMPNN). A student records are collected from the real time dataset. After that, preprocessing is carried out by using Canberra Normalization (CN) technique to remove duplicate data. Followed by, feature selection is performed using Elastic Net Regression (ER) to choose most relevant attributes. Finally, Hamann Indexive SGMPNN framework is employed for student academic performance prediction by measuring correlations between mobile phone usage and academic results. **Findings:** The results of the analysis demonstrate that the CNER-SGMPNN technique improves accuracy and sensitivity by 8%, while achieving reductions of 39% in specificity, 47% in RMSE, and 48% in MAPE when compared to the existing methods. **Novelty:** Multilayer Perceptron Neural Network is employed to predict students' performance with several layers. Hamann similarity index, maxout activation function, and stochastic gradient function are applied to this Neural Network. The Hamann similarity index enables in CNER-SGMPNN technique to meaningfully process binary-encoded behavioral data (like mobile usage data) and students' performance. This enhances the model's ability to learn significant relationships between student mobile usage behavior and academic performance. This helps to accurately predict the low academic and high academic performance students. The maxout activation function in Neural Network accurately predicts a low academic performance student by choosing the maximum value of Hamann similarity index outputs. The stochastic gradient function is employed for updating the hyperparameter (i.e. weight) of the Multilayer Perceptron Neural Network to minimize the error in the low academic performance prediction.

Keywords: Academic performance prediction; Attribute selection; Data normalization; Mobile phone usage; Multilayer perceptron neural network

1 Introduction

Predicting the students' academic performance is very much relevant in judging the quality of education and assessing the well-being of the students. The intensive use of smartphones and constant screen distraction are the more important obstacles against researching habits, short-term attention spans, and a defeat in the overall learning experience. Several solutions are suggested to overcome the enhanced accurate prediction. Proposed APP-DGNN[1] utilizes structural information extracted from interaction activities. The method was not effective in improving prediction accuracy. To decrease the error rate and simplify predictions, the GA-based decision tree classifier was presented in[2]. The approach did not effectively handle a large volume of data. To deal with these limitations, the technique proposed here makes use of an MLP-NN effectively providing accurate predictions of academic performance. The technique enhances predictive ability by utilizing data from student statistics and mobile usage, making the process more efficient and scalable.

Different research focus on the issue of academic performance prediction regarding smartphone use. The effect of smartphone usage on academic performance was evaluated[3], and outcomes were enhanced. However, strong predictive models were not employed. Formulated a network analysis model was presented in[4] to check the performance of student. However, the model did not evaluate the impact of mobile phone usage. To account for smartphone usage, machine learning algorithms were presented[5]. However, error minimization was not efficiently achieved by the machine learning approach. Mobile phone usage was investigated[6] to predict academic performance. However, comparative study among educational institutions in different areas was not conducted to observe the relationship between mobile phone addiction and academic performance. Structural modeling was designed[7] in predicting academic performance regarding smartphone addiction, but there was no application of ML methods to enhance the performance prediction. Proposed SAS-SV model was designed[8] for performance prediction using Pearson correlation. But the model did not deliver the best performance. Regression techniques were used[9] to analyze mobile phone usage and behavioral changes. However, it did not provide accurate assessment of the impact of mobile phone usage on predicting academic performance. Similarly, academic performance was estimated[10] using the Pearson correlation coefficient. However, the method did not achieve the accuracy needed for dependable performance prediction. A moderated mediation model[11] explained the relationship between mobile phone addiction and the performance of adolescents. However, model was unable to demonstrate the effect of depression on academic performance. To enhance performance for students, the connection between smartphone addiction and data forecasting was analyzed[12]. Similarly, SEM was applied[13] to examine mobile phone addiction but only applied for Chinese medical students. To analyze current academic data, the relationships of the data examined[14] but it was not easy to handle the large data samples. Machine learning algorithms[15] focused on performance prediction in high school students with large datasets. However, there was a lack of effective prediction techniques to enhance the accuracy. The study focused on smartphone use and academic performance in[16] for education. However, the correlation analyses were not employed. Statistical analysis was carried out in[17] for handling missing entries. However, the approach was not successful in strengthening accuracy for analyzing students' academic performance. A deep learning (DL) based model named Students Academic Performance Prediction Network (SAPPNet) was discussed in[18] with maximum accuracy. However, it did not reflect real-time conditions accurately. The systematic Literature Review (SLR) approach was investigated in[19] for Smartphone usage habits. However, it did not perform the correlation analysis between the mobile phone usage data and academic performance. ML was employed in[20] to minimize the dimensionality of data. However, the method did not effectively improve the accuracy. Table 1 shows the Summary of Literature Review.

1.1 Research Gap

Student academic prediction performance is an important concern in online learning analytics. Academic performance prediction is a proficient method to handle learning and evaluate the learning process. Conventional machine learning and deep learning-based models are made to perform preprocessing and feature selection. This method was unable to remove noisy data and select significant features. Several studies were developed for academic performance prediction. However, predicting academic performance is a demanding task that requires capturing the complex relationships in online learning activities and student attributes with higher prediction accuracy and sensitivity. Specificity was not considered in several studies. To overcome the gap, the CNER-SGMPNN method is introduced for academic performance prediction from mobile phone usage. The developed method reduces specificity while improving prediction accuracy and sensitivity through preprocessing, feature selection, and prediction techniques.

1.2 Major contribution of the paper

The key contributions of the CNER-SGMPNN method are,

Table 1. Summary of Literature Review

Reference no and Methods	Topic/Focus	Key findings	Merits	Demerits
APP-DGNN [1]	Academic performance was a vital issue in online learning analytics	Structural information derived from interaction activities	Increased accuracy	Higher specificity
GA-based decision tree classifier [2]	Focuses on students' grades and marks prediction	Predict students' academic performance	Higher accuracy	Specificity was not reduced
SAS-SV [8]	Estimate the relationship between smartphone addiction, sleep quality, and academic performance	Forecast academic performance	Specificity was lower	The optimal strategy was not employed
Moderated mediation model [11]	Detecting mobile phone addiction and academic performance of adolescents	Standard deviations were estimated	Predicting poor school performance was higher	Mobile phone addiction and related depression in adolescents were not considered
Data gathering [16]	Smartphone Use and Academic Performance	Finding a negative association	Specificity was minimized	Correlation analyses were not utilized
Statistical analysis [17]	Discover student data prediction performance	Predict study hours and attendance outcome	Enhance academic performance	Accuracy was not enhanced
SAPPNet [18]	Predict student performance, determine smartphone use and academic performance	Forecast final semester grade performance	Improve accuracy	Failed to consider real-time situation
SLR [19]	Determine smartphone use and academic performance	Capture students' experiences with smartphone use	Specificity was reduced	Failed to consider correlation analyses
ML [20]	Predicting student academic performance	Determine important attribute(s) in the student's data	Data dimensionality was minimized	Accuracy was reduced
Proposed CNER-SGMPNN technique	Predicting student academic performance based on mobile phone usage	Accurate predictions of student academic performance with minimal specificity	Higher accuracy, sensitivity and lower specificity	Investigate complex relationships between different social and environmental factors

- **Improved Accuracy and Sensitivity:** A Hamann Indexive SGMPNN is introduced to predict student academic performance. The model measures the similarity between mobile phone usage attributes and other relevant factors, resulting in improved prediction accuracy. The suggested technique includes Canberra-based information normalization for preprocessing and Elastic Net Regression for attribute selection which efficiently improves the accuracy.
- **Reduction of Specificity RMSE and MAPE:** Deep learning framework utilizes the Stochastic Gradient method, effectively lowering specificity, RMSE and MAPE and enhancing prediction reliability.

2 Methodology

The CNER-SGMPNN technique presented in this research aims to achieve enhanced accuracy in predicting student academic performance based on mobile phone usage. The flow of the performance prediction process is shown in Figure 1, which illustrates the proposed CNER-SGMPNN framework for accurate academic performance prediction.

The real-time student dataset, denoted as 'D', comprises student records '$d_1, d_2, d_3 \ldots\ldots d_m$' and associated attributes '$A_j = A_1, A_2, A_3, \ldots\ldots A_n$'. The preprocessing phase utilizes Canberra distance-based data normalization described in [21] which calculates relationships between data points and removes duplicates. Subsequently, feature extraction using Elastic Net Regression-based Attribute Selection is performed in [22] enabling the identification of significant attributes that contribute to prediction. Finally, the Hamann Indexive SGMPNN is applied to identifying the low academic performance students. Finally,

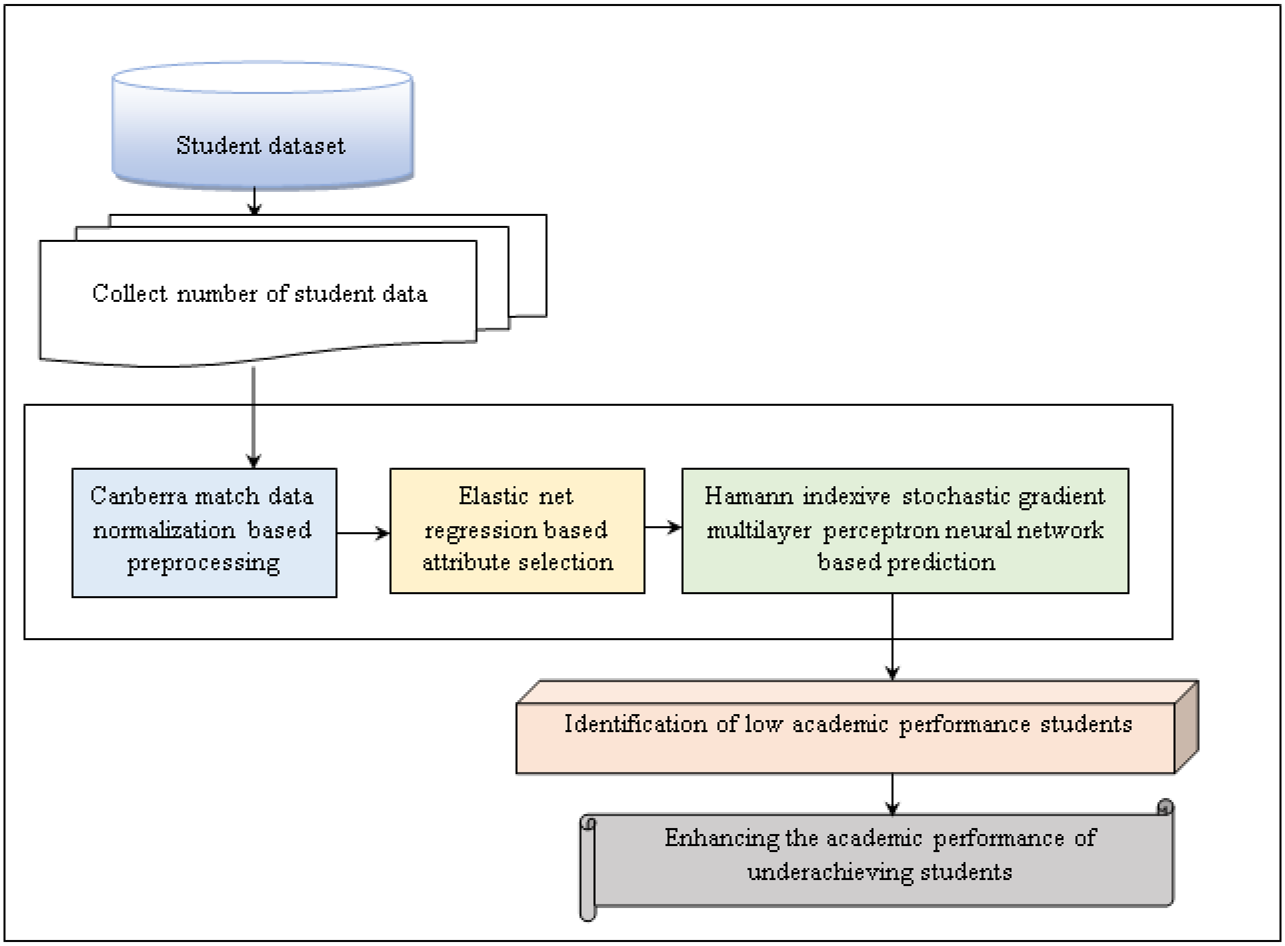

Fig 1. Flow diagram for the CNER-SGMPNN approach

the analysis is conducted to improve the low academic performance of students into higher achievement levels. This method improves prediction accuracy, providing solid insights into student academic performance.

2.1 Hamann Indexive Stochastic Gradient Multilayer Perceptron Neural Network based prediction

The Hamann Index-driven SGMPNN is used to predict students with low academic success based on their mobile phone utilization. This neural network is a fully connected artificial feed-forward deep learning model composed of multiple layers, including an input layer, an output layer, and several hidden layers. These layers work collaboratively to process the input data and generate accurate predictions.

The construction of a multilayer perceptron neural network comprised of input layer, hidden layers, and output layer. The architecture includes training samples $\{d_i, Y_i\}$ where d_i denotes training student data (i.e. mobile usage attributes) $\{d_1, d_2, d_3 \ldots\ldots d_m\}$ and 'Y'indicates a desired prediction output. Each layer consists of artificial neurons or perceptron or nodes that transfer input from one layer to another. Therefore, the activity of neurons is expressed as,

$$Z = \sum_{i=1}^{m} [d_i * \delta_t] + B_{Ik} \tag{1}$$

Where, neuronal activity in the input processing layer 'Z', 'd_i' denotes student data,'δ_t' denotes a weight between input and hidden layer, bias function $'B'$, which saved the value as $'1'$. The input has been transmitted to hidden layers, whereby attribute evaluation is conducted using the Hamann similarity index. The Hamann index is a statistical method to measure the similarity among mobile phone usage attributes and student academic attributes.

Let us consider the number of student data $d_i = \{d_1, d_2, d_3 \ldots \ldots d_m\}$ i.e. scores are derived from attributes. The Hamann index measures the similarity between the data with binary representations. First, input mobile phone usage attributes data and student academic attributes data are converted into the binary representation. Before applying the Hamann index, first score values are assigned to each numeric and categorical attribute. For example, Mobile phone usage time in a day attribute have four ranges and each assigned an integer score from 1 to 4. Score 1 is assigned to 0–2 hours of mobile phone usage per day. Score 2 is assigned to 2–4 hours of usage. Score 3 is assigned to 4–6 hours of usage. Finally, Score 4 is assigned to more than 6 hours of usage. These score values are represented as integers. Then applying a repeated division by 2 methods to converts a decimal number into binary numeral system, where each digit is either 0 or 1. In this method, select a number and divide it by 2. Then record the remainder (either 0 or 1) and update the number to the quotient. The final binary representation is obtained by reading the remainders from bottom to top. Based on this process, the binary representation of the input data is obtained using **eq (2)**.

$$d_i = [b_1, b_2 \ldots b_n] \tag{2}$$

Where, d_i denotes mobile usage data, $[b_1, b_2 \ldots b_n]$ denotes a binary representation of student data. Similarly, binary representations of student academic performance data 'd_t' are obtained using **eq (3)**.

$$d_t = [\, b_1{}', b_2{}' \ldots b_n{}'] \tag{3}$$

Then the Hamann index similarity is measured between the data in **eq (4)**.

$$HS = \frac{m - 2 * |d_i \Delta d_t|}{m} \tag{4}$$

Where, HS represents Hamann Similarity evaluates the relationship among two binary data.d_i denotes training student data (i.e. mobile usage attributes), testing data(i.e. student academic performance attributes) d_t, m indicates the total length of binary data, $|d_i \Delta d_t|$ denotes a difference between the position of bits in two binary data. The similarity value gives an output range of $0\ to\ 1$. Data are sorted into several categories based on their similarity value classes namely Marginal, Below Average, Average, Above Average, Good, and Excellent. Hence proposed deep learning method uses the max-out activation function for accurate prediction. The max out activation function is a piecewise linear function that computes the maximum value among a set of inputs as given in **eq (5).**

$$Z = \begin{cases} 1; & \arg\max HS \\ 0; & otherwise \end{cases} \tag{5}$$

Where Z denotes a predicted result, the max out activation function returns '1' when the similarity output is maximum (i.e. 1). Otherwise, the activation function is minimal, and it incorrectly predicts performance. The highest activation coefficient implies that student academic achievement may be precisely anticipated based on data similarities. Throughout the learning phase, the incorrect rate of data prediction is calculated using **eq (6)**.

$$E = \frac{1}{2}(Z_a - Z)^2 \tag{6}$$

Where error rate 'E' is measured as a squared difference between actual results 'Z_a' and predicted result'Z'. To minimize the error, the weight between the layers is updated using stochastic gradient functioning **eq (7)**.

$$\delta_{new} = \delta_t * \eta \left(\frac{\partial E}{\partial \delta_t}\right) \tag{7}$$

Where, δ_{new} indicates a modified weight, δ_t indicates the present weight, η indicates a learning speed, '$\frac{\partial E}{\partial \delta_t}$' signifies an incomplete derivative of the errors. 'E' with regard to the present weight 'δ_t'. This method is repeated until the algorithm has low error. If the minimum error is reached, the results are presented in an output layer utilizing **eq 8**.

$$Y = \sum_{i=1}^{n} H(t) * \delta_{ho} \tag{8}$$

From **eq (8)**, Y' represents the final output of deep learning results, $H(t)$ denotes a hidden layer output, 'δ_{ho}' represents the weight between the hidden and output layers. Prediction results with minimal error are obtained at the output layer resulting in increased accuracy.

Algorithm 1: Hamann Indexive Stochastic Gradient Multilayer Perceptron Neural Network based prediction
Input: Selected relevant student mobile usage attributes $d_i = d_1, d_2, d_3 \ldots..d_m$
Output: Accuracy

Step 1: Number of selected student mobile usage attributes $d_1, d_2, d_3 \ldots..d_m$ **at the input layer**
Step 2: For every student, datad_i
Step 3: Assign weight 'δ_t' **and** add bias 'B'
Step 4: End for
Step 5: For each mobile usage attributes'd_i'–**[hidden layer]**
Step 6: **For** each student academicperformanceattribute'd_t'
Step 7: Measure the HamannSimilarity index
Step 8: **End for**
Step 9: End for
Step 10: Apply maxout activation function '$A_{\max}$'
Step 11: **If**$(A_{\max} = 1)$**then**
Step 12: Low academic performance of students is correctly predicted
Step 13: **else**
Step 14: Low academic performance of students is incorrectly predicted
Step 15: **End if**
Step 16: **For every categorization outcome**
Step 17: Assess the mistake rate. 'E'
Step 18: Increase the weight 'δ_{new}'
Step 19: Identify the most appropriate weight to minimize errors
Step 20: **End for**
Step 21: Obtain the final results with minimum error **at the output layer**
End

3 Results and Discussion

Experimental assessments of the proposed CNER-SGMPNN technique were implemented using the Python programming language. The experimental evaluation utilized a real-time dataset collected from college students.

3.1 Performance analysis of low academic performance prediction versus mobile phone usage attributes

The low academic performance student is identified using the CNER-SGMPNN technique, which utilizes ten mobile phone usage attributes from the dataset. The survey, conducted to examine the relationship between mobile usage attributes and academic performance, involved a total of 1,115 students. Among them, 915 students were predicted as having better academic performance, while 200 students were identified as having low academic performance. The academic performance of the 915 high-achieving students is further identified into four groups. Each group represents the percentage contribution of individual attributes to the student's performance. The improved academic performance of these students is determined by higher and positive attribute scores. By utilizing strong attribute score values, the evaluation results for the academic performance of the 915 students are showcased, illustrating the prediction accuracy of the proposed model.

Figure 2 provides a graphical representation of the academic performance predictions for 915 students, based on their scores and corresponding academic performance percentages. The predictions are divided into performance ranges of 70–80%, 80–90%, 90–95%, and above 95%, with higher performance levels correlating with improvements in selected attributes. The results show that students who primarily use **mobile phones for academic purposes, limit their daily usage, and actively participate in class, library visits, sports, and other productive activities** exhibit significantly better academic outcomes. Furthermore, fostering effective communication between students and teachers promotes better learning, while spending quality time with family members encourages meaningful interactions, reducing excessive mobile phone usage.

Among 1115 students, 200 students were predicted as having low academic performance i.e. below average which is scoring less than 60%, based on ten mobile phone usage attributes. These attributes include: The purpose of mobile usage, Mobile

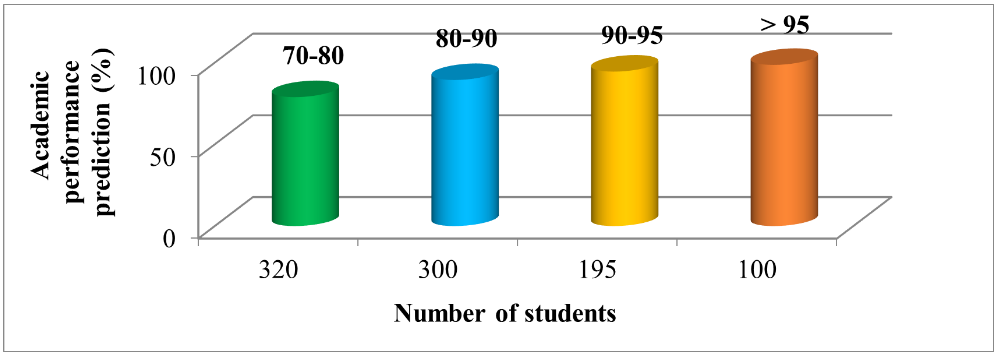

Fig 2. Analysis of academic performance prediction for 915 students

phone usage time in a day, the Impact of mobile phones on academics, Lack of concentration in classes and other activities, Depression in life compared with others on social media, Mobile phone usage affects the sleep distress, Depression due to incomplete assignments, Distraction during study time due to Smartphone usage, Spending more money on online shopping and Smartphone as primary source of entertainment. A detailed summary of all ten attributes before intervention is provided in the table 2.

Table 2. Mobile usage attributes before intervention

S. No	Mobile usage attributes
1	Depression in life comparing with others on social media
2	Impact of mobile phone on academics
3	Lack of concentration in classes and other activities
4	Purpose of mobile usage
5	Mobile phone usage affects sleep distress
6	Depression due to incomplete assignments
7	Mobile phone usage duration in a day
8	Distraction during study time due to Smartphone usage
9	Spending more money on online shopping
10	Smartphone- main form of entertainment

To enhance the academic performance of these 200 students, additional data was collected after six months by encouraging students to actively participate in both indoor and outdoor sports events such as cricket, chess, throw the ball, and similar sports. Students were also encouraged to contribute during class activities by involving them in snowball discussions, Jigsaw, role play, brainstorming, and worksheet assessments. The introduction of these new methods in teaching and encouraging learning activities resulted in the improvement of overall academic performance. From Table 1, the phrase 'other activities' within the original attribute 'Lack of Concentration in Classes and Other Activities' refers to activities such as sports and class participation. Therefore, the same attribute is used for both before and after the intervention. The only change lies in the score value. Before the intervention, the score for 'Lack of Concentration' was 1, indicating poor concentration. After the intervention, the score increased to 4, indicating improved concentration. Therefore, the improvement is achieved in overall academic performance after the intervention. The graphical results of before and after performance predictions are visually represented in Figure 3.

This indicates an inverse relationship between mobile usage attributes and academic performance predictions. The performance results of academic prediction before intervention range from 20-30%, 30-40%, 40-50%, and 50-60%. After intervention, improved performance results fall within the ranges of 50.2%–54.7%, 55.3%–59.8%, 60.2%–69.7%, and 70.3%–79.8%. Moreover, encouraging student participation in classroom activities and promoting better health through sports significantly contribute to improved academic performance levels.

The results of enhanced academic performance prediction with 200 students by selected attributes are determined using the Hamann Similarity (HS) index. This index measures the similarity between two variables, such as mobile usage attributes and academic performance, and indicates the extent to which changes in one variable relate to changes in the other. A higher similarity value signifies a strong positive correlation. For example, if changes in key attributes correspond to higher academic

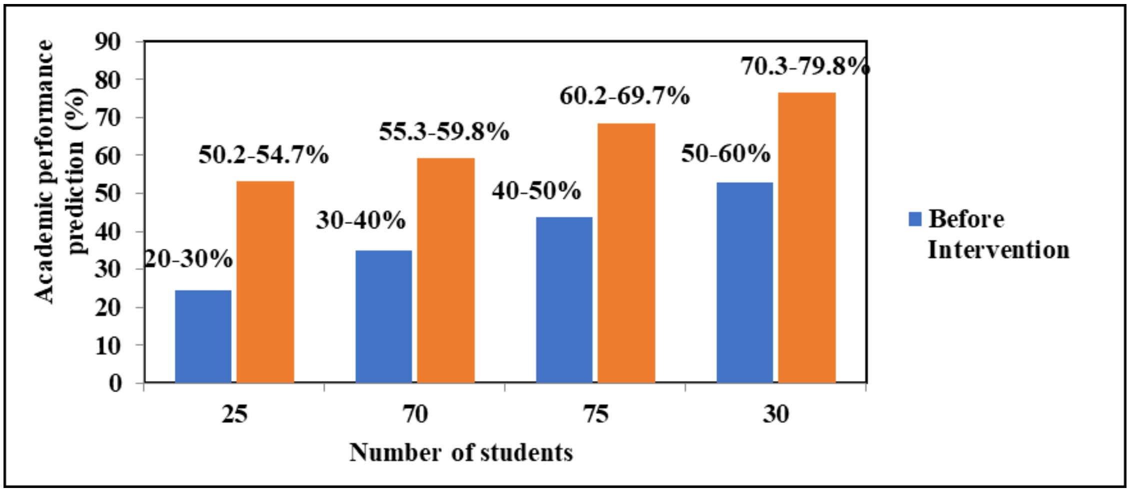

Fig 3. Before and after improving academic performance prediction for 200 students

performance and the HS index is high, it reflects a significant relationship. Conversely, lower similarity values indicate a weaker association. Academic performance predictions are compared for both pre and post-intervention scenarios using data collected over six months. The detailed results are presented in the table 3.

Table 3. HS between attributes and academic performance for category

Attributes	**Before Intervention**				**After Intervention**			
	HS for Category I	**HS for Category II**	**HS for Category III**	**HS for Category IV**	**HS for Category I**	**HS for Category II**	**HS for Category III**	**HS for Category IV**
Depression in life comparing with others on social media	0.856	0.756	0.823	0.808	0.622	0.605	0.589	0.612
Impact of mobile phone on academics	0.863	0.823	0.808	0.785	0.765	0.720	0.709	0.682
Lack of concentration in classes and other activities	***0.657***	***0.636***	***0.623***	***0.566***	***0.975***	***0.980***	***0.988***	***0.994***
Purpose of mobile usage	0.720	0.633	0.745	0.762	0.821	***0.968***	***0.981***	0.852
Mobile phone usage affects sleep distress	0.789	0.763	0.689	0.621	0.632	0.715	0.642	0.586
Depression due to incomplete assignments	0.726	0.712	0.696	0.635	0.658	0.685	0.653	0.615
Mobile phone usage duration in a day	0.763	0.723	0.702	0.696	0.656	0.639	0.615	0.589
Distraction during study time due to Smartphone	0.725	0.678	0.712	0.675	0.701	0.625	0.684	0.611
Spending more money on online shopping	0.638	0.620	0.615	0.608	0.578	0.565	0.541	0.523
Smartphone- main form of entertainment	0.746	0.736	0.722	0.714	0.686	0.671	0.637	0.624

Table 3 shows the Hamann Similarity (HS) index results for ten attributes in relation to academic performance prediction across four distinct categories, both before and after the intervention. Each attribute's HS value, measured on a scale from 0 to 1, represents its degree of similarity to academic performance. HS values above 0.9 imply a strong correlation, identifying these attributes as highly influential in predicting high academic outcomes.

Before the intervention, ten attributes were considered to compute HS. Among ten attributes, (HS) values of eight attributes such as Depression in life compared with others on social media, the impact of mobile phone usage on academics, sleep disturbances due to mobile phone use, depression caused by incomplete assignments, duration of mobile phone usage per day, Distraction during study time, excessive spending on online shopping, and the use of smartphones as the main form of entertainment were high for all categories. However, the HS values for two attributes such as lack of concentration in classes and other activities, and the purpose of mobile usage were low. However, HS values do not exceed above 0.9.

To address this issue, data were collected again after six months. In the after-intervention phase, participation in sports events and concentration in class activities replaced the previous negative attribute of lack of concentration in class and other activities. Additionally, there was an increase in the HS values for the attribute related to using mobile phones for study purposes.

The changes in these attributes significantly predict improved academic performance, leading to HS values exceeding 0.9 after the intervention, compared to the values recorded before the intervention.

3.2 Performance Analysis of CNER-SGMPNN Versus Student Academic DataTop of Form

The performance of the proposed CNER-SGMPNN technique is evaluated with CMN-GCCDSC [21], CNERBC [22], and existing APP-DGNN [1] based on the different parameters.

3.2.1. Accuracy

Accuracy is defined as the percentage of correctly predicted information about students to the overall amount of student records. It's given as a fraction and calculated using **eq (9)**.

$$Accuracy = \frac{Number\ of\ student\ data\ more\ accurately\ predicted}{Number\ of\ student\ data} * 100 \tag{9}$$

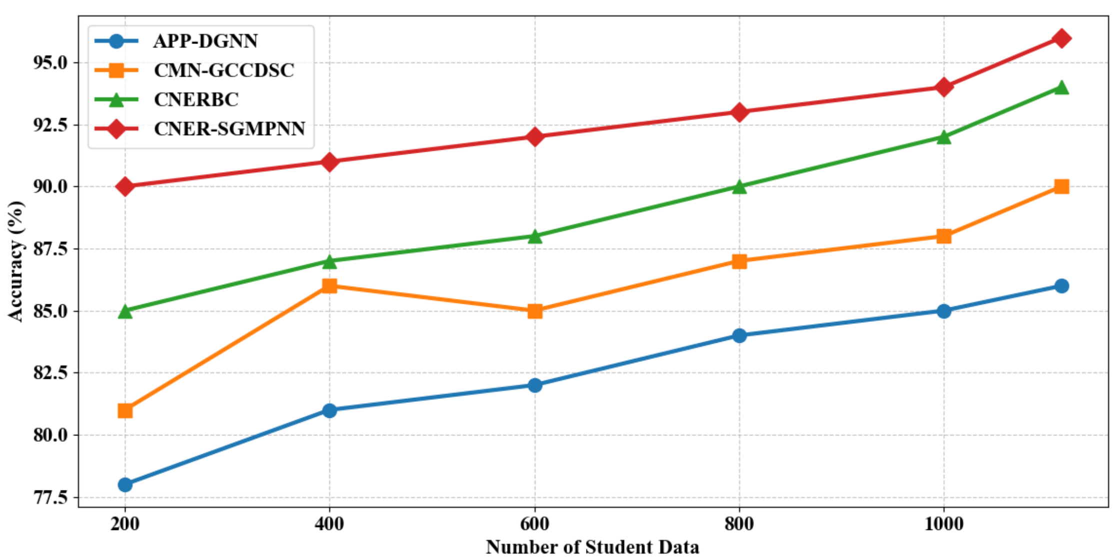

Fig 4. A comparative examination of accuracy

Figure 4 depicts the outcome of the analysis of accuracy with the quantity of student information. The results of accuracy were higher using CNER-SGMPNN than the other methods. When considering 200 student data in result analysis, the accuracy using CNER-SGMPNN was found to be 90%. Subsequently, 78%, 81%, and 85% of accuracy were observed by applying methods namely APP-DGNN [1], CMN-GCCDSC [21], and CNERBC [22] respectively. The percentage improvement is calculated as the ratio of the absolute difference between the accuracy of proposed CNER-SGMPNN model and that of existing methods, relative to the accuracy of existing methods. From the analysis, CNER-SGMPNN model increases the performance of accuracy by

15%, 11% and 6% compared to APP-DGNN [1], CMN-GCCDSC [21], and CNERBC [22] respectively. Similarly, performance improvements of the CNER-SGMPNN model were computed across student data samples ranging from 400 to 1115. The average of six experimental results was then used to evaluate the overall performance improvement of CNER-SGMPNN. Based on the accuracy calculation, the average of these six results shows that CNER-SGMPNN improved overall accuracy by 12% compared to APP-DGNN, 8% compared to CMN-GCCDSC, and 4% compared to CNERBC. The reason for higher accuracy is to apply the Hamann indexive stochastic gradient multilayer perceptron neural network. The Hamann index is used to measure the similarity among mobile phone usage attributes and student academic attributes. Based on similarity values, the student's academic performance level is correctly predicted with higher accuracy.

3.2.2. Sensitivity

Sensitivity measures the proportion of correctly predicted student data out of a total number of data. It is expressed as a percentage (%) and formulated using **eq (10)**.

$$Sensitivity = \frac{True\ positive}{True\ positive + False\ Nagative} * 100 \tag{10}$$

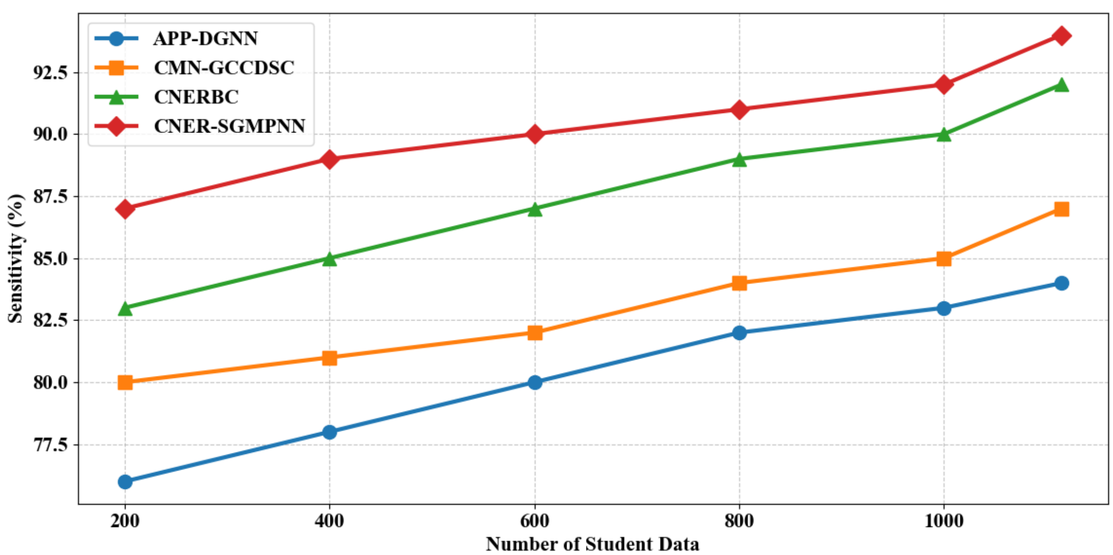

Fig 5. A comparative examination of sensitivity

Figure 5 shows a sensitivity investigation regarding student academic achievement forecasting. When considering 200 student data in result analysis, the sensitivity using CNER-SGMPNN was found to be 87%. Subsequently, 76%, 83%, and 80% of sensitivity were observed by applying methods namely APP-DGNN [1], CMN-GCCDSC [21] and CNERBC [22] respectively. The percentage improvement of is calculated as the ratio of the absolute difference between the sensitivity of the proposed CNER-SGMPNN model and that of existing methods, relative to the sensitivity of the existing methods. From the comparison, CNER-SGMPNN increases the percentage of sensitivity by 12%, 9% and 3%. Among the four methods, CNER-SGMPNN demonstrates superior sensitivity performance. Similarly, performance improvements of sensitivity using CNER-SGMPNN model is computed across student data samples ranging from 400 to 1115. The average of six experimental results is then used to evaluate the overall performance improvement of sensitivity using CNER-SGMPNN. Based on the sensitivity calculation, the average of these six results shows that CNER-SGMPNN improved overall sensitivity by 12% compared to APP-DGNN [1], 9% compared to CMN-GCCDSC [21], and 3% compared to CNERBC [22] respectively. This enhancement is achieved by applying the Hamann index to the Multilayer Perceptron Neural Network for analyzing the mobile phone usage attributes and student academic attributes. This helps to effectively identify academic performance levels such as marginal, below average, average, above average, good, and excellent. For each method, different results were observed with respect to number of student data. The overall results show that CNER-SGMPNN improves sensitivity performance by 12% compared to APP-DGNN[1], 9% compared to CMN-GCCDSC[21] and 3% compared to CNERBC[22] respectively.

3.2.3. Specificity

Specificity can be described as the amount of improperly predicted student information divided by the total quantity of student information. It is expressed as a fraction and calculated using **eq (11)**.

$$Specificity = \frac{Number\ of\ student\ data\ incorrectly\ predicted}{Number\ of\ student\ data} * 100 \tag{11}$$

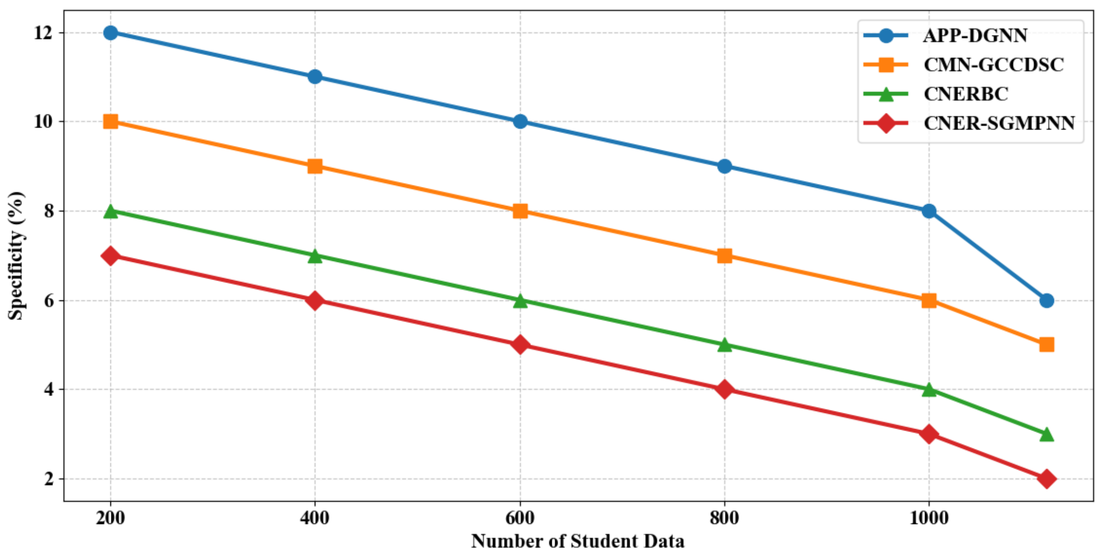

Fig 6. A comparative examination of specificity

Figure 6 depicts the outcome of the evaluation for specificity. For results analysis, 200 student data were considered in first run to compute the specificity. The performance of specificity was found to be 7%, 12%, 10%, and 8%using proposed CNER-SGMPNN, APP-DGNN[1], CMN-GCCDSC[21] and CNERBC[22] respectively. The performance improvement of specificity using CNER-SGMPNN is computed based taking the absolute difference between the sensitivity of the proposed CNER-SGMPNN model and that of existing methods and dividing it by the sensitivity of the existing methods. From the calculation, the performance of specificity using CNER-SGMPNN s reduced by 42%, 30% and 13% when compared to APP-DGNN[1], CMN-GCCDSC[21] and CNERBC[22]. Similarly, performance improvements of specificity using CNER-SGMPNN model is computed across student data samples ranging from 400 to 1115. The average of six experimental results is used to evaluate the overall performance improvement of specificity using CNER-SGMPNN. Based on the specificity calculation, the average of these six results shows that CNER-SGMPNN reduces overall specificity by 54% compared to APP-DGNN, 42% compared to CMN-GCCDSC, and 20% compared to CNERBC. The reason for the lesser specificity is to employ the Stochastic Gradient method in the deep learning technique. Weights are updated between the layers to minimize the prediction error.

3.2.4. Root mean square error:

This metric measures the accuracy of a predictive model by calculating the square root of the average squared differences between the actual and predicted academic performance outcomes, based on the total number of student records. The corresponding mathematical formula is given below.

$$RMSE = \left[\sqrt{\frac{\left(Y_{act} - Y_{pre}\right)^2}{m}} \right] \tag{12}$$

Where, $RMSE$ indicates a root mean square error, $\mathbf{Y_{act}}$ denotes the actual results, $\mathbf{Y_{pre}}$ denotes a predicted result, m denotes a total number of student data.

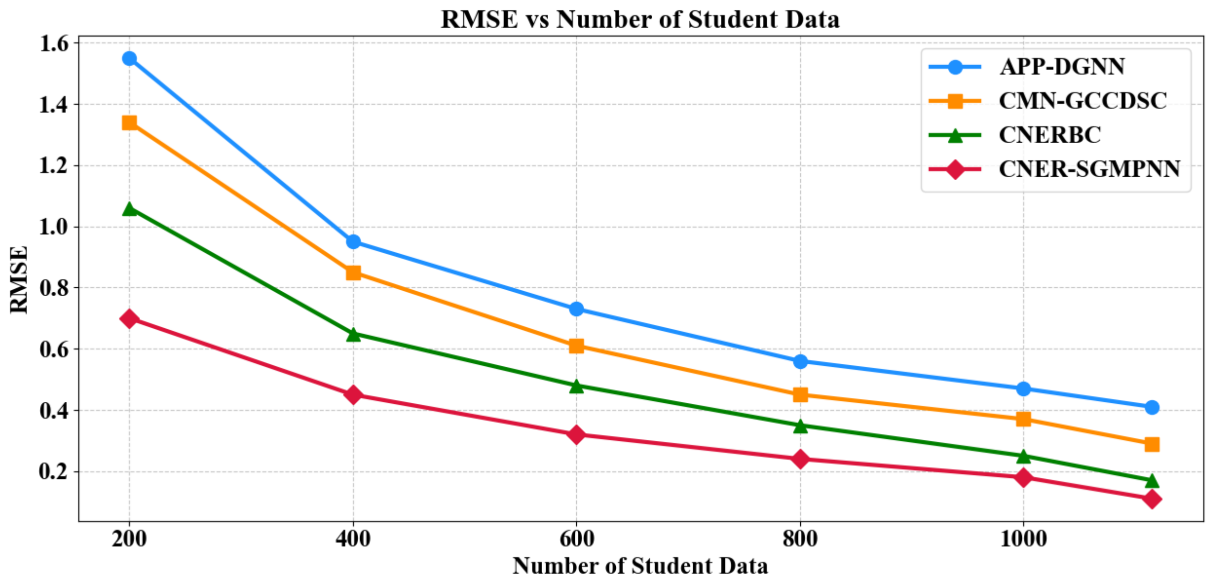

Fig 7. A comparative examination of RMSE

Figure 7 presents the evaluation results for RMSE using a dataset of 1115 student records. For each input data, various performance results of RMSE were observed for all four methods namely, APP-DGNN [1], CMN-GCCDSC [21], CNERBC [22] and proposed CNER-SGMPNN. To evaluate overall improvement, observed values were compared to the existing methods. The comparison indicates that the RMSE of the CNER-SGMPNN model is reduced by 59% compared to APP-DGNN [1], by 50% compared to CMN-GCCDSC [21], and by 32% compared to CNERBC [22]. The reduction in RMSE is the result of applying the Stochastic Gradient approach, which optimizes the model by minimizing prediction errors during training.

3.2.5. Mean absolute Percentage Error:

Mean Absolute Percentage Error (MAPE) is a metric that measures the accuracy of a model in the student academic performance prediction. It calculates the average of the absolute percentage differences between the actual and predicted values.

$$MAPE = \frac{1}{m} \left| \frac{Y_{act} - Y_{pre}}{Y_{act}} \right| * 100 \tag{13}$$

Where, $MAPE$ indicates Mean Absolute Percentage Error, $\mathbf{Y_{act}}$ denotes the actual results, $\mathbf{Y_{pre}}$ denotes a predicted result, m denotes a total number of student data. Lesser percentage of MAPE values indicates more accurate predictions.

Figure 8 presents the MAPE analysis versus student data. The CNER-SGMPNN technique demonstrated lesser MAPE compared to other methods. When analyzing results based on 200 student records, CNER-SGMPNN achieved an MAPE of 0.05%. In contrast, APP-DGNN [1], CMN-GCCDSC [21], and CNERBC [22] achieved 0.11%, 0.095%, and 0.075% MAPE, respectively. Each method was tested through multiple iterations using different input data volumes. The average of six results indicates that the CNER-SGMPNN outperformed other approaches by reducing the MAPE by 60% compared to APP-DGNN [1], 51% compared to GCCDSC [21], and 33% compared to CNERBC [22].

4 Conclusion

Student academic performance serves as a crucial indicator of both individual success and holistic educational impacts. This research work introduced the **CNER-SGMPNN technique** to accurately predict students' academic performance by solving the major issues of data processing, attribute relevance, and accuracy of the performance prediction. The student data is

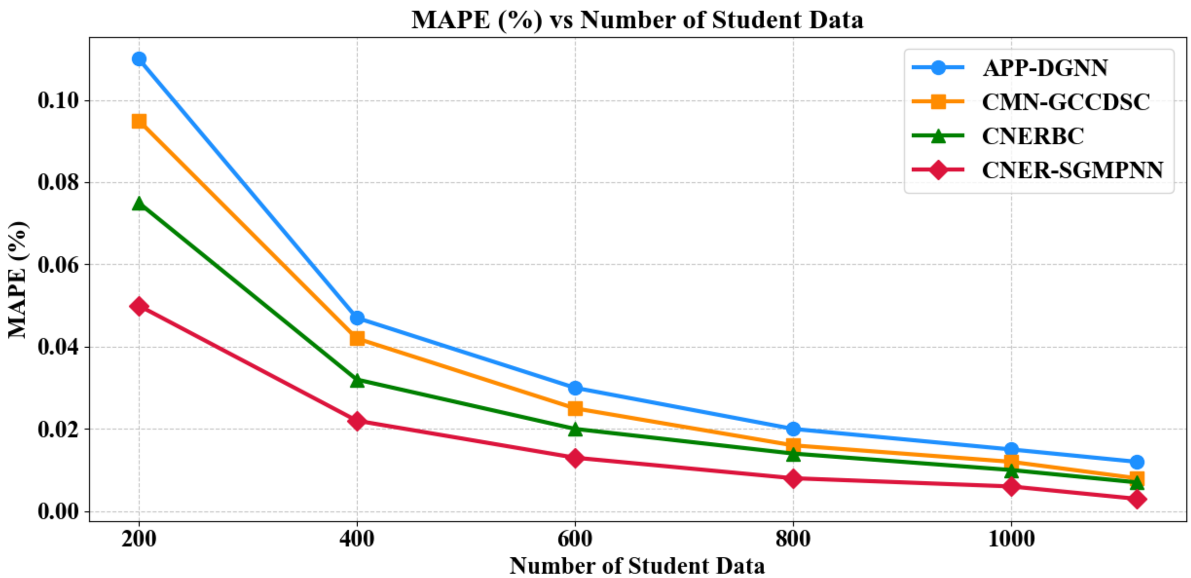

Fig 8. A comparative examination of MAPE

preprocessed by using the CN technique to remove outlier data. Subsequently, the attribute selection process is performed with the aid of ER by identifying the most relevant attributes. Finally, Hamann indexing SGMPNN is used to improve the accuracy of low academic performance predictions. The similarity among smartphone usage attributes and student academic attributes is determined by the Hamann index. Based on this analysis, precise student academic performance prediction is performed. The stochastic Gradient method is applied in deep neural networks to minimize specificity. A comprehensive performance analysis that took into account such factors as sensitivity, specificity, and accuracy made it evident that the proposed technique was useful. In comparison to existing techniques, the CNER-SGMPNN technique recorded an 8% increase in both accuracy and sensitivity and reduced specificity by 39%. In addition, the RMSE and MAPE performance of CNER-SGMPNN technique is reduced by 47% and 48%. The results show that students who are more active in physical activities and participate more in classroom activities have better concentration and retention of knowledge. The CNER-SGMPNN technique, therefore, emphasizes the need for a balanced approach to active participation in physical activities and focused academic engagement for improving student performance. These results are useful for educators and policymakers looking to improve academic outcomes through data-driven strategies. However, the other potential variables that influence academic performance were not focused. Also, several parameters such as F1 score, precison and recall not handled for preprocessing and feature selection. As machine learning is integrated into educational systems, ethical concerns regarding data privacy and security are recognized as increasingly important.

So, it is recommended for future research to further investigate other potential variables that influence academic performance including various datasets allows the system to understand complex relationships between student behavior, mobile fasting, academic engagement, learning environments, emotional, and time management skills. This would enable comprehensive research and provide further understanding into the complex relationship between different personality-related, social, and environmental factors affecting the educational performance of a student. Advanced AI techniques such as transformer-based models, and graph neural networks are explored to improve predictive accuracy, F1 score, precision and recall. Finally, in future research, ethical considerations, such as student data privacy, should be prioritized.

5 Acknowledgement

The authors would like to sincerely acknowledge the support of the Department of Science and Technology (DST), Government of India, for providing the FIST (Fund for Improvement of S&T Infrastructure) program funding. This assistance has significantly contributed to the research and development activities outlined in this work.

References

1) Huang Q, Zeng Y. Improving academic performance predictions with dual graph neural networks. *Complex & Intelligent Systems*. 2024;10:1–19. https://doi.org/10.1007/s40747-024-01344-z.
2) Hussain S, Khan MQ. Student-Performulator: Predicting Students' Academic Performance at Secondary and Intermediate Level Using Machine Learning. *Annals of Data Science*. 2023;10:637–655. https://doi.org/10.1007/s40745-021-00341-0.
3) Wang JC, Hsieh CY, Kung SH. The impact of smartphone use on learning effectiveness: A case research of primary school students. *Education and Information Technologies*. 2023;28:6287–6320. https://doi.org/10.1007/s10639-022-11430-9.
4) Liu C. The unique role of smartphone addiction and related factors among university students: a model based on cross-sectional and cross-lagged network analyses. *BMC Psychiatry*. 2023;23:1–12. https://doi.org/10.1186/s12888-023-05384-6.
5) Elhai JD, Yang H, Rozgonjuk D, Montag C. Using machine learning to model problematic smartphone use severity: The significant role of fear of missing out. *Addictive Behaviors*. 2020;103:1–7. https://doi.org/10.1016/j.addbeh.2019.106261.
6) Rashid J, Aziz AAA, Rahman AA, Saaid SA. The Influence of Mobile Phone Addiction on Academic Performance Among Teenagers. *Jurnal Komunikasi: Malaysian Journal of Communication*. 2020;36(3):408–424. https://doi.org/10.17576/JKMJC-2020-3603-25.
7) Abbasi GA, Jagaveeran M, Goh YN, Tariq B. The impact of type of content use on smartphone addiction and academic performance: Physical activity as moderator. *Technology in Society*. 2021;64:1–11. https://doi.org/10.1016/j.techsoc.2020.101521.
8) Rathakrishnan B, Singh SSB, Kamaluddin MR, Yahaya A, Nasir MAM, Ibrahim F, et al. Smartphone Addiction and Sleep Quality on Academic Performance of University Students: An Exploratory Research. *International Journal of Environmental Research and Public Health*. 2021;18(16):1–12. https://doi.org/10.3390/ijerph18168291.
9) Yadav MS, Kodi SM, Deol R. Impact of mobile phone dependence on behavior and academic performance of adolescents in selected schools of Uttarakhand, India. *Journal of Education and Health Promotion*. 2021;10:1–7. https://doi.org/10.4103/jehp.jehp_915_20.
10) Bravo-Sánchez A, Morán-García J, Abián P, Abián-Vicén J. Association of the Use of the Mobile Phone with Physical Fitness and Academic Performance: A Cross-Sectional Research. *International Journal of Environmental Research and Public Health*. 2021;18(13):1–11. https://doi.org/10.3390/ijerph18031042.
11) Bai C, Chen X, Han K. Mobile phone addiction and school performance among Chinese adolescents from low-income families: A moderated mediation model. *Children and Youth Services Review*. 2020;118:1–11. https://doi.org/10.1016/j.childyouth.2020.105406.
12) Alinejad V, Parizad N, Yarmohammadi M, Radfar M. Loneliness and academic performance mediates the relationship between fear of missing out and smartphone addiction among Iranian university students. *BMC Psychiatry*. 2022;22:1–13. https://doi.org/10.1186/s12888-022-04186-6.
13) Zhang CH, Li G, Fan ZY, Tang XJ, Zhang F. Mobile Phone Addiction Mediates the Relationship Between Alexithymia and Learning Burnout in Chinese Medical Students: A Structural Equation Model Analysis. *Psychology Research and Behavior Management*. 2021;14:455–465. https://doi.org/10.2147/PRBM.S304635.
14) Sapci O, Elhai JD, Amialchuk A, Montag C. The relationship between smartphone use and students' academic performance. *Learning and Individual Differences*. 2021;89:1–9. https://doi.org/10.1016/j.lindif.2021.102035.
15) Rajendran S, Chamundeswari S, Amitan A, Sinha. Predicting the academic performance of middle- and high-school students using machine learning algorithms. *Social Sciences & Humanities Open*. 2022;6(1):1–15. https://doi.org/10.1016/j.ssaho.2022.100357.
16) Amez S, Baert S. Smartphone use and academic performance: A literature review. *International Journal of Educational Research*. 2020;103:1–8. https://doi.org/10.1016/j.ijer.2020.101618.
17) Islam F, Krishna A, Kumar D, Kumari S. Factors Influencing Academic Performance: An Empirical Study Using Predictive Analytics. *Multidisciplinar (Montevideo)*. 2025;3(51):1–7. https://doi.org/10.62486/agmu202551.
18) Junejo NUR, Huang Q, Dong X, Wang C, Zeb A, Humayoo M, et al. SAPPNet: students' academic performance prediction during COVID-19 using neural network. *Scientific Reports*. 2024;14:1–16. https://doi.org/10.1038/s41598-024-75242-2.
19) Sarumaha MS. Smartphone Use and Academic Performance among Undergraduate Students: Analysis of Systematic Reviews. *International Journal of Multidisciplinary Approach Research and Science*. 2024;2(2):638–645. https://doi.org/10.59653/ijmars.v2i02.657.
20) Oppong SO. Predicting Students' Performance Using Machine Learning Algorithms: A Review. *Asian Journal of Research in Computer Science*. 2023;16(3):128–148. https://doi.org/10.9734/AJRCOS/2023/v16i3351.
21) Belina RR, Beena TLA, Savarimuthu C. Canberra Match Normalization-Enhanced Decision Stump Classifier for Predicting Academic Performance in the Context of Smartphone Addiction. *International Journal of Modern Education and Computer Science (IJMECS)*. 2025;17(1):59–71. https://doi.org/10.5815/ijmecs.2025.01.05.
22) Belina RR, Beena TLA. Elasticnet Regressive Bagging Classification for Student Academic Performance Prediction Based on Smartphone Addiction. *International Journal of Intelligent Systems and Applications in Engineering*. 2024;12(4):1709–1716. Available from: https://ijisae.org/index.php/IJISAE/article/view/6470/5305.