IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 5, October 2025, pp. 4113~4122
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i5.pp4113-4122 ❒ 4113

Educational data mining approach for predicting student performance and behavior using deep learning techniques

Muniappan Ramaraj[1], Sabareeswaran Dhendapani[2], Jothish Chembath[3], Selvaraj Srividhya[4], Nainan Thangarasu[5], Bhaarathi Ilango[1]
[1]Department of Computer Science, Rathinam College of Arts and Science, Coimbatore, India
[2]Department of Computer Science, L.R.G. Government College of Arts and Science for Women, Tiruppur, India
[3]Department of Computer Science, Presidency University, Bengaluru, India
[4]Department of Computer Science, KPR College of Arts Science and Research, Coimbatore, India
[5]Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, India

Article Info

Article history:

Received Sep 28, 2024
Revised Jun 27, 2025
Accepted Aug 6, 2025

Keywords:

Artificial neural networks
Deep learning
Educational data mining
Long short-term memory
Predictive analytics
Student performance prediction

ABSTRACT

Educational data mining (EDM) uncovers insights from large datasets collected from various educational platforms, such as online learning systems, student information databases, and classroom tools. EDM helps educators identify hidden patterns that improve teaching strategies, personalize learning experiences, and predict student performance. Predicting student success has become a key focus of EDM, allowing institutions to implement targeted interventions and personalized support. The dataset included academic achievement grades from 1,001 students enrolled in various courses during the fall semester across multiple years, to demonstrate how proposed models provide more accurate predictions compared to traditional machine learning methods. Models such as you only look once (YOLO), fast region-based convolutional neural networks (Fast RCNN), artificial neural networks (ANNs), and long short-term memory (LSTM) networks are used to capture complex, non-linear relationships within the data. The comparative analysis shows that these deep learning models significantly outperform traditional techniques, such as decision trees and support vector machines (SVMs). The results indicate that proposed method offers improved predictive accuracy, enabling educational institutions to identify at-risk students and deliver tailored interventions. This study highlights the potential of enhanced method to transform personalized education and enhance student success by better understanding individual learning needs and behaviors.

Corresponding Author:

Muniappan Ramaraj
Department of Computer Science, Rathinam College of Arts and Science
Coimbatore, India
Email: ramaraj.phdcs@gmail.com

1. INTRODUCTION

Educational data mining (EDM) is a growing field dedicated to uncovering valuable patterns and insights from extensive educational datasets. By effectively utilizing this data, institutions can forecast student outcomes, detect students at risk, and customize educational interventions to meet individual needs [1]. Recent progress in machine learning and deep learning has led to the creation of highly accurate models for predicting student performance [2]. Although traditional machine learning models, such as decision trees and support vector machines (SVMs), have been applied to this task, deep learning techniques have demonstrated significant potential due to their capacity to handle large datasets and uncover intricate, non-linear relationships [3]. This paper seeks to investigate the application of deep learning models for predicting student performance, with a particular emphasis on artificial neural networks (ANNs) and long short-term memory (LSTM) networks [4]. In many virtual learning environments (VLEs), however, datasets can be sparse or imbalanced, making simpler models such as decision trees, logistic regression, or even ensemble methods like random forests more practical in real-world applications [5].

Learning analytics (LA) typically focuses on real-time, actionable insights for immediate interventions, while EDM emphasizes discovering patterns in large datasets through data mining techniques. The problem arises when attempting to combine the real-time focus of LA with the data-driven discovery of EDM, as each operates on different temporal and methodological planes [6]. A deep cognitive diagnosis model (DCDM) for predicting students' performance focuses on enhancing how accurately we can assess student knowledge based on their responses to various assessments [7]. Another issue is the need for large-scale, high-quality data to train such models effectively. In many educational contexts, data may be scarce, noisy, or imbalanced, especially in terms of assessments for specific learning domains or minority student groups [8]. Many deep learning-based knowledge tracing (DLKT) models are trained on specific types of data (e.g., online learning platforms and standardized tests). The survey seeks to address the problem of how well DLKT models generalize across diverse learning contexts. However, the problem remains that there is limited research exploring the direct role artificial intelligence (AI) can play in enhancing academic outcomes by focusing on both study strategies and learning disabilities [9]. A novel machine learning model, random grouping-based deep multi-modal learning (RG-DMML), which is coupled with an ensemble learning algorithm. This model integrates various data sources, such as academic records and demographic information, and applies deep learning techniques to enhance prediction accuracy [10]. Educational institutions struggle to identify at-risk students early enough to intervene effectively.

2. LITERATURE REVIEW

EDM has seen rapid growth over the last decade, driven by the increasing availability of educational data from online learning platforms, student information systems, and other digital tools. Early research focused on rule-based systems and statistical models, which, while effective in certain scenarios, struggled to scale with increasing data complexity. Rathi *et al.* [11] has presents the hybrid approach combining the self-supervised robust optimization algorithm (SS-ROA) and deep LSTM networks. This model leverages the strengths of deep learning in handling time-series data while optimizing feature selection and model training using the SS-ROA technique. Ding [12] has illustrate on deep learning models can analyze student movements through video data, providing real-time corrections or feedback on technique and posture. In music, AI-driven models can assess pitch, timing, and expression during performances, offering students detailed feedback on areas for improvement. Aulakh *et al*. [13] aims to examine the intersection of e-learning and EDM during the COVID-19 pandemic. It explores various EDM methods applied in e-learning, such as clustering, classification, and regression analysis.

Sarker *et al.* [14] analyzing students' academic performance through EDM has emerged as a valuable approach for improving educational outcomes and institutional decision-making. Feng and Fan [15] has investigate how EDM can improve the learning process by evaluating learning behaviors, predicting student success, and visualizing data in a way that supports decision-making in education. Deng *et al*. [16] has introduces a novel deep learning-based predictive model, capable of analyzing various factors such as self-esteem levels, tendencies towards individualism, and their combined impact on performance metrics. Lam *et al*. [17] introduces a robust framework that leverages machine learning techniques to accurately predict student performance, enabling proactive identification of learners at academic risk. By utilizing algorithms such as k-means, hierarchical clustering, and density-based spatial clustering of applications with noise (DBSCAN), the study seeks to uncover patterns that can inform educators about the diverse needs of their students. Peng *et al.* [18] the achievement of this research lies in its ability to facilitate targeted interventions, personalized learning pathways, and ultimately enhance educational outcomes.

Rejeb *et al*. [19] aims to examine how ChatGPT is being utilized in various educational contexts and to assess its influence on teaching methods, learning experiences, and overall educational outcomes. Bhardwaj *et al.* [20] demonstrates that deep learning models, such as convolutional neural networks (CNNs) and LSTM networks, are effective tools for predicting and analyzing student engagement in e-learning environments. It aims to identify patterns of engagement, predict student behaviors, and provide personalized interventions to improve learning outcomes. Al Ka'bi [21] has introduces a novel AI algorithm and deep learning techniques tailored for enhancing the quality of higher education. Lin *et al*. [22] aims to streamline learning processes, improve educational outcomes, and optimize institutional management by providing personalized learning experiences, predictive analytics, and automated administrative tasks. Farhood *et al.* [23] contributes to the field of EDM by introducing generative adversarial networks (GANs) as a novel approach for improving student outcome predictions. The focus is on extracting meaningful patterns and insights from textual or communication data generated during learning processes, such as online discussions, written assignments, or feedback. Riaz *et al.* [24] introduces TransLSTM, a novel hybrid architecture combining the strengths of LSTM and Transformer models to perform fine-grained suggestion mining.

3. METHOD

The methodology for predicting student performance through EDM, this study employs a multi-step methodology utilizing deep learning techniques [25]. The approach begins with data collection, where academic records, demographic details, and behavioral patterns are aggregated. The data undergoes preprocessing to clean and normalize it, followed by feature selection to identify the most relevant attributes for prediction [26].

Figure 1 illustrates on the predicting student performance using deep learning in EDM involves several key steps. Initially, a dataset comprising academic records, demographic details, and behavioral data is collected and preprocessed to handle missing values and normalize features. To identify the most relevant features for prediction, feature selection is carried out, followed by splitting the data into training and testing sets to ensure reliable model evaluation. Advanced deep learning models, such as you only look once (YOLO), fast region-based convolutional neural networks (Fast RCNN), ANNs, and LSTM networks, are employed to capture complex patterns within the data [27]. These models are trained on the training set and assessed on the testing set, using metrics like accuracy, precision, and recall to gauge their performance. A comparative analysis is performed against traditional machine learning models, including decision trees and SVMs, to highlight the superior predictive accuracy of deep learning techniques. This methodology aims to provide precise insights into student performance, enabling more effective and targeted educational interventions.

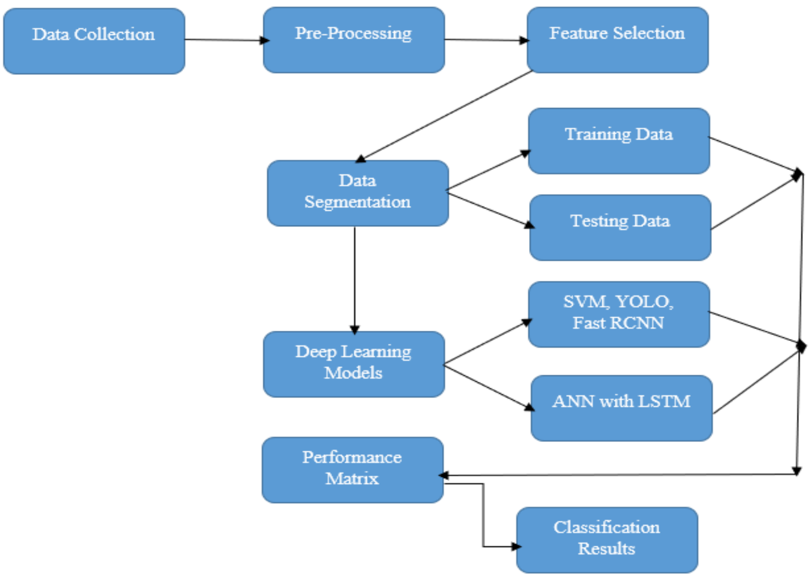

Figure 1. Proposed educational data mining models

3.1. SVM method used for EDM with spatial pyramid pooling

SVM is a powerful machine learning algorithm used for both classification and regression tasks. SVM is often applied to predict student performance by classifying students into different performance categories or predicting continuous scores [28]. Given a dataset of student features (such as grades, attendance, and demographic data), the goal is to classify students into categories like pass/fail, high/medium/low performance, or predict their final scores. The primary objective of SVM is to identify a hyperplane that optimally separates data points (students) into distinct classes. In student performance prediction, the hyperplane separates students based on their performance levels.

$$\omega^T x + b = 0 \tag{1}$$

$$f(x) = \omega^T x + b \tag{2}$$

Where ω is the weight vector (which determines the orientation of the hyperplane), x is the feature vector (input data, such as student features), b is the bias (offset from the origin), $\omega^T x + b = 0$ defines the hyperplane, and this is the decision boundary. If f(x)≥0 the student is classified into one category (e.g., pass), If f(x)<0 the student is classified into the other category (e.g., fail).

3.2. YOLO method used for EDM with spatial pyramid pooling

EDM, to the direct application of YOLO for student performance prediction is unconventional, as YOLO is fundamentally designed for image-based tasks. However, with some creative modification, YOLO-like architectures could theoretically be adapted for EDM tasks, especially if image-like data representations (e.g., heatmaps, time series, or visual patterns of student activity) are used [29]. n traditional YOLO, the objective is to predict bounding boxes around objects in an image and classify them. T The algorithm segments the image into an S×S grid, where each grid cell predicts multiple bounding boxes along with corresponding confidence scores and class probabilities.

$$SPP = \lambda_p \sum_{i=1}^{n}(P(TP_i) - P(PP_i))^2 + \lambda_c \sum_{i=1}^{n}(TC_i - PC_i)^2 + \lambda_f \sum_{i=1}^{n}\sum_{j=1}^{m}(f_{i,j}^{t} - f_{i,j}^{p})^2 \quad (3)$$

Where $\lambda_p, \lambda_c, \lambda_f$ are hyperparameters that control the relative importance of each loss term, $(P(TP_i) - P(PP_i))$ are the true and predicted probabilities of student iii belonging to a certain performance class, $(f_{i,j}^{t} - f_{i,j}^{p})$ are the true and predicted feature values for student i and feature j.

3.3. Fast RCNN method used for EDM with spatial pyramid pooling

Fast RCNN is a computer vision algorithm typically used for object detection in images. While it's not directly applicable to EDM, it can creatively adapt the principles of Fast RCNN for student performance prediction. The idea is to leverage its underlying framework for analyzing segmented data regions and making predictions, and map these concepts onto the features and performance prediction tasks in EDM.

$$L_c = -\sum_{i=1}^{n}\sum_{j=1}^{m} y_{ij} \log P\left(P_{ij}\right) \quad (4)$$

Where y_{ij} is the true performance class for student i's region j, $\log P\left(P_{ij}\right)$ is the predicted probability of the true class for region j. For each student, we make predictions for each region of features and then aggregate these to make a final decision about the student's overall performance. The algorithm can classify performance or predict scores for each feature set and then aggregate these predictions to make a final decision on student performance.

3.4. ANN with LSTM method used for EDM with spatial pyramid pooling

ANNs combined with LSTM units are widely used for time series prediction and sequential data modeling. In EDM, this combination can be highly effective for student performance prediction, especially when there is a temporal aspect to the data (e.g., predicting performance over multiple semesters or assessments). A neural network with fully connected layers, typically used for learning from non-sequential, static data. In EDM, an ANN can be used to model relationships between student features (e.g., grades, attendance, and assignment scores) and their final performance. LSTMs are particularly useful in modeling time-dependent relationships, such as a student's performance over multiple periods or tasks.

$$L_c = -\sum_{i=1}^{n} y_i \log P(P_i) \quad (5)$$

$$L_r = {}^{1}\!/_{n} \sum_{i=1}^{n}(y_i - \widehat{y_i})^2 \quad (6)$$

Where y_i is the true performance class for student i, $\log P(P_i)$ is the predicted probability for the true class. Then, next equation y_i is the true performance class for student I, $\widehat{y_i}$ is the predicted performance score. For each student, the LSTM processes the sequential features, and the ANN layers make the final performance prediction based on the learned representation.

4. RESULTS AND DISCUSSION

The ANN and LSTM models demonstrated superior performance in predicting student outcomes compared to traditional machine learning methods. The LSTM, in particular, excelled at capturing temporal patterns in the data, leading to higher accuracy in predicting long-term student performance. When compared to decision trees and SVMs, deep learning models showed a marked improvement in prediction accuracy. The ANN and LSTM models reduced the error rate by approximately 10-15%, highlighting their effectiveness in identifying non-linear relationships and complex patterns in student data.

4.1. Classification accuracy

The classification accuracy for student performance prediction in the real time dataset, to find the accuracy is a commonly used metric to evaluate the performance of a classification model. Accuracy measures the proportion of correct predictions made by the model relative to the total number of predictions.

$$A = {}^{1}/_{n} \sum_{i=1}^{n} 1\,(\hat{y_i} = y_i) \tag{7}$$

Where n is the total number of students, $\hat{y_i}$ is predicted class either 0 or 1, y_i is the true class of i.

In the context of student performance prediction, accuracy measures how well the model classifies students into the correct performance categories (e.g., pass/fail, high/medium/low performance). Let's say the model is predicting whether a student will pass or fail based on their features (such as grades, attendance, and assignments). If the model classifies a student as passing, and the student actually passes, it is a true positive (TP). If it predicts failure, and the student fails, it is a true negative (TN).

4.2. Precision, recall, and F-measures

Precision calculates the proportion of TP predictions among all positive predictions (including both TP and false positives (FP)). It addresses the question: "Out of all the students predicted to succeed, how many actually did?" Recall, also referred to as sensitivity or the true positive rate (TPR), measures the ratio of TP predictions to all actual positives (TP and false negatives (FN)). It answers: "Out of all the students who actually succeeded, how many were correctly predicted by the model?" The F1-score, which is the harmonic mean of precision and recall, offers a single metric that balances the two. This score is particularly valuable when there is a need to balance precision and recall, such as when both FP and FN have significant consequences.

$$Precision = \frac{TP}{TP+FP} \tag{8}$$

$$Recall = \frac{TP}{TP+FN} \tag{9}$$

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{10}$$

Where TP refer to the number of cases where positive outcomes are correctly predicted (e.g., students correctly identified as passing). FP represent instances where the model incorrectly predicts a positive outcome (e.g., students predicted to pass but actually fail). FN are the cases where the model wrongly predicts a negative outcome (e.g., students predicted to fail but actually pass). The F1-score ranges between 0 and 1, with 1 signifying perfect precision and recall.

4.3. Receiver operating characteristic

The receiver operating characteristic (ROC) curve is an effective tool for assessing the performance of a classification model, especially when predicting student outcomes (such as pass/fail or different performance categories). It is a graphical representation that illustrates the trade-off between the TPR and the false positive rate (FPR) as the decision threshold changes. The ROC curve plots the TPR against the FPR for varying thresholds, with each point on the curve reflecting a different threshold used by the model to classify students as passing or failing. The x-axis represents the FPR, while the y-axis represents the TPR.

$$TPR = \frac{TP}{TP+FN},\ FPR = \frac{FP}{FP+TN} \tag{11}$$

This represents one point on the ROC curve. By varying the threshold, you generate additional TPR and FPR values to plot the entire curve.

4.4. Time calculation

EDM for student performance prediction, calculating the time complexity of the algorithms and the overall prediction process can provide insights into the efficiency of the model. Time complexity generally refers to the amount of computational time an algorithm takes as a function of the length of the input.

$$o(n.p^2) \tag{12}$$

Where n is the number of instances and p is the number of features. The computation involves calculating the coefficients using the least squares method.

Table 1 represents that the comparison of various methods for student performance calculation reveals distinct differences in accuracy and processing time. Among the techniques evaluated SVM, Fast RCNN, ANN with LSTM, and YOLO. YOLO demonstrates the highest accuracy, achieving 0.86 during training and 0.93 during testing, while also exhibiting the shortest processing time of 2.7 seconds for training and 1.84 seconds for testing. Fast RCNN follows closely, with training and testing accuracies of 0.76 and 0.87, respectively, taking slightly longer at 3.7 seconds for training. ANN with LSTM shows competitive performance with accuracies of 0.74 for training and 0.87 for testing, though it requires more time than YOLO at 3.1 seconds for training. SVM, while effective, records the lowest accuracy of 0.73 during training and 0.82 during testing, taking 3.4 seconds for training. Overall, YOLO stands out as the most efficient and accurate method in this comparison.

In the Figure 2, the performance comparison of various methods like SVM, Fast RCNN, ANN with LSTM, and YOLO illustrates that YOLO achieves the highest accuracy of 0.93 during testing while also being the fastest, requiring only 1.84 seconds. Fast RCNN and ANN with LSTM provide competitive accuracies of 0.87 but take longer, with Fast RCNN at 2.7 seconds and ANN with LSTM at 2.02 seconds. In contrast, SVM shows the lowest accuracy of 0.82 during testing and a training time of 3.4 seconds. Overall, the data indicate that YOLO is the most effective method in terms of both accuracy and processing time.

Table 1. Comparison of student performance data for testing and training process

Methods	Training		Testing	
	Accuracy	Time	Accuracy	Time
SVM	0.73	3.4	0.82	2.9
Fast RCNN	0.76	3.7	0.87	2.7
ANN with LSTM	0.74	3.1	0.87	2.02
YOLO	0.86	2.7	0.93	1.84

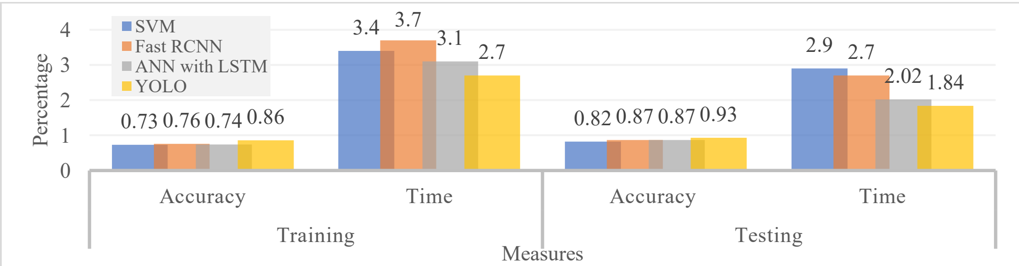

Figure 2. Performance comparison of proposed methods

Table 2 presents a comparative analysis of various methods for evaluating student performance, specifically focusing on precision, recall, and F-measure during training and testing phases. Among the methods assessed, Fast RCNN demonstrates the highest precision (0.752) and F-measure (0.721) during training, indicating its superior ability to identify relevant instances. The ANN with LSTM closely follows, with a training precision of 0.748 and an F-measure of 0.723, showcasing its effectiveness in handling sequential data. SVM yields respectable results, with a training precision of 0.735, while YOLO shows the lowest performance across metrics, particularly with a training precision of 0.706. During testing, Fast RCNN again leads with a precision of 0.725, followed by ANN with LSTM at 0.723. Overall, Fast RCNN and ANN with LSTM exhibit consistent performance, suggesting their potential suitability for applications requiring reliable performance metrics in educational contexts.

In Figure 3 compares the performance metrics precision, recall, and F-measure of various methods used to evaluate student performance. The methods analyzed include SVM, Fast RCNN, ANN with LSTM, and YOLO, with results presented for both training and testing phases. This analysis highlights the effectiveness of each method in accurately assessing student performance metrics.

Table 2. Overall performance comparison of the proposed measures

Methods	Training			Testing		
	Precision	Recall	F-measure	Precision	Recall	F-measure
SVM	0.735	0.735	0.704	0.693	0.697	0.687
Fast RCNN	0.752	0.746	0.721	0.725	0.734	0.712
ANN with LSTM	0.748	0.746	0.723	0.723	0.734	0.732
YOLO	0.706	0.713	0.692	0.712	0.701	0.685

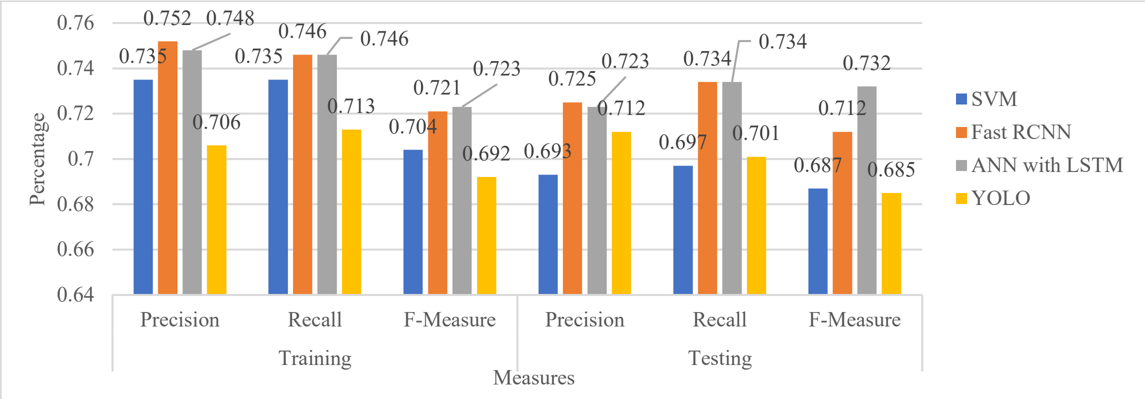

Figure 3. Overall performance comparison of the different measures

The performance of various methods for student performance prediction is illustrated on the Table 3 through their ROC values, which indicate the models' effectiveness in distinguishing between classes. Among the evaluated techniques, the SVM achieved the highest training ROC value of 0.804 and a testing ROC of 0.81, suggesting a robust model. Fast RCNN follows with a notable training ROC of 0.863; however, its testing ROC drops to 0.746, indicating potential overfitting. The ANN with LSTM achieved a training ROC of 0.826 but also faced a decline in testing performance at 0.717. Finally, the YOLO method demonstrated the lowest performance overall, with training and testing ROC values of 0.713 and 0.693, respectively. This comparison highlights the varying effectiveness of these models, emphasizing the need for further optimization, particularly for those with lower testing ROC scores.

Figure 4 illustrates the performance comparison of various methods using ROC values derived from a real-time dataset for student performance prediction. The evaluated techniques include SVM, Fast RCNN, ANN with LSTM, and YOLO. Among these, SVM stands out with the highest training ROC of 0.804 and a testing ROC of 0.81, indicating its reliability. Fast RCNN has a strong training ROC of 0.863, but its testing ROC declines to 0.746, suggesting potential overfitting. The ANN with LSTM achieves a training ROC of 0.826, with a testing ROC of 0.717, reflecting a similar trend. Conversely, the YOLO method shows the lowest performance, with training and testing ROC values of 0.713 and 0.693, respectively. This comparison underscores the varied effectiveness of these models and highlights the need for optimization, especially for those with lower testing ROC scores.

Table 3. Performance comparison of ROC applied on student performance dataset

Methods	Training ROC	Testing ROC
SVM	0.804	0.81
Fast RCNN	0.863	0.746
ANN with LSTM	0.826	0.717
YOLO	0.713	0.693

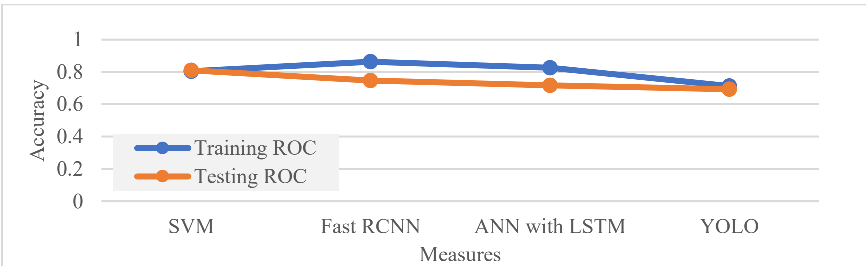

Figure 4. Performance comparison of ROC in real time dataset

5. CONCLUSION

This study demonstrates the transformative potential of deep learning in the realm EDM. In this article has proposed many advanced deep learning techniques introduced such as YOLO, Fast RCNN, ANNs, and LSTM networks, the research reveals that deep learning models significantly enhance the accuracy of student performance predictions compared to traditional machine learning methods. The deep learning approaches employed effectively capture intricate, non-linear relationships in diverse data sources, including academic assessments, demographic information, and student behaviors. The comparative analysis shows that these models outperform conventional techniques like decision trees and SVMs in terms of predictive accuracy. The findings underscore that deep learning can offer a more nuanced understanding of student performance and behavior, which is crucial for identifying at-risk students and implementing timely, personalized interventions. This capability allows educational institutions to better tailor their support strategies and improves overall student success. By integrating deep learning into educational practices, institutions can move beyond one-size-fits-all solutions and develop more effective, individualized approaches to learning. This study highlights the significant potential of deep learning to revolutionize personalized education, offering deeper insights into student needs and enhancing educational outcomes.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Muniappan Ramaraj	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓
Sabareeswaran Dhendapani		✓	✓	✓	✓	✓		✓	✓	✓			✓	✓
Jothish Chembath	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓
Selvaraj Srividhya		✓			✓	✓	✓	✓		✓	✓	✓		✓
Nainan Thangarasu	✓	✓		✓	✓	✓		✓	✓	✓			✓	✓
Bhaarathi Ilango	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest related to this work.

INFORMED CONSENT

We have obtained informed consent from all participants involved in the study.

ETHICAL APPROVAL

The authors have obtained the necessary permissions from the institutional ethics committee to conduct this work.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] X. Liu, "The educational resource management based on image data visualization and deep learning," *Heliyon*, vol. 10, no. 13, 2024, doi: 10.1016/j.heliyon.2024.e32972.
[2] B. Alnasyan, M. Basheri, and M. Alassafi, "The power of deep learning techniques for predicting student performance in virtual learning environments: a systematic literature review," *Computers and Education: Artificial Intelligence*, vol. 6, 2024, doi: 10.1016/j.caeai.2024.100231.
[3] R. Cerezo, J.-A. Lara, R. Azevedo, and C. Romero, "Reviewing the differences between learning analytics and educational data mining: towards educational data science," *Computers in Human Behavior*, vol. 154, 2024, doi: 10.1016/j.chb.2024.108155.
[4] L. Gao, Z. Zhao, C. Li, J. Zhao, and Q. Zeng, "Deep cognitive diagnosis model for predicting students' performance," *Future Generation Computer Systems*, vol. 126, pp. 252–262, 2022, doi: 10.1016/j.future.2021.08.019.

[5] X. Song, J. Li, T. Cai, S. Yang, T. Yang, and C. Liu, "A survey on deep learning based knowledge tracing," *Knowledge-Based Systems*, vol. 258, 2022, doi: 10.1016/j.knosys.2022.110036.
[6] A. Bressane *et al.*, "Understanding the role of study strategies and learning disabilities on student academic performance to enhance educational approaches: a proposal using artificial intelligence," *Computers and Education: Artificial Intelligence*, vol. 6, 2024, doi: 10.1016/j.caeai.2023.100196.
[7] K. Okoye, J. T. Nganji, J. Escamilla, and S. Hosseini, "Machine learning model (RG-DMML) and ensemble algorithm for prediction of students' retention and graduation in education," *Computers and Education: Artificial Intelligence*, vol. 6, 2024, doi: 10.1016/j.caeai.2024.100205.
[8] S. Maniyan, R. Ghousi, and A. Haeri, "Data mining-based decision support system for educational decision makers: Extracting rules to enhance academic efficiency," *Computers and Education: Artificial Intelligence*, vol. 6, 2024, doi: 10.1016/j.caeai.2024.100242.
[9] A. Harb *et al.*, "Diverse distant-students deep emotion recognition and visualization," *Computers and Electrical Engineering*, vol. 111, 2023, doi: 10.1016/j.compeleceng.2023.108963.
[10] T. Shaik, X. Tao, C. Dann, H. Xie, Y. Li, and L. Galligan, "Sentiment analysis and opinion mining on educational data: a survey," *Natural Language Processing Journal*, vol. 2, 2023, doi: 10.1016/j.nlp.2022.100003.
[11] S. Rathi, K. K. Hiran, and S. Sakhare, "Affective state prediction of E-learner using SS-ROA based deep LSTM," *Array*, vol. 19, 2023, doi: 10.1016/j.array.2023.100315.
[12] J. Ding, "Deep learning perspective on the construction of SPOC teaching model of music and dance in colleges and universities," *Systems and Soft Computing*, vol. 6, 2024, doi: 10.1016/j.sasc.2024.200137.
[13] K. Aulakh, R. K. Roul, and M. Kaushal, "E-learning enhancement through educational data mining with COVID-19 outbreak period in backdrop: a review," *International Journal of Educational Development*, vol. 101, 2023, doi: 10.1016/j.ijedudev.2023.102814.
[14] S. Sarker, M. K. Paul, S. T. H. Thasin, and M. A. M. Hasan, "Analyzing students' academic performance using educational data mining," *Computers and Education: Artificial Intelligence*, vol. 7, 2024, doi: 10.1016/j.caeai.2024.100263.
[15] G. Feng and M. Fan, "Research on learning behavior patterns from the perspective of educational data mining: evaluation, prediction and visualization," *Expert Systems with Applications*, vol. 237, 2024, doi: 10.1016/j.eswa.2023.121555.
[16] J. Deng, X. Huang, and X. Ren, "A multidimensional analysis of self-esteem and individualism: A deep learning-based model for predicting elementary school students' academic performance," *Measurement: Sensors*, vol. 33, 2024, doi: 10.1016/j.measen.2024.101147.
[17] P. X. Lam, P. Q. H. Mai, Q. H. Nguyen, T. Pham, T. H. H. Nguyen, and T. H. Nguyen, "Enhancing educational evaluation through predictive student assessment modeling," *Computers and Education: Artificial Intelligence*, vol. 6, 2024, doi: 10.1016/j.caeai.2024.100244.
[18] J. Peng *et al.*, "DeepRisk: a deep learning approach for genome-wide assessment of common disease risk," *Fundamental Research*, vol. 4, no. 4, pp. 752–760, 2024, doi: 10.1016/j.fmre.2024.02.015.
[19] A. Rejeb, K. Rejeb, A. Appolloni, H. Treiblmaier, and M. Iranmanesh, "Exploring the impact of ChatGPT on education: a web mining and machine learning approach," *The International Journal of Management Education*, vol. 22, 2024, doi: 10.1016/j.ijme.2024.100932.
[20] P. Bhardwaj, P. K. Gupta, H. Panwar, M. K. Siddiqui, R. M.-Menendez, and A. Bhaik, "Application of deep learning on student engagement in e-learning environments," *Computers & Electrical Engineering*, vol. 93, 2021, doi: 10.1016/j.compeleceng.2021.107277.
[21] A. Al Ka'bi, "Proposed artificial intelligence algorithm and deep learning techniques for development of higher education," *International Journal of Intelligent Networks*, vol. 4, pp. 68–73, 2023, doi: 10.1016/j.ijin.2023.03.002.
[22] C.-C. Lin, E. S. J. Cheng, A. Y. Q. Huang, and S. J. H. Yang, "DNA of learning behaviors: A novel approach of learning performance prediction by NLP," *Computers and Education: Artificial Intelligence*, vol. 6, 2024, doi: 10.1016/j.caeai.2024.100227.
[23] H. Farhood, I. Joudah, A. Beheshti, and S. Muller, "Advancing student outcome predictions through generative adversarial networks," *Computers and Education: Artificial Intelligence*, vol. 7, 2024, doi: 10.1016/j.caeai.2024.100293.
[24] S. Riaz, A. Saghir, M. J. Khan, H. Khan, H. S. Khan, and M. J. Khan, "TransLSTM: a hybrid LSTM-Transformer model for fine-grained suggestion mining," *Natural Language Processing Journal*, vol. 8, Sep. 2024, doi: 10.1016/j.nlp.2024.100089.
[25] Y. Wang, Y. Zhang, M. Liang, R. Yuan, J. Feng, and J. Wu, "National student loans default risk prediction: A heterogeneous ensemble learning approach and the SHAP method," *Computers and Education: Artificial Intelligence*, vol. 5, 2023, doi: 10.1016/j.caeai.2023.100166.
[26] R. Song, F. Pang, H. Jiang, and H. Zhu, "A machine learning based method for constructing group profiles of university students," *Heliyon*, vol. 10, no. 7, 2024, doi: 10.1016/j.heliyon.2024.e29181.
[27] M. Bilal, H. Israr, M. Shahid, and A. Khan, "Sentiment classification of roman-urdu opinions using naïve Bayesian, decision tree and KNN classification techniques," *Journal of King Saud University - Computer and Information Sciences*, vol. 28, no. 3, pp. 330–344, 2016, doi: 10.1016/j.jksuci.2015.11.003.
[28] X. Guo *et al.*, "Longitudinal machine learning prediction of non-suicidal self-injury among Chinese adolescents: a prospective multicenter Cohort study," *Journal of Affective Disorders*, vol. 392, 2026, doi: 10.1016/j.jad.2025.120110.
[29] S. Gupta and J. Choudhary, "An efficient test suit reduction methodology for regression testing," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 34, no. 2, pp. 1336–1343, 2024, doi: 10.11591/ijeecs.v34.i2.pp1336-1343.

BIOGRAPHIES OF AUTHORS

Muniappan Ramaraj is working as an Assistant Professor in the Department of Computer Science at Rathinam College of Arts and Science, Coimbatore. He holds a Ph.D., degree in Computer Science at Bharathiar University in the year of 2020 with specialization in data mining with image process and also fuzzy logic in the image analysis. His research areas are data mining, image processing, fuzzy logic, pattern recognition, and deep learning concept. He has published more research article in the reputed various national and international journals and also filed the patents in the same field. He has a reviewer of many international journals including with IEEE, ASTESJ, and JERS. He can be contacted at email: ramaraj.phdcs123@gmail.com.

Dr. Sabareeswaran Dhendapani is working as an Assistant Professor in the Department of Computer Science at L.R.G. Government College of Arts and Science for Women, Tiruppur. He holds a Ph.D., degree in computer science at Karpagam Academy of Higher Education in the year of 2020 with specialization in data mining. His research areas are data mining, image processing, fuzzy logic, and pattern recognition. He has published more research article in the reputed various national and international journals and also filed the patents in the same field. He can be contacted at email: sabaredhandapani@gmail.com.

Dr. Jothish Chembath completed his Ph.D. from Karpagam deemed to be University, Coimbatore, Tamilnadu in the year 2020, under the title "Next web page prediction using enhanced preprocessing and ensemble clustering based hybrid Markov model using web log data". More than 15 research journals have been published in reputed journals which are either Scopus indexed, Web of Science, and international journals of repute, which would help the Internet mechanism to predict the user's intention and interest, when internet is browsed using mathematical model algorithms after cleaning the internet data, thereby giving accurate predictions using mathematical models. He also has attended 3 international conferences held in the Sultanate of Oman, Malaysia, and Cochin. Currently author is working in Presidency University, Bengaluru. The teaching and research experience is around two decades starting from 2003 to toll date. He can be contacted at email: jothishchembath12@gmail.com.

Selvaraj Srividhya is working as an Associate Professor in the Department of Computer Science at KPR College of Arts Science a Research, Coimbatore. She has more than a decade of teaching experience in different verticals. She is a reviewer in various international journals and life-time member in IAENG. Her research area includes data mining, machine learning, and deep learning. She has about 10+ research papers published in national and international journals. She also presented more than 15 papers in national and international conferences and also published many books. She can be contacted at email: vidhyasai14@gmail.com.

Dr. Nainan Thangarasu is currently working as an Assistant Professor in the Department of Computer Science, at Karpagam Academy of Higher Education, Coimbatore. He is greatly fascinated with the advanced computing technology and research programs is cluster computing, cryptography and network security, cloud computing, artificial intelligent system, information security in large database and data mining as well as the strong teaching experience. His doctoral dissertation also focuses on advanced security systems with cloud computing, and he has published more than 13 publications in reputed journals, which he finds would be a great addition to the success of your teaching and research department. He can be contacted at email: drthangarasu.n@kahedu.edu.in.

Bhaarathi Ilango is working as an Assistant Professor in the Department of Computer Science at Rathinam College of Arts and Science, Coimbatore. He holds a NET. Qualification in Computer Science at NTA in the year of 2023. with specialization in data mining with image process and also fuzzy logic in the image analysis. His research areas are data mining, in the various fields. He has published more research article in the reputed various national and international journals and also filed the patents in the same field. He can be contacted at email: bhaarathi.cs@rathinam.in.