A Deep Residual Auto-Encoding Approach with Metaheuristic Optimization for Educational Quality Improvement

G.Sathya[1]*, S.Vijayprasath[2],S.Gopalakrishnan[3],Sridevi T[4]

[1]Assistant Professor, Department of Electronics and Communication Engineering, P.S.R. Engineering College, Sivakasi-626140, Virudhunagar, Tamil Nadu, India. Email: sathyag@psr.edu.in

[2]Department of Electronics and Communication Engineering PSNA College of Engineering and Technology,

Dindigul 624 622, Tamil Nadu India. Email ID :vijayprasathme@psnacet.edu.in

[3]Professor, Department of Electronics and Communication and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS), Thandalam, Chennai-602105, Tamil Nadu, India. Email: gopalakrishnans.sse@saveetha.com

[4]Assistant Professor, Department of Electrical and Electronics Engineering, Peri institute of technology, Chennai - 600 048,Tamil Nadu, India. Email: devikamal05@gmail.com

Abstract: The technology that deals with analyzing the educational data is in vogue in the modern educational setting to help to determine the students' behavior and learning process. The prediction of students' learning outcomes is essential because it assists the educators in situations in such ways as; early identification of learners who are performing poorly, formulation of the necessary approaches to address the matter, and therefore improving the learners' performances. Therefore, despite the development of many concepts in this field, some of them have vagueness in most of the results, low compatibility with complicated data structures, or poor quality of other aspects such as optimization that makes predictions less accurate. Hence, the proposed work develops a smart framework, AcadPredictor that employs advanced deep learning and optimization models for the prediction of students' academic performance. In the proposed system, the Regression Auto-Encoding Residual Network (Hyb-RARNet) model based on regression, auto-encoder and residual model is developed, which allows the model to distinguish as well as detect the sample data pattern and optimize the EDM. In order to further improve the prediction performance of classifier, the Honey Levy Explorer Optimization (HLEO) is used in this study for optimally computing the learning rate, which significantly enhances the speed and accuracy. Moreover, the Kaggle open source students dataset has been used to test and examine the outcomes of the proposed system using a range of evaluation measures

Keywords: : Education Data Mining (EDM), Students Performance Prediction, Deep Learning, Quality Education, Optimization, Innovation, Academics, Classification.

Introduction

The most fundamental indicator in educational advancement in any nation is student academic performance. In a nutshell gender, age, the teaching staff, and the students' learning all have an impact on children's academic progress (Fazil et al., 2024; Hachani et al., 2024). In higher learning, determining how well students are doing academically has attracted a lot of attention. Put in distinct ways, the

degree to which students meet both short- and long-term learning goals is referred to as their performance. According to its rankings, an outstanding academic record is an essential requirement for admission to a renowned university. Consequently, when an institution has a solid educational record and track record, its position in rankings rises (Arifin et al., 2024). From the point of view of the student, maintaining to obtain exceptional academic results boosts their chances of landing a job, as one of the main factors firms consider is outstanding academic performance. Education Data Mining (EDM) has been a source of many ideas for the recent changes in education. Several investigations have resulted in finding and developing new possibilities for creating technologically advanced learning systems that are made to fulfill students' requirements. The present methods and implementation tactics employed by EDM are vital elements for enhancing the learning atmosphere (Kocsis & Molnár, 2024). On this matter, considering the educational environment and machine learning methods - EDM helps to understand the student learning environment. The field of EDM studies the investigation, study and use of data mining techniques. Multidisciplinary approaches are used by the data mining discipline for success accomplishment. It offers a comprehensive process for drawing insightful conclusions from unprocessed data. The context in which modern educational facilities operate is complicated and highly competitive. Therefore, a great deal of institutions today confront issues in performance analysis, high-quality education, developing strategies for learning evaluation, and recognizing future demands (Forero-Corba & Bennasar, 2024).

Universities use student intervention techniques to help learners overcome hurdles they experience while they study. Entry-level student performance predictions and following periods in which students performance prediction aid institutions in the development and evolution of efficient approaches to intervention that benefit educators as well as administrators together. Student academic performance is another well-known cog that can increase the rate of educational development in any country and also can be used for evaluating both individual success and the achievements of educational establishments. In this related context, students' performances depend on so many parameters such as their gender, their age, the performance of teachers and teaching methods, and their own study habits. However, determining the reasons for student success in academic activities has been an area of interest in the education sector. This predictive capability is very important as it provides the simulative approaches to the objectives that capture the short- term and long-term learning desires of the students (Chen et al., 2024; Stasolla et al., 2025).

There seems to be a rising trend in the amount of students who got notices and have been forced to leave education. One of the possible causes is that students might be unable to properly evaluate and predict their capacity to choose suitable courses. Because the performance of students is a criterion for high-quality institutions that are founded on a great profile of their academic achievements, it is a vital responsibility for higher education facilities. Student performance can be measured in a number of ways. The educational program and educational assessment can be measured to figure out the performance of the students. Still, almost all of the research demonstrated that the achievement of students is determined by their completion. Academic achievements mean a lot to universities most especially to ranking because they are part of the primary determinants (Fazil et al., 2024). Achievements made in academics by the students benefit the institution through the upgraded recognition which leads to better rankings that increase resource mobilization, faculty and other student intakes.

From the students' point of view academic performance is critical in their lives as students as it determines the outcome in future (Deepa, 2025). Academic records are believed to be among the most crucial aspects that employers consider before hiring employees. It can therefore be concluded that good grades directly translate into improved employability to pave way to better and more opportunities. This is especially relevant in the present day world where employment opportunity is limited and one needs to have a leverage and academic qualifications offer the leverage (Luo et al., 2024; Sengupta et al., 2025). Consequently, it is evidenced that academic performance has significant worth both to the institution and the individual learner, which is why initiatives aimed at identifying and improving student outcomes are of great significance.

Over the last few years, the use of deep learning systems has been seen as a highly effective way of predicting students' performance levels. Deep learning models are promising due to their high capacity and skills in learning complex patterns and characteristics within the datasets which could not be captured by linear or logistic models Institutions' educational data in forms of student's engagement with the learning material, participation in the forums, performance in the quizzes, and collaboration tools usage can be analyzed with deep learning (Sheela et al., 2024; Boddu & Manimaran, 2025). The key is to train models on these different sources of data and allow them to make more accurate predictions of students' likely future performance; detect the poor performers early on, and suggest helpful interventions. Several benefits come with the application of deep learning in the case of prediction of students' performance. They are capable of handling and synthesizing data from different sources where features can also involve time series analysis of student behaviors along with spatial characteristics of activities (Faruque et al., 2024; Anjireddy & Jaisharma, 2024). For example, LSTM networks are known to be very effective in dealing with time series data due to which they can be effectively used for analyzing time series data of student activities. CNNs, on the other hand, would be well suited for tasks such as the identification of patterns within the data stream, like the nature and frequency of student engagement with content. This type of calibration gives a better insight of many factors that affect student performance, thus, offering better prediction results (Sheela et al., 2024).

However, the current state of deep learning methods for student performance predictions has the following main limitations. A major one on them is that these models are hard to interpret and this is because deep learning architectures are in most cases known as 'black boxes' whereby educators and stakeholders have to take a leap of faith when working with such models as they cannot explain what the model is doing or how it arrived at a given decision (Namoun & Alshanqiti, 2020; Silambarasan et al., 2024). Furthermore, deep learning practice involves finding reliable labeled data which is scarce in learning domains.

This shortcoming restricts the models' performance in transferring knowledge from the source to the target domain. Another issue consists of the computational requirements the training of these models consumes – data necessitate better hardware and sophisticated software, often unavailable at certain educational facilities (Vellela et al., 2022). In addition, there is an issue of over fitting, machines learn the data well enough to predict on the training data set but are less accurate when called upon to predict on other data sets, especially when the data set is not large enough or diverse enough (Waheed et al., 2020; Akour et al., 2020). Finally, these models also have drawbacks such as over-representing the inputs' prejudice in the process of decision-making and therefore arriving at biased decisions that harm specific students. Both these aspects form the core of changes that are required to make deep learning models more feasible and fair in the educational context. The unique contributions of the proposed work are given below:

- This research is to provide a new, state of the art high accuracy prediction method for student performance prediction in educational systems by employing and testing hybrid deep learning and optimization related techniques to develop the AcadPredictor model.
- In this way, the proposed approach combines the separate Hybrid Regression Auto-Encoding Residual Network (Hyb-RARNet) which brings together regressions auto-encodings and residual network techniques to reveal intricate data patterns and explore forecast accuracy in education data mining.
- The model parameters performance and accuracy are highly improved using the Honey Levy

Explorer Optimization (HLEO) algorithm, which fuses levy flight models to the honeybee foraging mechanism.

- In the analysis stage, a full system evaluation is performed to demonstrate the performance of the approach using several evaluation indicators.

The other portions of this paper are segregated into the remaining sections: Section 2 investigates some of the recent learning models used for students' performance prediction in educational context. It also provides clear insights about the problems and challenges correlated to the conventional prediction models. Then, the detailed overview and methodological explanations are presented with the flow in Section 3. Moreover, the performance outcomes, comparative results and findings are all presented along with datasets and evaluation metrics in Section 4. Finally, the overall paper is summarized with the future work in Section 5

Related Works

Here, we get more involved with the machine learning and deep learning models used for the purpose of performance forecast of students in learning institutions. It is on this premise that we look at the subcategories of these methodologies, not sparing any aspect: application, efficacy, or otherwise, that comes with them. At the same time, the paper analyses the issues and limitations associated with the use of such sophisticated approaches to the predictive analytics application in education (Wanda & Diqi, 2024). In this paper, we classified the recent developments in student performance prediction along with their drawbacks as a way of identifying the current state and direction for enhancement of this significantly important area in educational studies.

Albreiki, et al (Albreiki et al., 2021) conducted a systematic literature review to investigate different types of machine learning techniques used for student performance prediction. In this study, the major impacts of systemic performance indicators and advanced data collecting technologies on the evaluation and improvement of educational systems are addressed in this research. The concept of big data analytics in education has emerged as a consequence of its concentration on using cutting-edge machine learning and data mining methods to analyze and track large amounts of learning data. The study's main objective is to improve students' performance with the use of static and dynamic data information. Moreover, the study also highlights the possibilities for advancements in education data mining that use these techniques, but it additionally points out a lack of information available in the literature about corrective actions that give users immediate feedback. Dien, et al (Dien et al., 2020) investigated and analyzed different kinds of data mining methodologies for predicting students' performance in educational settings. It is essential that institutions of learning forecast how well students do academically because doing so can help individuals learn more effectively. This study indicates a method for deep learning that builds a model that predicts the success of pupils of future semesters using current educational achievement results as inputs from previous academic years. It does this by using the convolutional neural network on 1D data (CN1D) and the Long Short Term Memory (LSTM). Data transformation techniques are additionally taken into consideration to boost this model's capacity for prediction by strengthening its applicability over different time periods, settings, and individual performance instances. Despite the advantages, the suggested study has certain limitations such as lack of generalizability in attribute analysis, and lack of interpretability in prediction, which makes the system more complex for training. Ren, et al (Ren & Yu, 2024) has incorporated a new Learning Ability Adaptive Algorithm (LASA) for building an automatically generated student performance prediction model. The proposed framework is comprised of two principal parts one of which is learning ability modeling and the other part is the distributed alignment. By using this, the knowledge acquisition of learners is precisely estimated in this system. In essence, LAM is founded on what written replies to exercises contain in terms of specification of the students' fundamental abilities to learn, these are viewed as parametric distributions. Through integrating different response data from the variety of pupils, it learns these variables and transforms the acquired information into one new unified feature space through LAM. This change in structure affords a way to approach the problem of data heterogeneity characteristic of long-term educational data. Subsequently, in order to qualify the numerous asymmetric adjustments made on the newly generated uniform features, LTDA employs a number of procedures for equalizing the distribution of these features across academic decades. This coordination is necessary for reducing that impact of distribution shifts that are likely to threaten model performance at some interval.

Nanavaty, et al (Nanavaty & Khuteta, 2024) conducted a comprehensive study for predicting students' performance based on their learning abilities. The study offers a detailed understanding of technology-mediated learning activity pointing to all types of interactions with the educational content that can comprise learning materials, forums, quizzes, and collaboration platforms. The research has recorded an impressive 75% certainty in projecting learning outcomes therefore underlining the prospect of assembling various features to deliver an accurate estimate of performance. This high level of predictive accuracy suggests that the study has a rich understanding of the various factors that come into play in learning environments, specifically in online learning, and specifically the factors that impact the student success. The work successfully connects the theoretical analysis of knowledge in the field with its application, providing useful recommendations for education professionals and stakeholders. It does so to show how evidence-based approaches can be used to design content that is not only fair, but also suitable for the learning styles of students, and needs that each may have, when learning online. This is done not only to improve the general approach to obtaining knowledge as a whole but also focused on reaching out to low achievers at risk, with the general idea done in the belief that it can help erase divides in terms of educational opportunities by making available, easily accessible, and user-friendly online learning instruments. The conclusions suggested that the task of promoting academic accomplishment and active engagement requires a learning environment that is becoming less universally categorized and more adaptable to individual learners.

Lin, et al (Lin & Guo, 2024) utilized a multi-view hyper-graph neural network known as IMHNN is developed as part of the study to boost the student performance prediction study. IMHNN also uses hyper-graphs, which helps to find the high-order interactions between different student behaviors and does not suffer from missing key interactions as conventional approaches do. An attentional convolutional network is integrated to make the various behaviors enjoy focus weights for maintaining the model quality that depicts the significance of different activities. Besides, there is a feature called residual network adopted to maintain the elimination of transition smoothing and enhance the generalization ability of the model. We have also provided the experimental results of a real educational dataset which contributes to verify that the proposed IMHNN yields more better performance compared with the prior methods, which also supplies evidence to support that the research can successfully apply complicated behavioral data and enhanced MHNN structure. Nevertheless, this study has limitations that are worthy of mention. First, architectural complexity of hyper-graph neural networks and attentional convolutional networks may shift the processing demands up and, therefore, hurt accessibility for institutions with fewer computational resources. Furthermore, the derived model performs better; however, the does the delivered output of the executed analysis which may present a problem for the educators who are not knowledgeable in deep learning. This may also pose a limitation to the applicability of the model due to environment and classroom constraints in educational institutions. Finally, the integral use of educational contexts and students' population type restricts the possibility to compare the outcomes and find out other correlated factors. More empirical trials are necessary to determine whether the presented strategy works well in other contexts and whether it is possible to enhance the interpretability of the results and their usability for educational practitioners.

Sun, et al (Sun et al., 2024) provided an analysis of the effects of the various antecedent conditions that are combined to affect primary school students' performance with special emphasis on executive functions, physical fitness, and demographics. Thus, by incorporating these variables, the study provides a multiple perspective on the factors that affect students' performance. In particular, the strength of the work is the application of the methods of the ensemble learning to the prediction of the academic performance. There also exists ensemble learning, in which several separate learners are used to generate a final prediction and the performance difference between a solitary learner and several combined learners is compared carefully. Power of ensemble techniques over the individual ones is also supported while emphasizing the possibility for more accurate and less error-prone prediction of the academic performances. Besides, this comparative procedure is beneficial in increasing the comprehension of the separate and combined impact of the various factors and proving the effectiveness of modern machine learning approaches in the educational field. The findings of this study allow for the discussion of educational approaches and initiatives based on the notion of the diverse and complex nature of organizations that can enhance outcome by targeting all the aspects that are involved in learning performance. Zhang, et al (Zhang et al., 2024) introduced a new type of student dropout prediction model, the Image Convolutional and Bi-directional Temporal Convolutional Network (IC-BTCN), is proposed to solve the problem of high dropout rates in MOOCs. Built on learning clickstream data, the proposed model, the IC-BTCN, employs state-of-art machine learning to find out which of the students are potentially dropping out. The background of the conducted study refers to the attractiveness of the MOOCs, as they convey large amounts of teaching materials and affordable learning approaches to millions of learners globally. The major challenge that have been identified to affect the MOOC's include; However, MOOCs have some specific drawbacks despite the mentioned benefits they possess high dropout rates compared with more traditional forms of offline education. IC-BTCN resolves this problem by incorporating image convolutional form with bi-directional temporal convolutional networks, making the device powerful in analyzing the students' interactions and engagement tendencies. It not only raises the accurate level of dropout prediction, but also brings out some insightful information that would definitely contribute to the establishment of proper measures to increase students' retention rate by educators. Thus, this study raises the possibility of incorporating complex data analysis methodologies into on-line learning systems in a way that addresses dropout and other issues concerning MOOCs' effective implementation.

Luo, et al (Luo et al., 2024) seeks to present a learning outcome evaluation of blended courses offered in a Chinese university with emphasis on utilizing learning data to recommend a new categorization of these courses. Through expectation–maximization algorithm, the students were properly grouped by their behaviors in online learning, thus, different blended courses were grouped under the dominant behavior that the students displayed. This method offers a protocols-based unambiguous criterion for defining blended courses hence it is easy for machine learning systems to employ log data from Learning Management Systems for classification. Apart from improving the knowledge about students' behaviors in the context of blended learning, this study also presents the new and more interpretable way to perform the classification of students' behaviors and make them applicable to the machine learning approaches. Further, the researchers created and verified a general predictive model through the random forest algorithm for estimating learning performance from the students' online activities. On using this model, improvements were realized across-discipline and amongst various strata of students, further increasing the model's generalizability. The results of this research show that, in the future, by using more detailed information about users' activity on the Internet, it is possible to enhance the accuracy of positioning subjects in courses that would indicate a higher academic performance, which is worth considering by educators and policymakers who are interested in customizing web-based instruction.

Kukkar, et al (Kukkar et al., 2023) presented a new Student Academic Performance Predicting (SAPP) system which the researcher proposes to improve on the existing systems for accuracy. The SAPP system also has a five-tier architecture that includes stacked LSTM network with RF and GB methods in a four-layer structure. This hybrid approach aims to improve the precision of predicting students' pass or fail outcomes by leveraging the strengths of each method: LSTM being suitable for use in sequential data, RF for being effective in classification problems, and GB for increasing the prediction accuracy through boosting. Thus, the performance of the proposed SAPP system is compared with the existing prediction models using public Open University Learning Analytics Dataset (OULAD) along with an additional self-compiled emotional dataset. An addition of emotional data focuses on a different level which might reach the psychological and emotional factors that cause or affect performance. The comparison proves that the application of SAPP system exceeds the practices of the traditional approach, which suggests that it has the possibility to generate more accurate and exceptional forecast than other models. As such, the findings of this research highlight the need for multi-authored data integration as well as the applicability of modern machine learning approaches for student outcomes modeling, which can benefit educational organizations interested in enhancing their students' success rates (Joshi, Puthuparampil, Kini, & Baisil, 2025).

Asselman, et al (Asselman et al., 2023) introduced a Performance Forecasting Algorithm (PFA) which combine Random Forest, AdaBoost, and XGBoost algorithms. When comparing the result of experimentation with three different datasets, it is thus found that the scalable XGBoost model has a better performance and offers a much higher prediction accuracy as compared to the PFA algorithm or any other models like KNN, DT, or SVM. This research underlines the use of ensemble methods in E-learning EDM specifically, to emphasize the fact that such approaches can mark great possibilities of yielding better and more reliable prognosis on the performance level of students. There exist several research gaps that have not been covered even with the current developments in technological approaches such as the application of machine learning and deep learning in analyzing students' performance. Current models' drawbacks include complexity and interpretability, which may entail difficulties in using the predictions made by the models for educators and other stakeholders. Some of these models, such as deep learning architectures need a large amount of labelled data and a lot of computational power which is, at times, hard to come by in educational environments. In addition, the challenge of overfitting is still there; that is, while the models perform very well on the training data set, they do not perform well on new data. Moreover, these models may end up developing specific biases in the training dataset that may be prejudicial when used to predict the efficiency of students (Gopalan, 2025). Analyzing the factors affecting the students' performance, there is no information on how various type of data, including emotional and behavior data, could enhance the general understanding of the factors that affect the performance of the students. Most of existing models fail to consider the time varying student data and distribution changes across period, which in addition to one or more features can cause more deterioration of the accuracy of prediction (Yuetqi, Lin, Mamat, & Noorani, 2025).

The rationale for the proposed student performance prediction system is to overcome these limitations and improve the prediction's reliability, equity, and explainability. In turn, the proposed system strives to achieve the added values of each of the employed models like Random Forest, AdaBoost and XGBoost to have the more accurate and reliable predictions. To complement the conventional academic data, features from a manually selected set of emotional data along with the psychological state of the students is included, which is hoped to offer a wholesome scenario to encompass the facets that affect the performance of the students.

Proposed Methodology

This section provides a detailed discussion of the proposed AcadPredictor model, which is indeed a highly elaborate model that meets the purpose of accurately predicting students' performance with much efficiency. Embedded at its center is the Hyb-RARNet (Regression Auto-Encoding Residual Network) a deep learning model developed to identify subtle patterns and correlations from the data, and more importantly make accurate forecast on student academic performance. The proposed Hyb-RARNet model integrates regression, auto-encoding, and residual architectures for an extensive predictive engine. Regression methods are used to analyze the dependence of a target vector on input variables and allow for estimating the student's grades taking into account certain academic, demographic, and behavioral characteristics. Auto-encoding processes expand on the model's capability to learn hidden features of input information and identify patterns that might not be observable from regression models. On the other hand, residual links allow gradients to flow smoothly from the output deeper to the input layers while the vanishing gradient problem is solved effectively for the training of deep learning neural networks to be efficient. Further, besides the strong predictive model of the Hyb-RARNet, the Honey Levy Explorer Optimization (HLEO), a new optimization model is also employed in this study. This technique is derived from honey badger foraging behavior and Stochastic search Levy Flight property and thus provides a different way of performing parameter tuning as well as optimization. HLEO works by successive search for the best estimate of the parameters in the search space with deterministic and stochastic methods. First, the algorithm perform the exhaustive search of all previously defined areas of interest by using the deterministic optimization methods in order to observe the more promising zones for additional exploration. Fig 1 shows the overview of the proposed AcadPredictor model.

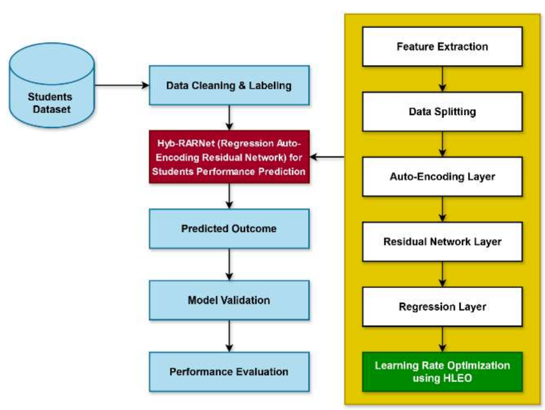

Figure 1. Flow of proposed AcadPredictor

Next, stochastic search tactics mimicking the Levy Flight behavior are used to move in the parameter space non-linearly and thus search for various areas and possible solutions. AcadPredictor plans to make full utilize of the Hyb-RARNet model and HLEO optimization to provide an accuracy and robust model for future student performance prediction. The combination of such an intricate deep learning algorithm along with newer and more efficient optimization methods enables the AcadPredictor to diagnose intricate patterns within the data sets and provides pertinent recommendations to educators and the stakeholders of the academic sphere. With the use of systematic testing and proving, AcadPredictor aims to innovate the previously explored way of predicting learners' performance and thus, become a solution for data-supported decision making and improving educational results for learners.

The proposed work can be considered as a significant improvement concerning the existing approaches to student performance prediction, as well as several new additions that differentiate it from the prior research. Essentially, the inclusion of the Hyb-RARNet deliberates an innovative strategy in the application of a prediction model. This work overcomes some of the restrictions of a traditional regression model, while applying various structures such as regression, auto-encoding, and residuals from the ResNet and thus, gives a profounder interpretation of the patterns lying in the data of students. Such an approach to modeling helps Hyb-RARNet capture complex interactions and dependencies between various academic, demographic, and behavioral predictors, which in turn fosters improved performance predictions' validity and stability.

In addition, compared with the conventional modeling and optimization approach in data science, the proposed framework includes a wide range of aspects of students' data information. In a way that simple algorithms that primarily focus on a single academic record, attendance prediction generator offers a more subtle and complex look at students' performance based on many factors such as attendance records, extracurricular activities, demographic information, and much more data that the AcadPredictor uses. Not only does this multi-dimensional approach improves the high degree of precision of the predictions but also greatly improves the understanding of factors that underlie students' performance, which in turn increases the chance that educators and other stakeholders will take actions relevant to improving student achievement. Overall, the proposed work contains the great potential to reinvigorate the student performance prediction analysis by incorporating diverse research starting from the deep learning approaches to the optimization techniques and including new efficient data analysis methodologies. Since, the AcadPredictor can inform action in decision-making processes and contribute to the enhancement of educational results, it is one of the first aforesaid approaches to support student success in an academic context more efficiently through data analysis.

3.1 Hyb-RARNet (Regression Auto-Encoding Residual Network)

The Hyb-RARNet (Regression Auto-Encoding Residual Network) is the new generation of deep learning structure used in predictive modeling that aims to precise various jobs including the one outlined in the proposal such as the student performance prediction. This innovative framework seamlessly integrates three distinct components: are regression, auto-encoding, and residual networks, it makes use of the advantages of each to get better performance and stability at prediction. The network starts with an input layer in order to transmit feature vectors that define different inputs like, academic records, demographics, behavior, etc. The regression component of Hyb-RARNet is made up of a few fully connected layers that are responsible for the regression of the obtained input data. These layers perform the task of converting input features to continuous output like the student's grade or performance scores. At the same time, the autoencoder part contains the encoder and decoder layers, which try to extract the main features of the data. The encoder thus reduces the features of the input to a dimensionality embodying all the informative characteristics while the decoder strives to create a replica of the input after mapping through the latent space.

Connections between selected layers are retained with the aim of helping gradients to flow through all the layers and overcome the vanishing gradient issue. These connections enable the network to learn residual mappings and thus, restrict the model's capability of learning intricate mappings in the data. Pressing at last results in an output layer that delivers the last estimations involving the use of both the regression and auto-encoding parts of the network. The regression component of Hyb-RARNet is also aimed at learning a continuous function to predict continuous variables, like students' grade or performance scores. This component contains fully connected layers that make use of techniques of regression analysis to capture dependencies between input and output variables of a given problem. Consequently, the auto-encoding part works in parallel to the generation and tries to capture a reliable generalized representation of the given input data. This is the basic concept of auto-encoders and is accomplished through the encoder-decoder model where the features of the input are compressed into the latent space and yet they preserve relevant information.

After data collection, the input dataset having n number of features is represented in the form of $s_1, s_2 \ldots s_n$. As a consequence, the data normalization is carried out for generating the preprocessed data based on the mean and standard deviation value as represented in below:

$$N_D = \frac{S-\mu}{\sigma} \tag{1}$$

Where, N_D indicates the normalized data, S is the input students data, μ and σ are the mean and standard deviation values. Then, the normalized students' data can be used for prediction, where the regression layer operation is performed with the use of ReLU activation function for feature represented as shown in below:

$$x_{\mathbb{R}} = \mathrm{ReLU}(\mathfrak{w}_{\mathbb{R}} \times N_D + \mathfrak{B}_{\mathbb{R}}) \tag{2}$$

Where, $x_{\mathbb{R}}$ indicates the regression feature representation, $\mathfrak{w}_{\mathbb{R}}$ is the weight value and $\mathfrak{B}_{\mathbb{R}}$ denotes the bias value. The output of this layer is the predicted regression result, which is estimated with the learning weight $\mathfrak{w}_{\sigma}$ and bias values $\mathfrak{B}_{\sigma}$ as represented in below:

$$\hat{\sigma}_{\mathbb{R}} = \mathfrak{w}_{\sigma} \times x_{\mathbb{R}} + \mathfrak{B}_{\sigma} \tag{3}$$

Where, $\hat{\sigma}_{\mathbb{R}}$ indicates the regression result. Consequently, the auto-encoding layer operation is performed, which comprises both encoding and decoding operation. At first, the encoding layer is used to encode the normalized students' data into the latent form based on ReLU activation function. This operation is mathematically represented as shown in below:

$$x_{\mathcal{E}} = \mathrm{ReLU}(\mathfrak{w}_{\mathcal{E}} \times N_D + \mathfrak{B}_{\mathcal{E}}) \tag{4}$$

$$\mathfrak{w}_{\mathcal{E}}^{k+1} = \mathfrak{w}_{\mathcal{E}}^{k} - \mathfrak{X} \times \frac{\partial \mathbb{L}}{\partial \mathfrak{w}_{\mathcal{E}}} \tag{5}$$

$$\mathfrak{B}_{\mathcal{E}}^{k+1} = \mathfrak{B}_{\mathcal{E}}^{k} - \mathfrak{X} \times \frac{\partial \mathbb{L}}{\partial \mathfrak{w}_{\mathcal{E}}} \tag{6}$$

Then, the decoding layer operation is performed with encoded representation according to the sigmoid activation function as shown in below:

$$S' = \varphi(\mathfrak{w}_{\mathfrak{D}} \times N_D + \mathfrak{B}_{\mathfrak{D}}) \tag{7}$$

$$\mathfrak{w}_{\mathfrak{D}}^{k+1} = \mathfrak{w}_{\mathfrak{D}}^{k} - \mathfrak{X} \times \frac{\partial \mathbb{L}}{\partial \mathfrak{w}_{\mathfrak{D}}} \tag{8}$$

$$\mathfrak{B}_{\mathfrak{D}}^{k+1} = \mathfrak{B}_{\mathfrak{D}}^{k} - \mathfrak{X} \times \frac{\partial \mathbb{L}}{\partial \mathfrak{w}_{\mathfrak{D}}} \tag{9}$$

Moreover, the residual connection enhance the information flow in the network, where the residual mapping function is used to obtain the residual information as shown in the following equation:

$$x_{\mathbb{R}} = N_D + \mathrm{ReLU}(\mathfrak{w}_{\mathbb{R}} \times N_D + \mathfrak{B}_{\mathbb{R}}) \tag{10}$$

$$\mathfrak{w}_{\mathbb{R}}^{k+1} = \mathfrak{w}_{\mathbb{R}}^{k} - \mathfrak{X} \times \frac{\partial \mathbb{L}}{\partial \mathfrak{w}_{\mathbb{R}}} \tag{11}$$

$$\mathfrak{B}_{\mathbb{R}}^{k+1} = \mathfrak{B}_{\mathbb{R}}^{k} - \mathfrak{X} \times \frac{\partial \mathbb{L}}{\partial \mathfrak{w}_{\mathbb{R}}} \tag{12}$$

At the end of process, the total loss function is estimated by integrating the regression and auto-encoding models with a hyper-parameter $\eth$, as mathematically computed in below:

$$\mathbb{L} = \eth\mathbb{L}_{\mathbb{R}} + (1 - \eth)\mathbb{L}_{AE} \quad (13)$$

Where, $\mathbb{L}$ indicates the total loss function, $\mathbb{L}_{\mathbb{R}}$ is the loss function of regression model, and $\mathbb{L}_{AE}$ denotes the loss function auto-encoding model.

3.2 Honey Levy Explorer Optimization (HLEO)

The HLEO technique is a novel optimization strategy policy that aims at improving the performance of the Hyb-RARNet model by estimating the values of the learning rate during the training phase. This method combines and harmonizes the exploration feature of the Honey Badger Algorithm with the efficient search function of Levy Flight, and well applied it on how to fine-tune the learning rate for the deep learning models. Hence, the main advantage of HLEO over other optimization approaches is in its originality of melding the efficiency of two great algorithms. Unlike many other algorithms, the Honey Badger Algorithm is widely regarded as possessing exceptional exploration characteristics, owing to the honey badger's proactive searching patterns. Mitigating the weaknesses of the global and local searches, it optimally executes the search in the solution space. On the other hand, Levy Flight, which is a random walk pattern having drawn out step sizes, improves the search since they unlock the algorithm from shallow optima making it to search a wider area of the solution space.

With this HLEO successfully maintains a reasonable balance between greediness and exploration that is so important for learning rate in a high-dimensional space, which is typical for deep learning models. Another factor that makes the choice of this type of learning rate optimal is the ability to adjust the learning rate online during the running of the program to ensure that it converges faster, does not trap in local minima and can learn as much as it can from the data set that has been supplied to it. Besides, this form of adaptive approaches not only optimizes the training of Hyb-RARNet but also the performance of the model when prediction accuracy is concerned due to its flexibility in determining the most desirable learning rate. The proposed HLEO model that integrates Honey Badger Algorithm and Levy Flight to find the best learning rate makes HLEO a unique and valuable contribution to the development of deep learning for the prediction of students' performance in the field of education.

HBA is an optimization technique designed based on the features of the honey badger, which is a courageous creature in the process of searching for food as it does not give up until it is successful. The dynamic search behavior of honey badgers is characterized by two main approaches: What's more, it has the digging mode and the honey mode, where the former refers to the exploration mode in the algorithm and the latter is for the exploitation mode. With regard to foraging for food, the honey badger uses two strategies that can be considered different behaviors. In digging mode, the badger uses its sense of smell to dig insisting on finding its prey including mice that may be hidden in the surface of the ground. It then digs continually to get to the food source once it identified it. This technique falls under the exploration phase of the HBA in that the algorithm goes out to search space to look for potential optimal solutions on its own. In this phase, we have the honey badger algorithm searching the new solution space as the honey badger would search for food by digging the ground at different places. This makes the search wider and more diverse, excludes the program lock-in to local optima before reaching the global ones and increases the efficiency of finding them.

The second of the tactics is quite opposite to the first one and it is called the honey mode. In this mode, it accompanies the honey guide bird which can point to a beehive but cannot get to the honey. Since the honey badger is a follower of the bird, the kill will be lead directly to the honey badger. It is the exploitation phase where the HBA algorithm mainly works on feathering and local optimization of the obtained results in order to find the optimum solution. The dynamic search modes of the HBA are further enhanced by the durability, versatility, and courage of the animal, which is the honey badger. Despite their quite small size, honey badgers will face much larger predators and fight with them if necessary, and they are also capable climbers to get at birds nests and beehives. This is typical in the environment, and is clearly demonstrated in the flexibility and boldness that the algorithm displays when it can move from exploration to exploitation while holding a balanced approach that is fundamental in achievement of the optimization. This makes the HBA flexible in responding to the types of optimization problems as well as appropriately taking exploitative steps to fine tune the solutions discovered by a more exploitative step and the exploration steps to look for as many different solutions as can be found.

In the proposed model, the number of honey badgers are initialized with respect to the population size and respective positions as represented in below:

$$p_i = \ell b_i + \delta \times (ub_i - \ell b_i) \qquad (14)$$

Where, p_i indicates the position of honey badger, ℓb_i denotes the lower bound value, ub_i represents the upper bound value, and δ denotes the random number. As a consequence of this, the intensity is estimated according to the concentration and distance measures as shown in below:

$$\mathcal{I}_i = \mathfrak{z} \times \frac{\mathfrak{F}}{4\pi d_i^2} \qquad (15)$$

$$\mathfrak{F} = (p_i - p_{i+1})^2 \qquad (16)$$

$$d_i = p_{prey} - p_i \qquad (17)$$

Where, $\mathcal{I}_i$ indicates intensity, $\mathfrak{z}$ represents the random number, $\mathfrak{F}$ is the source strength, and d_i denotes distance. Then, the density factor that is used to control the time variable randomness is estimated according to the Levy flight model as mathematically represented in the following equation:

$$k = \frac{\xi\alpha}{|\gamma|^{\frac{1}{\varpi}}} \times \frac{Mx_{itr} - it^{\tau}}{Mx_{itr}} \qquad (18)$$

Where, α, γ are the random values, ξ indicates the scaling factor, ϖ denotes the Levy component, τ is the controlling parameter, Mx_{itr} is the maximum number of iterations, and it is the current iteration. In order to avoid from local optimal, the flag value is generated for adjusting the searching direction in the region as shown in the following equation:

$$f = \begin{cases} 1 & \text{if } \mathfrak{A} \leq 0.5 \\ -1 & \text{Otherwise} \end{cases} \qquad (19)$$

Where, $\mathfrak{A}$ is the random parameter. Finally, the position of all agents are updated according to the following equation:

$$p_{new} = p_{prey} + f \times \zeta \times k \times d_i \qquad (20)$$

Where, p_{new} indicates the newly updated solution, and ζ denotes the random number. By adopting this optimization technique, the learning rate is optimally computed for an effective students' performance prediction

Results and Discussion

This section provides a clear justification of the performance measures and the findings of the developed AcadPredictor model. This also involves qualitative analysis that covers descriptions of the datasets used and counts of the measures that were used to assess the model. In doing so, the discussion seeks to draw out and exaggerate the performance strengths of the AcadPredictor model while pinpointing areas for improvement in comparison with using current models' predictive accuracy and other performance parameters. Hence, through the analysis of the different levels of the model based on student lectures, this section will show the flexibility and generalizability of the model in predicting student academic performance. As shown in Table 1, the dataset in this particular study covers different variables that describe students' demographic and academic characteristics, as well as some behavioral characteristics necessary for modeling students' academic performance. The AGE feature is numeric, and it has values from 1 to 10 and corresponds to the students' age. The GENDER feature is dichotomous, as it will be either 'male' or 'female' pertaining to the student. HS_TYPE is binary feature where It indicates that the student has attended a particular sort of high school 'Yes' or 'No'. In the same manner, the SCHOLARSHIP feature is binary where 1 represents that the student receives a scholarship. Basic education levels of the fathers and mothers are represented by FATHER_EDU and MOTHER_EDU respectively which are categorical variables with High, Low, and Average. The SIBLINGS feature is nominal, that is whether the number of siblings is more than 5, equal to 5 or less than 5. Both the children indicator (KIDS) and the parents' employment status (MOTHER_JOB and FATHER_JOB) are categorical observations with two values: Yes/No. About education background is assessed by PREV_STUDY which is categorical variable with answers High, Low and Average about previous study, and PREV_EXAM which is also categorical variable containing the same answers about previous exam performance. Other variables are also binary; whether a student takes notes (NOTES), listens (LISTENS), and enjoys discussion (LIKES_DISCUSS). CLASSROOM reflects performance of students in the classroom and expectations of the teachers, CUML_GPA is cumulative grade point average and EXP_GPA is the expected grade point average, all are categorical variables with High, Low and Average. The feature of COURSE_ID is numeric as different courses correspond to different course ids. At last, the GRADE feature is categorized whereby a student's performance is as passed as low, good, average or drop. We explained that this set of student records is quite vast and forms a strong base on which we can apply and compare different machine learning models to forecast student academic performance.

Table 1. Kaggle students' dataset description

Feature type	Description
Age	Numerical value
Gender	0 or 1 (Male or Female)
Hs_Type	Yes or No
Scholarship	Yes or No
Father_Edu	Categorical value (low, average or high)
Mother_Edu	Categorical value (low, average or high)
Sibilings	Categorical
Mother_Job	Yes ot No
Father_Job	Yes or No
Prev_Study	Categorical value (low, average or high)
Prev_Exam	Categorical value (low, average or high)
Notes	Yes or No
Listens	Yes or No
Likes_Discuss	Yes or No
Classroom	Categorical value (low, average or high)

Cuml_GPA	Categorical value (low, average or high)
Exp_GPA	Categorical value (low, average or high)
Course_ID	Numerical value
Grade	Categorical value (low, average or high)

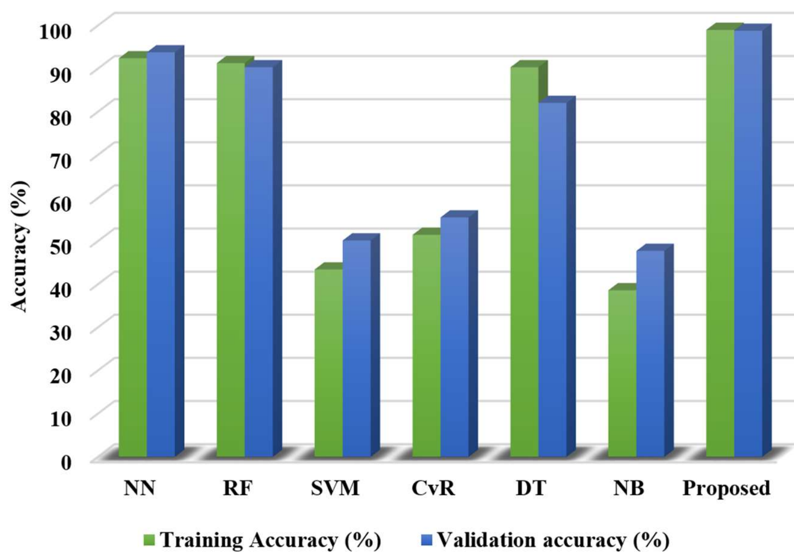

Figure 2. Comparison among the conventional machine learning and proposed AcadPredictor models using Kaggle dataset

Fig 2 and Table 1 illustrate the comparison between the previous machine learning models and the AcadPredictor model by using a Kaggle dataset. The table particularly describes the percentage of the training and the validation accuracy of several models for easy comparison of the performance of the models. The piece of research focuses on the use of Neural Network (NN) model in analyzing data, and the model has a high training accuracy of 92%. It shows that the model is not very sensitive to the variations of the dataset and thereby, exhibits good generalization capability. The Random Forest (RF) model also gives a good performance having training accuracy of 91%. 3% with valuation accuracy of 90%. 34% and this is confirmation of its efficiency in the handling of large data set. A further comparison of the findings of the present study with those of the literature on similar topics revealed that the performance of the proposed NN, CNN, and DA models is very high than that of the Support Vector Machine (SVM), and Naive Bayes (NB) model, with 43% training accuracy only. 44% and 38. 62% respectively, and validation accuracy of 50. 17% and 47. 75%. The validity measures come out to be 75%, where the identified measures are somewhat useful, but not very powerful. Interestingly, the moderate training accuracy of 51% for the CvR and 49% for the DT models are obtained. 48% and 90. 34%, and validation accuracy of 55. 51% and 82. 06%, respectively. On the other hand the proposed AcadPredictor model performs extremely well with the training accuracy of 99% and validation accuracy of 98%. 8%. The AcadPredictor model linked in this study is much more accurate in its predictive ability than other models for similar calculations, and indeed, it employs new and innovate methods for prediction. The performance of the proposed AcadPredictor model is equally impressive, given that the model offers accurate results of students' performance, compared with the other machine learning methods. This better performance has more to do with its structure and with the use of as many and different data feeds as possible, making prediction more accurate.

Table 1. Comparison based on training and validation accuracy

Methods	Training Accuracy (%)	Validation accuracy (%)
NN	92.41	93.79
RF	91.3	90.34
SVM	43.44	50.17
CvR	51.48	55.51
DT	90.34	82.06
NB	38.62	47.75
Proposed	99	98.8

Table 2. Comparison based on prediction accuracy

Methods	Prediction accuracy (%)
LSTM + Embedding	77
LSTM + Attention	76
Fused LSTM	75
Aspect Attention + GRU	84
Layered LSTM	85
LSTM + AE	89
Proposed model	99

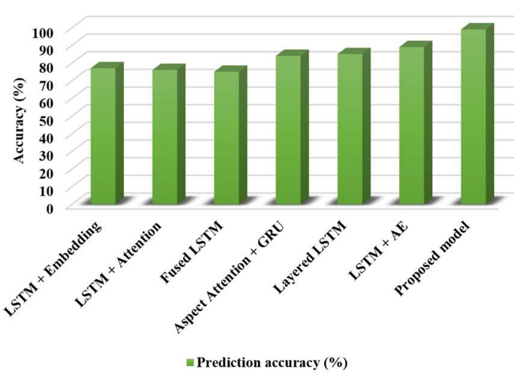

Figure 3. Comparative analysis with hybrid models based on prediction accuracy

The accuracy comparison of different models is discussed in Table 2 and Fig 3 that indicates the effectiveness of the proposed model. The table captures the performance of several other complex advanced machine learning and deep learning models thus providing a good reference point of performance when it comes to predicting the student outcomes. With the LSTM (Long Short-Term Memory) based models, the best one is LSTM + Embedding model that estimates the prediction accuracy of 77%, the LSTM + Attention achieves the accuracy of 76% only. Both of these models capitalize on LSTM's attribute as a good algorithm when dealing with sequential data, however, their results suggest that there is potential for enhancement. The accuracy of combined multiple LSTM is 75 % which is slightly low than the previous one and hence proves that just connecting multiple LSTM do not improve the efficacy of prediction. On the other hand, the models which adopt attention mechanism and different architectural structure perform significantly richer. Aspect Attention + GRU model has a prediction rate of 84% showing the importance of focus aspect to improve the performance of the model. If multiple

LSTM layers, to learn deeper temporal dependencies, are applied then the accuracy shoots up to 85%. In the case of LSTM with the additional Auto-encoder feature, a barely respectable accuracy of 89% is achieved.

This kind of performance which is exceptional could be attributed to the fact that the model used has a better architecture than ordinary models in that it is capable of recognizing complex patterns within the data and is able to combine multiple types of predictions in a bid to improve on the accuracy aspect. This paper has provided empirical evidence in support of the proposed model in as much as students' academic performance prediction is concerned while at the same time revealing how it can be practically implemented in the educational environment due to high prediction accuracy. The utilization of reliable predictions to populate the suggestions would mean that the interventions would be more efficient which again implies high success rates among the students and low dropout rates. As shown in this extensive comparison, it is imperative to incorporate the strategies of the most recent machine learning methods while creating better and more reliable forecasting models in educational data mining.

Table 3. Comparison with hybrid deep learning models

Methods	Precision (%)	Recall (%)	F1-score (%)
LSTM + Embedding	90	62	73
LSTM + Attention	82	85	83
Fused LSTM	89	83	85
Aspect Attention + GRU	93	90	91
Layered LSTM	87	88	87
LSTM + AE	83	86	87
Proposed model	98.8	98.9	99

Table 3 and Fig 4 show the comparison of different hybrid deep learning models with all the associated metrics of precision, recall, and F1-score, which shows that the proposed model has the highest score for each of the metrics. It is essential to note that these metrics are almost mandatory for evaluating the effectiveness of the developed predictive models, especially for the educational data mining where the needed accuracy represents a balanced criterion in the predictive models for making correct prognosis. Another remarkable model is LSTM + Embedding with the precision of 90% that proves the effectiveness of the algorithm in terms of identifying positive samples. However, its recall value is 62%, which means it actually fails to identify quite a number of actual 'positive' instances, and its F1-score stands at 73% which is a little better than the binary accuracy score. This compares its ability of precise search and comprehensive search and shows that it does not perform well in both fields.

Specifically, the LSTM + Attention model conceives comparatively average metrics with 82% of precision, 85% of recall, and an F1-score of 83%. It assists it converge on the places that contain the most information about the input data and thus increases its ability to make predictions. The Fused LSTM model which is essentially a compilation of several LSTM networks were slightly better with a precision of 89%, recall 83 % and an F1-score of 85%. Compared to LSTM, this model also has the advantage of aggregating multiple outputs for better recall and, at least at first glance, minimal loss of precision. Aspect Attention and GRU presents one of the finest conventional work correlation examination efficiency with the precision of 93%, the recall of 90%, and F1-score of 91%. The overall accuracy is high, and this model combines aspect-level attention with GRUs to capture features well, which is proven very useful. The Layered LSTM model, and the LSTM + AE (Autoencoder) model are also quite effective with f1-score of 0.87 each. These models benefit from deeper layers as well as auto-encoding methods in capturing data patterns of higher complexity as well as eliminating noise.

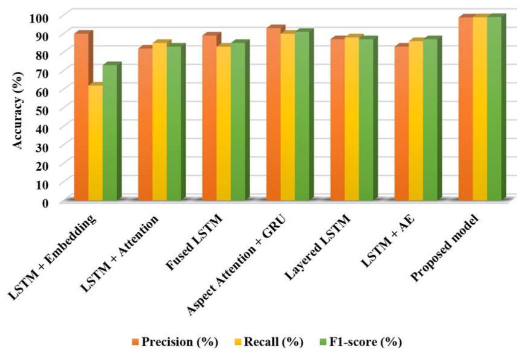

Figure 4. Comparison based on precision, recall and f1-score with hybrid deep learning models.

Conclusion

The proposed work is mainly focused on the important issue of predicting student academic achievement with advanced deep learning and optimization models in our study. We show that this model had greater generalizability as the estimator performed significantly better than a host of traditional models in every task with real-world educational data. The proposed framework uses the Hyb-RARNet model by merging regression, auto-encoding and residual network techniques to capture fine-grained data patterns that translate into better prediction capability. We applied the HLEO algorithm, inspired by natural foraging and search behaviors, and Levy flight patterns to optimize parameters in our predictive models and therefore improve their efficiency and accuracy. We empirically evaluated our models by measuring precision, recall, f1-score and prediction accuracy on different datasets. This paper presents a milestone in the field of educational data mining by providing a completely new, effective solution for student academic performance prediction. Such models have very good potential to improve educational outcomes and reduce dropout rates and thus support overall goals of educational quality and student success. Further research can be done by integration of data from sources like social media interactions, information about extracurricular activities, or psychological measurements. These additional data might put a student's academic environment into a holistic perspective and hence further improve predictions.

6. Author Declarations

Conflict of Interest

The authors declare that they have no conflict of interest.

Competing Interests

The authors have no competing interests to declare relevant to this article's content.

Funding Details

No funding was received to assist with the preparation of this manuscript.

Data Availability

Data sharing not applicable to this article as no datasets were generated or analyzed during the current Study.

References

Akour, M., Alsghaier, H., & Al Qasem, O. (2020). The effectiveness of using deep learning algorithms in predicting students achievements. Indonesian Journal of Electrical Engineering and Computer Science, 19, 387–393.

Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student performance prediction using machine learning techniques. Education Sciences, 11, Article 552.

Anjireddy, D., & Jaisharma, K. (2024). Performance comparison of flood prediction using recurrent time delay neural network and K-nearest neighbor algorithm. In 2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (pp. 1–5).

Arifin, A., Suryaningsih, S., & Arifudin, O. (2024). The relationship between classroom environment, teacher professional development, and student academic performance in secondary education. International Education Trend Issues, 2, 151–159.

Asselman, A., Khaldi, M., & Aammou, S. (2023). Enhancing the prediction of student performance based on the machine learning XGBoost algorithm. Interactive Learning Environments, 31, 3360–3379.

Boddu, Y., & Manimaran, A. (2025). Design of an iterative method for time series forecasting using temporal attention and hybrid deep learning architectures. IEEE Access.

Chen, W., Shen, Z., Pan, Y., Tan, K., & Wang, C. (2024). Applying machine learning algorithm to optimize personalized education recommendation system. Journal of Theory and Practice of Engineering Science, 4, 101–108.

Deepa, N. (2025). Accurate multi language interpretation word imperative prediction in cross linguistic using novel modal and minimal rule-based technique in comparison with vector space approach. In 2025 Tenth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (pp. 1–6).

Dien, T. T., Luu, S. H., Thanh-Hai, N., & Thai-Nghe, N. (2020). Deep learning with data transformation and factor analysis for student performance prediction. International Journal of Advanced Computer Science and Applications, 11.

Faruque, S. H., Khushbu, S. A., & Akter, S. (2024). Unlocking futures: A natural language driven career prediction system for computer science and software engineering students. arXiv. https://arxiv.org/abs/2405.18139

Fazil, A. W., Hakimi, M., Shahidzay, A. K., & Hasas, A. (2024). Exploring the broad impact of AI technologies on student engagement and academic performance in university settings in Afghanistan. RIGGS: Journal of Artificial Intelligence and Digital Business, 2, 56–63.

Fazil, M., Rísquez, A., & Halpin, C. (2024). A novel deep learning model for student performance prediction using engagement data. Journal of Learning Analytics, 1–19.

Forero-Corba, W., & Bennasar, F. N. (2024). Techniques and applications of machine learning and artificial intelligence in education: A systematic review. RIED-Revista Iberoamericana de Educación a Distancia, 27.

Hachani, A., Mallek, M., & Jmal, Y. (2024). Enhancing academic success: Performance early prediction using machine learning algorithms. EasyChair Preprint, 2516–2314.

Kocsis, Á., & Molnár, G. (2024). Factors influencing academic performance and dropout rates in higher education. Oxford Review of Education, 1–19.

Kukkar, A., Mohana, R., Sharma, A., & Nayyar, A. (2023). Prediction of student academic performance based on their emotional wellbeing and interaction on various e-learning platforms. Education and Information Technologies, 28, 9655–9684.

Lin, X., & Guo, Y. (2024). Student's performance prediction based on an improved multi-view hypergraph neural network. Journal of Experimental & Theoretical Artificial Intelligence, 1–17.

Luo, X., Wang, W., Xu, Y., Lai, Z., Jin, X., Zhang, B., et al. (2024). A deep convolutional neural network for diabetic retinopathy detection via mining local and long-range dependence. CAAI Transactions on Intelligence Technology, 9, 153–166.

Luo, Y., Han, X., & Zhang, C. (2024). Prediction of learning outcomes with a machine learning algorithm based on online learning behavior data in blended courses. Asia Pacific Education Review, 25, 267–285.

Namoun, A., & Alshanqiti, A. (2020). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. Applied Sciences, 11, Article 237.

Nanavaty, S., & Khuteta, A. (2024). A deep learning dive into online learning: Predicting student success with interaction-based neural networks. International Journal of Intelligent Systems and Applications in Engineering, 12, 102–107.

Ren, Y., & Yu, X. (2024). Long-term student performance prediction using learning ability self-adaptive algorithm. Complex & Intelligent Systems, 1–30.

Sengupta, S., Das, R., & Chakrabarti, S. (2025). Hybrid data mining techniques for predicting student academic performance in e-learning to avoid drop-out (HDL-SP): An efficient data mining technique to forecast academic performances of students. SN Computer Science, 6, Article 162.

Sheela, M. S., Amirthayogam, G., Hephzipah, J. J., Gopalakrishnan, S., & Chand, S. R. (2024). Machine learning based lung disease prediction using convolutional neural network algorithm. Mesopotamian Journal of Artificial Intelligence in Healthcare, 50–58.

Sheela, M. S., Gopalakrishnan, S., Begum, I. P., Hephzipah, J. J., Gopianand, M., & Harika, D. (2024). Enhancing energy efficiency with smart building energy management system using machine learning and IoT. Babylonian Journal of Machine Learning, 80–88.

Silambarasan, E., Naresh, E., Asha, V., & Lamani, M. R. (2024). Energy efficient heterogeneous clustering scheme using improved golden eagle optimization algorithm for WSN-based IoT. International Journal of Information Technology.

Stasolla, F., Zullo, A., Maniglio, R., Passaro, A., Di Gioia, M., Curcio, E., et al. (2025). Deep learning and reinforcement learning for assessing and enhancing academic performance in university students: A scoping review. AI, 6, Article 40.

Sun, Z., Yuan, Y., Xiong, X., Meng, S., Shi, Y., & Chen, A. (2024). Predicting academic achievement from the collaborative influences of executive function, physical fitness, and demographic factors among primary school students in China: Ensemble learning methods. BMC Public Health, 24, Article 274.

Vellela, S. S., Balamanigandan, R., & Praveen, S. P. (2022). Strategic survey on security and privacy methods of cloud computing environment. Journal of Next Generation Technology, 2.

Waheed, H., Hassan, S.-U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. Computers in Human Behavior, 104, Article 106189.

Wanda, P., & Diqi, M. (2024). DeepNews: Enhancing fake news detection using generative round network (GRN). International Journal of Information Technology.

Zhang, X., Wang, X., Zhao, J., Zhang, B., & Zhang, F. (2024). IC-BTCN: A deep learning model for dropout prediction of MOOCs students. IEEE Transactions on Education.

Gopalan, A. (2025). Compact Bi-slot Patch Antenna with Tapered Edges for Ka-Band Applications Featuring Machine Learning-Assisted Performance Prediction. International Journal of Environment, Engineering and Education, 7(3), 196-207.

Joshi, U., Puthuparampil, R. R., Kini, S., & Baisil, S. (2025). Bridging silos on a budget: how interprofessional education shapes collaborative attitudes across low-and middle-income countries-a systematic review. BMC Medical Education, 25(1), 1602.

Yuetqi, L., Lin, G. S. S., Mamat, R. B., & Noorani, T. Y. (2025). Student's perception of hybrid team-based and case-based learning in undergraduate endodontics education. BMC Oral Health, 25(1), 43...