International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Novel Chaotic Optimized Boost Long Short-Term Memory (COB-LSTM) Model for Students Academic Performance Prediction in Educational Sectors

K. Sankara Narayanan[1], A. Kumaravel[2]

Submitted: 18/07/2023 **Revised**: 07/09/2023 **Accepted**: 24/09/2023

Abstract: Due to the vast amount of data in educational databases, predicting student performance is a significant and complex task. Educational Data Mining (EDM) is used to do this, which creates techniques for locating data produced from educational settings. Understanding students' performance and their learning environment is accomplished using these techniques. Educational institutions sometimes wonder to examine that how many students will pass or fail in order to make the appropriate plans. It has been noted that in earlier studies many researchers focused on choosing the best algorithm for proper classification and neglect looking for solutions to issues that arise during the data mining and analysis phases, including high dimensionality, unbalanced classes, and sorting error. Therefore, the proposed work aims to develop a novel framework, referred to as, Chaotic Optimized Boost Long Short Term Memory (COB-LSTM) for students' academic performance prediction and classification. By using the public UCI education training dataset, a novel method for forecasting and analyzing students' educational achievement is developed in this work. After obtaining a dataset, the Chaotic Henry Gas Solubility (CHGS) optimization technique is used to choose the best traits for a classification that is accurate and low in data dimensionality. Then, a classification method based on the hybrid Boost integrated Deep Long Short Term Memory (Bo-LSTM) is used to accurately predict student performance. Several metrics and datasets have been employed in this study to test and evaluate the performance and results of the proposed COB-LSTM model.

Keywords: *Educational Data Mining (EDM), Students' Performance Prediction, Chaotic Henry Gas Solubility (CHGS) optimization, Boost integrated Deep Long Short Term Memory (Bo-LSTM), UCI Public Dataset*

1. Introduction

In present times, student academic performance prediction has become a significant area of research in educational information mining [1, 2], which uses machine learning techniques to examine the data of educational systems. It can be difficult to determine the academic success of students because it depends on a number of distinct elements. Forecasting student scores [3] and marks based on their prior academic performance is a well-known and beneficial application in the field of education data mining. It is an excellent source of knowledge that might be applied in a variety of methods in order to improve the level of education across the country. According to pertinent research [4, 5], there are several techniques for predicting academic performance that are designed for carried out enhancements in the managerial and educational professionals of academic institutions. It is very important for learners as well as educators to predict students' academic achievement. By modifying their instruction and education system, instructors and social administrators can support students' learning plans by drawing on the prediction of learning outcomes [6]. Students' academic success can be affected by a number of variables, including their educational growth, personal traits, and behaviors related to academic activities. The majority of institutions now have trouble pulling in prospective students as a result of the more competitive academic markets. Therefore, it is crucial to do research on students' academic achievements [7-9] in order to assist their development and raise the standard of higher education, both of which will ultimately improve institutions' repute. The results of the prediction can also be used by instructors to determine which kind of learning behaviors are best for every group of students and how effectively to satisfy their requirements by offering them additional support. The forecast's results can also help students understand more about their performance and subsequently create effective learning strategies. One strategy to raise academic requirements and deliver superior educational services is through accurate prediction of student performance.

Educational data mining (EDM) [10, 11], which focuses on using data mining techniques to analyze educational features, includes forecasting student academic

[1]*Research Scholar, School of computing, Bharat Institute of Higher Education and Research, Chennai- 600073, India. Tamil Nadu, India. Email: ksankaranarayananphd@gmail.com*
[2]*Professor, School of computing, Bharat Institute of Higher Education and Research,*
Chennai-600073 ,Tamil Nadu, India. Email :drkumaravel@gmail.com

achievement. EDM can be separated into three distinct categories based on the aim of prediction, which includes forecasting of educational development, estimation of dropouts, and estimation of overall average student grade. The machine learning algorithms perform well on classification [12, 13] and prediction tasks, but in practice, the selection of an ML algorithm relies greatly on the data, including the kind of data element (that is, continuous or discrete), the model's final result, and other data properties. As a result, conducting an empirical investigation is required to identify the best machine learning method for any given application. Deep learning [14, 15] has made it possible for people to handle and analyze educational data that has been gathered from many sources, and this field of study is continuing to grow. A number of statistical approaches, data extraction, data visualization, and machine learning tools are used to examine educational data. The purpose of the research analytics produced from academic data is to examine information received from academic databases. Learning-management systems evaluate data, refine teaching techniques, and change the environment in which data is received.

These techniques are advantageous to a range of professions by attaining a variety of objectives, such as recognizing patterns, forecasting behavior, or uncovering changes. This makes it possible for teachers to give pupils the best learning techniques for keeping track of their development [16, 17]. The performance of the pupils can be predicted using machine learning algorithms using data from educational institutions and academic records. The learning environment must be transformed using contemporary methodology and useful techniques. The education data mining model is crucial for understanding the learning environment for students because it examines both the educational environment and machine learning methodologies. The current research use a variety of machine learning techniques to predict students' academic success in educational institutions [18]. They suffer from serious drawbacks like poor forecast accuracy, a complex system, protracted processing times, and costly overhead. The proposed work therefore aims to develop a new model for projecting students' academic progress. The original research objectives of this work are given below:

- A new approach for predicting and analyzing students' education performance is developed in this work using the public UCI education training dataset.
- Once a dataset has been obtained, the Chaotic Henry Gas Solubility (CHGS) optimization technique is used to select the best characteristics for a classification that is both accurate and low in data dimensionality.
- Next, a classification method based on the hybrid Boost integrated Deep Long Short Term Memory (Bo-LSTM) is used to forecast student performance with a high degree of accuracy.
- In order to validate and assess the performance of the proposed COB-LSTM model, a number of metrics and datasets have been used in this study.

The following sections comprise the remaining parts of this article: A complete analysis of the literature is presented in Section 2 for assessing the various machine learning-based frameworks for forecasting student performance. The proposed COB-LSTM is thoroughly discussed in Section 3. The overall performance evaluation is carried out using the comparison findings in Section 4. The summary of the study, the findings, and the suggestions are offered in Section 5.

2. Related Works

In order to anticipate student performance and implement corrective actions to enhance their performance, educational data mining is being extensively applied to extract usable knowledge from a vast number of datasets in higher educational environments. By considering the student's past experiences and educational achievement at educational institutions, machine learning approaches can forecast a student's performance. Machine learning [19, 20] is specifically used to estimate the likelihood that students will not succeed, allowing educators and administrators to quickly intervene on behalf of their students. A prominent topic for research in the fields of data mining, AI, big data, and educational data mining has received a lot of attention in the past. The goal of diverse study methods is to create a forecasting framework that will predict grades, scores, academic ratings, and academic suggestions. To examine and present the data, many methods and tools are employed.

Hussain, et al [21] applied a machine learning technique to predict the students’ academic performance prediction. This system has generated an excess of data, which must be carefully analyzed in order to extract the finest usable knowledge for management and subsequent growth. Ha, et al [22] conducted an empirical study on several machine learning techniques used for students’ academic performance prediction. Imran, et al [23] utilized a supervised learning techniques for predicting students’ performance in their academics. Enhancing the standard of educational procedures in order to improve student performance is one of the main difficulties. In order to meet the needs of students who do inadequately, instructors might improve their teaching methods. They can also provide deserving learners more support. Students who have a good sense of how well or badly they would perform in a course may find the prediction findings helpful in developing strategies. A long-term goal of all educational institutions worldwide is to increase retention of students. Increased retention has a variety of beneficial effects, including better job possibilities for alumni and improved college reputation and ranking. In order to assess the feature ranks, this study utilized a filter approach employing an information gain-based selection algorithm. It determines the characteristics to include in the performance model for students. Each feature is given a rank value during the feature selection process based on its impact on the classification of the information at hand. The ensemble method is used in this paper to develop the performance prediction model for students. A learning strategy called the ensemble technique blends several models to lower classification errors and improve the precision of inaccurate classifiers.

Abu Saa, et al [24] intended to analyze different factors used to assess the performance of students in their high education. According to the study's findings, the most prevalent variables may be divided into four primary categories: students' e-Learning engagement, statistics, prior grades and academic performance, and social information. The results also showed that artificial neural networks, decision trees, and Naive Bayes classifiers are the most often utilized data mining techniques to anticipate and analyze learners' aspects. Tomasevic, et al [25] investigated about the different types of supervised data mining techniques for students’ performance prediction. In order to solve the problem of predicting student examination performance, a thorough analysis and comparison of cutting-edge machine learning approaches was conducted in this study. The need for a tagged dataset makes supervised machine learning algorithms more appropriate for this type of prediction problem than unsupervised ones because studied data samples must be comprehended. The three primary groups of supervised machine learning methods—similarity-based, model-based, and probabilistic methods—were all taken into consideration. By identifying students with similar prior results, the similarity-based technique predicts performance on the test. The subsequent strategy is a model-based strategy that is motivated by the determination of hidden correlation between input learning data that make up the underlying model. Mengesh, et al [26] developed a support decision making system for validating and assessing students’ performance in university admission systems. The goal of this research is to assist higher education institutions in making wise admissions decisions by assessing students' academic potential before granting admission. Here, the four best prediction models such as Artificial Neural Network (ANN), Naïve Bayes (NB), Decision Tree (DT), and Support Vector Machine (SVM) are validated to make appropriate decisions while predicting students’ performance in academics.

Jalotta, et al [27] analyzed several machine learning and statistical techniques used for students’ performance prediction. Instead of only providing a basic categorization, statistical techniques typically have a correct underlying probability model that offers possibilities of belonging to each class. A statistical technique known as correlation analysis is used to determine the strength of the relationship between two numerically determined continuous variables. Regression analysis explains the relationship between an independent variable and a dependent variable in terms of numbers. The Bayesian Model employs frequentist methodology and the application of probability on data is the core of this approach. Calculations using the Bayesian approach focus solely on the hypothesis' probability. Injadat, et al [28] discussed about the feature analysis models for education data mining system. This study aims to forecast students' final scores in order to spot those who might require assistance earlier in the course. In order to analyze several machine learning approaches for forecasting students' academic success in the learning environment, the authors in [29] published a thorough study. In this review, the performance of the students in the learning environment is examined using a variety of data sources. Additionally, the PISA dataset was utilized in this study for system validation, with sample data being gathered from various institutions across nine different nations. Furthermore, early prediction aids in improving student activity logs, instruction effectiveness and academic potential. Waheed, et al [30] [32] [33] used a deep Artificial Neural Network (ANN) approach with a collection of hand-crafted characteristics to examine the way students learn in a classroom setting. The major goal of this effort is to develop a new framework for supporting the technologically enhanced learning platform, which is primarily focused on predicting students' academic success. According to this concept, student performance is anticipated to fall into the distinction, with-drawn, and pass-fail categories.

3. Proposed Methodology

The suggested Chaotic Optimized Boost LSTM (COB-LSTM) technique is thoroughly explained in this part in order to more accurately forecast student performance. This work's original contribution is the creation of a brand-new framework for evaluating students' academic performance in educational institutions using complex and cutting-edge optimization and classification approaches. The UCI public dataset is taken into account for system design and implementation in this framework. Then, feature optimization and selection are performed using the unique Chaotic Henry Gas Solubility (CHGS) optimization technique. The hybrid Boost integrated Long Short Term Memory (Bo-LSTM) technique is used to forecast student performance with high accuracy and minimal system complexity after getting the optimized feature set. The proposed COB-LSTM differs from other existing prediction systems in that it is straightforward to develop, simple to understand, takes little time to train and evaluate the features, and predicts outcomes with a high degree of accuracy. The suggested COB-LSTM based students' performance prediction system's work flow model is depicted in Fig1 and includes the following stages of operations:

- Preprocessing and normalization
- Feature selection using CHGS
- Students' performance prediction using Bo-LSTM
- Performance evaluation

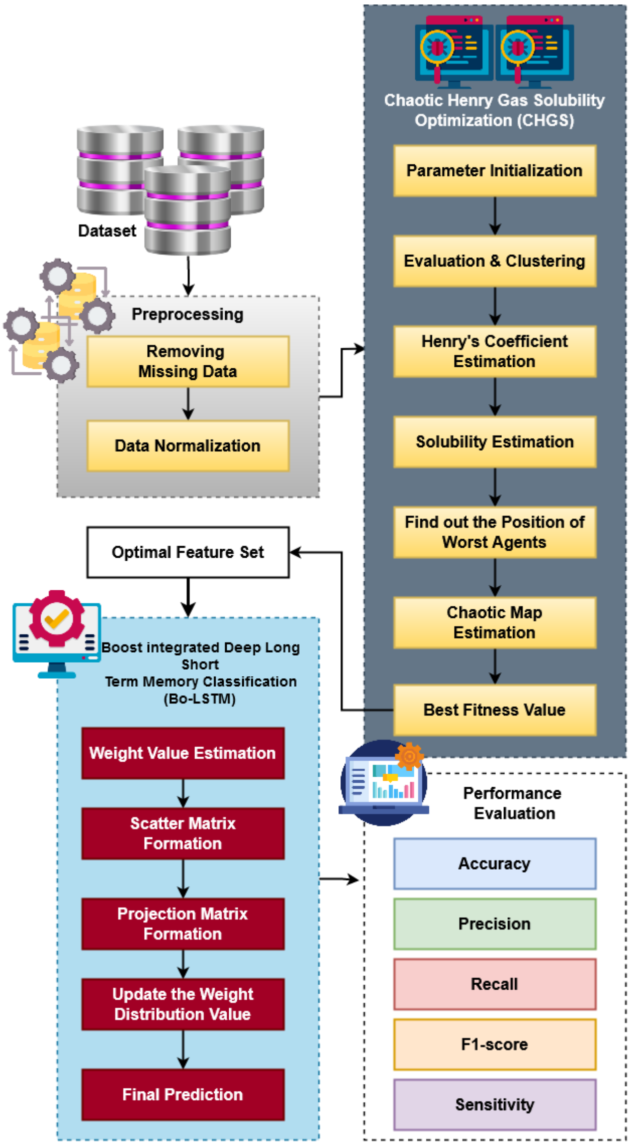

Fig 1. Flow of the proposed students' performance prediction system

Preprocessing is essential for data mining. Its goal is to transform unprocessed data into a format that mining algorithms can use. It includes the following operations:

- Integration
- Cleaning
- Discretization
- Balancing

Data integration is the process of combining information from various sources into a single repository. To guarantee data consistency, missing and inaccurate information are treated during the cleaning phase. This study's dataset was free of missing data, outliers, and other errors. The desired data are converted from numerical values to nominal values using the discretization technique. Following data pre-processing, the data balancing strategy is used in this phase to address the problem of class inequalities. When there are significantly fewer instances of one class than occurrences of any other class or classes, this is known as the class unbalanced problem. When the data is unbalanced, traditional classification algorithms perform well for majority classes because they focus more during classification on instances belonging to the majority of the class and less on those belonging to the minority class.

3. Chaotic Henry Gas Solubility (CHGS) Optimization for Feature Selection

The CHGS optimization technique is used in this stage to successfully reduce the dataset's dimensionality. High data dimensionality frequently leads to classifier complexity increases, slower training speeds, and longer training times. Choosing the most necessary and pertinent features is therefore more crucial for producing an accurate forecast or categorization. Thus, feature selection and data dimensionality reduction in the proposed study are accomplished using the CHGS-based optimization technique. In prior investigations, many feature optimization strategies were used, but they struggled with the main issues of poor search effectiveness, high processing complexity, and low convergence rate. In order to construct a comprehensive and successful model for feature optimization, the suggested study has this objective. This technique is developed based on the Henry's law that is specifically used for solving the real-world problems with low computational burden. This technique includes the following stages of operations:

- Parameter Initialization & Setup
- Clustering
- Best gas evaluation
- Henry's coefficient estimation
- Solubility estimation
- Position step updation
- Position identification of worst agents
- Chaotic map estimation

- Best optimal solution

Determine N, the total number of gases, and their beginning positions. Gas partial pressure needs to be set to zero. The following formula can be used to determine the position of the i^{th} gas, which is indicated by the sign $H_{i,j}$:

$$W_i(m+1) = H_{mn} + \delta w(H_{mx-H_{mn}}) \quad (1)$$

Where, δ indicates the random number, H_{mn} and H_{mx} are the lower and upper limits. In this method, selecting the number of clusters is important, and the same cluster should contain gases of the identical type. The best gas across all clusters is selected, as well as the best gas within each cluster. The gases can be ranked using the objective function. The position of the best gas among clusters is indicated by W_{best}, while the location of the best gas within each cluster j is shown by the symbol $W_{j,best}$. The following equation is used for determining each cluster's Henry's coefficient:

$$G_i(m+1) = G_i(m).\exp\left(-V_j\left(\frac{1}{M(m)} - \frac{1}{M^{\vartheta}}\right), M(m) = \exp\left(-\frac{m}{itr}\right)\right) \quad (2)$$

Where, $G_i(m+1)$ indicates the coefficient value, itr represents the iteration, $M(m)$ represents the temperature, and M^{ϑ} is the constant value. Consequently, the solubility is estimated for each gas as represented in the following model:

$$\mathcal{S}_{i,j}(m) = \mathscr{b}wG_j(m+1)w\mathfrak{p}_{i,j}(m) \quad (3)$$

Where, $\mathfrak{p}_{i,j}$ is the partial pressure, i indicates the gas belong to the cluster j, and $\mathscr{b}$ denotes the constant value. After that, the position updation is carried out by using the following equation:

$$W_{i,j}(m+1) = W_{i,j}(m) + \mathscr{f} \times \tau \times \rho \times (W_{j,best}(m) - W_{i,j}(m) + \mathscr{f} \times \tau \times \omega \times \left(\mathcal{S}_{i,j}(m) \times W_{best}(m) - W_{i,j}(m)\right) \quad (4)$$

$$\rho = \vartheta \times \exp\left(\frac{\mathscr{f}_{best}(m)+e}{\mathscr{f}_{i,j}(m)+e}\right), e \quad (5)$$

Where, τ is the random number, ρ, ω, and ϑ are the constants, $\mathscr{f}_{i,j}(m)$ defines the fitness of the gas, and $\mathscr{f}_{best}(m)$ indicates the best fitness value. Moreover, the number of worst agents are estimated with their updated positions. After that, the chaotic map functions including the chebyshev, circle, Gauss, iterative, logistic, sine, sinusoidal, and singer are computed for an effective optimization. The best optimal solution for selecting the most relevant characteristics from the provided data is found using the CHGS technique. It is also utilized to enhance categorization performance through better testing and training procedures.

4. Boost Integrated LSTM (Bo-LSTM) Classification for Performance Prediction

After feature selection, the Bo-LSTM classification methodology is used for predicting the students' performance results. It integrates the functions of an ensemble-based boost learning and deep LSTM classification methods for an accurate decision making. Since, the previous machine learning and deep learning techniques have the major difficulties in terms of high system complexity, increased training time for learning, low efficiency, and high overfitting. Therefore, the proposed work aims to implement a new classification model, named as, Bo-LSTM for an efficient students' performance prediction. In this technique, the input parameters such as base learner $\mathcal{L}$, and Number of rounds $\mathcal{K}$ are initialized with the attributes of the dataset as shown in below:

$$\mathcal{D} = \{(a_1, b_1), (a_2, b_2) \dots (a_x, b_x)\} \quad (6)$$

Where, $\mathcal{D}$ indicates the given dataset, and $(a_1, b_1), (a_2, b_2) \dots (a_x, b_x)$ are the attributes. Then, the weight value of the classifier is initialized by using the following model:

$$\mathcal{D}_s(i) = \begin{cases} \frac{1}{2x} & \text{if}, b_i = 0 \\ \frac{1}{2y} & \text{if}, b_i = 1 \end{cases} \quad (7)$$

Where, x and y are the number of samples for the data classes 0 and 1 respectively. Consequently, the mean estimation, and scatter matrix formulation operations are performed by using the following equations:

$$\mathscr{m}_k = \frac{1}{N}\sum_{i=1}^{N} \mathcal{D}_k(i)a_i \quad (8)$$

$$\mathscr{s}_k = \sum_{i=1}^{N} \mathcal{D}_k(i)(a_i - \mathscr{m})(a_i - \mathscr{m})^{\mathcal{K}} \quad (9)$$

Where, $k = 1,2 \dots \mathcal{K}$, N indicates the data samples, $\mathscr{m}_k$ represents the mean value, and $\mathscr{s}_k$ indicates the generated scatter matrix. After that, the eigen value u_i decomposition is performed as represented in the following equation:

$$\mathscr{s}u_i = \psi_i u_i, i = 1,2 \dots d' \quad (10)$$

Where, d' represents the dimensionality without losing any properties. Moreover, the projection matrix is constructed with the eigen vector U_v as defined in the following model:

$$U_v = [u_1, u_2 \dots u_{d'}] \in \mathfrak{R}^{d \times d'} \quad (11)$$

Furthermore, the LSTM is used as the base learner to forecast the data as represented in below:

$$g_k = \mathrm{LSTM}(\mathcal{D}, U_v, \mathcal{D}_s) \quad (12)$$

Where, g_k indicates the hypothesis value, $\mathcal{D}$ is the provided data, and $\mathcal{D}_s$ represents the distribution function. Subsequently, the error and weight value for g_k are computed by using the following mathematical models:

$$\varepsilon_k = W_{i\sim\mathcal{D}_i}[g_k(a_i \neq b_i)] \quad (13)$$

$$\omega_k = \frac{1}{2} \mathrm{In} \frac{1-\varepsilon_k}{\varepsilon_k} \quad (14)$$

According to the estimated weight value for the base learner, the weight value is updated for the distributed with the normalization factor as defined in below:

$$\mathcal{D}_{k+1}(i) = \frac{\mathcal{D}_k(i)}{H_i} \times \begin{cases} \exp(-\omega_k), & \text{if} g_k(a_i) = b_i \\ \exp(\omega_k), & \text{if} g_k(a_i) \neq b_i \end{cases} \quad (15)$$

The final prediction result of the Bo-LSTM classifier is in the following form:

$$G(a) = \begin{cases} 1, & \text{if} \sum_{k=1}^{\mathcal{K}} \omega_k g_k(a) > 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

By using this classification algorithm, an accurate performance prediction results are obtained that is used to analyze the students' academic achievements.

4. Results and Discussion

By using the public UCI education dataset, the suggested COB-LSTM model's performance outcomes are validated and compared. The dataset description [31] details are provided in Table 1.

Table 1. Dataset description

Attributes	Description
Age	Age of the student
Gender	Gender of the student
Educational institution	Student's education institution
Address	Residential address of the student
Fedu	Father's education status
Medu	Mother's education status
Fjob	Father's job
Mjob	Mother's job
Fsize	Family size
Reason	Reason to select the education institution
Tra_time	Travel time from home to institute
Failures	Number of failures in the previous classes
Activities	Extracurricular activities
Health	Health status
Absences	Number of school absences
Grade 1	First period grade
Grade 2	Second period grade
Grade 3	Final grade

$$\mathrm{Accuracy} = \frac{Tp+Tn}{Tp+Tn+Fp+Fn} \quad (17)$$

$$\mathrm{Precision} = \frac{Tp}{Tp+Fp} \quad (18)$$

$$\mathrm{Recall} = \frac{Tp}{Tp+Fn} \quad (19)$$

$$\mathrm{F1-score} = \frac{2*\mathrm{Precision}*\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}} \quad (20)$$

$$\mathrm{MCC} = \frac{Tp*Tn-Fp*Fn}{\sqrt{(Tp+Fp)(Tp+Fn)(Tn+Fp)(Tn+Fn)}} \quad (21)$$

$$\mathrm{Kappa} = \frac{x_o-x_e}{1-x_e} \quad (22)$$

Where, Tp – true positives, Tn – true negatives, Fp – false positives, Fn – false negatives, x_o – observed agreements, and x_e – expected agreements. The suggested COB-LSTM model's overall performance is validated in Table 2 and Fig. 2 by employing various parameters in relation to the pass, fail, and average classes. According to an analysis of the observations, the proposed COW-LSTM technique offers better performance results for the given data with high accuracy and precision values up to 98.8%. The overall performance of the proposed COB-LSTM model is significantly improved in this study with the help of appropriate data handling and feature optimization techniques.

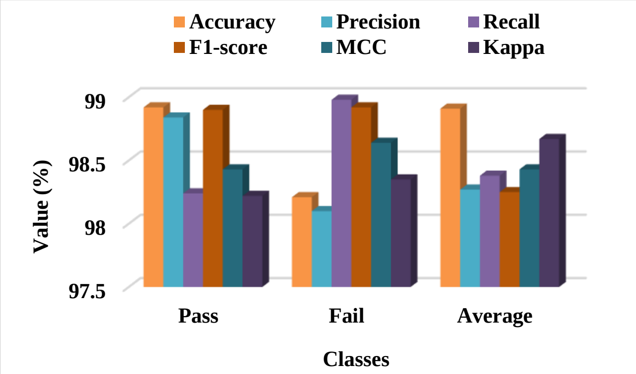

Fig 2. Performance analysis for the overall dataset

Table 2. Overall performance analysis

Predicted classes	Accuracy	Precision	Recall	F1-score	MCC	Kappa
Pass	98.92	98.84	98.24	98.9	98.43	98.22
Fail	98.21	98.10	98.98	98.92	98.64	98.35
Average	98.91	98.27	98.38	98.25	98.43	98.67

The suggested COB-LSTM technique's accuracy and precision values are validated in Table 3 both with and without the use of feature selection methods. Then, in Fig. 3 and Fig. 4, respectively, their graphical illustrations are shown. Some well-known methodologies, including SVM, NB, KNN, 1D-CNN, and ANN, are taken into consideration for this assessment and contrasted with the suggested model. According to the final results, the suggested COB-LSTM technique outperforms the current models with high precision and accuracy values up to 98.9% and 98.5%, respectively. The outcomes of learning models that use feature selection strategies have significantly improved as compared to those without. Since, the feature selection plays a vital role in the classification or prediction system, which supports to improve the learning speed of classifier by optimally picking the most reliable features for classification.

Table 3. Accuracy and precision performance analysis with and without feature selection

Techniques	With FS		Without FS	
	Accuracy	Precision	Accuracy	Precision
SVM	85	82	79	80
NB	81	87	77	85
KNN	81	78	80	72
1D-CNN	90	89	85	90
ANN	86	85	83	86
Proposed	98.9	98.5	95	96

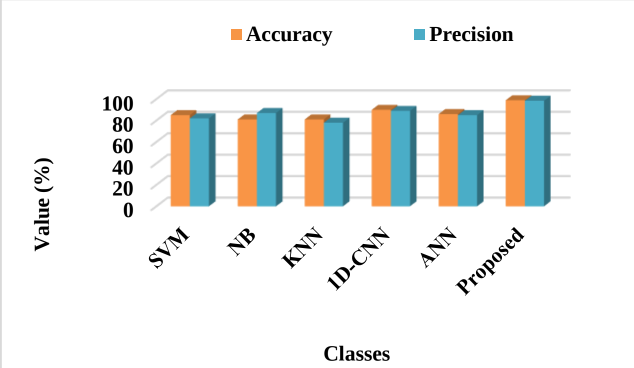

Fig 3. Performance analysis for the overall dataset with feature selection

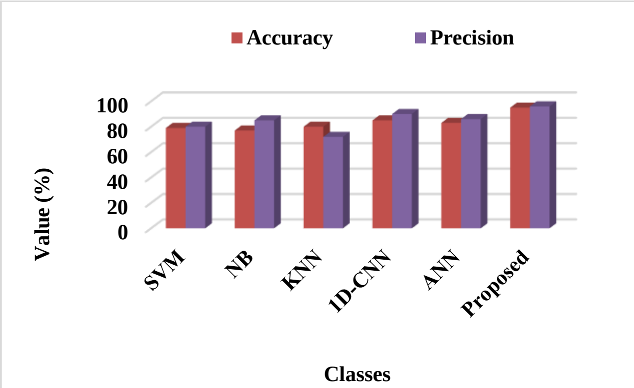

Fig 4. Performance analysis for the overall dataset without feature selection

The suggested COB-LSTM technique's training and testing performance outcomes are validated in this study and compared with samples of 75% and 25%, respectively. With 75% of the training dataset, Table 4 and Fig. 5 support the COB-LSTM technique's performance results. The testing results are similarly validated by Table 5 and Fig. 6 using 25% data samples. Overall evaluations lead to the conclusion that the proposed COB-LSTM improves performance outcomes for all classes, with training results outperforming test results.

Table 4. Performance evaluation of COB-LSTM using training dataset (75%)

Classes	Accuracy	Precision	Recall	F1-score	MCC	Kappa
Pass	99.12	98.9	99	98.8	98.7	99.1
Fail	98.96	98.5	98.9	98.87	98.68	98.92
Average	98.95	98.86	98.87	98.69	98.99	99.1

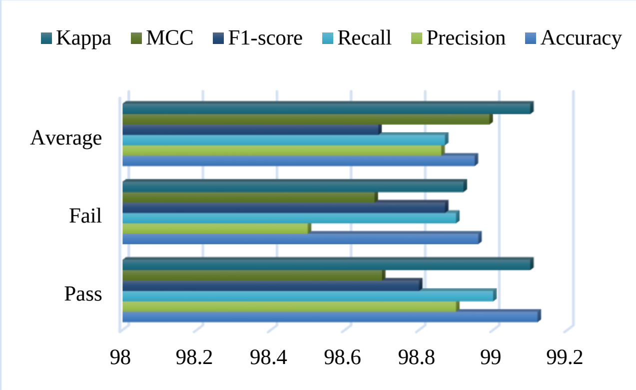

Fig 5. Performance evaluation with the training dataset (75%)

Table 5. Performance evaluation of COB-LSTM using training dataset (25%)

Classes	Accuracy	Precision	Recall	F1-score	MCC	Kappa
Pass	98.89	98.85	98.69	98.64	98.58	99
Fail	98.95	98.48	98.61	98.35	98.78	98.94
Average	98.85	98.67	98.76	98.67	98.57	98.69

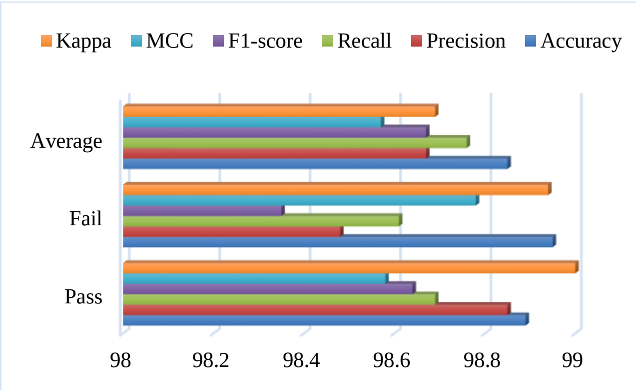

Fig 6. Performance evaluation with the training dataset (25%)

5. Conclusion

This study creates a brand-new framework called the COB-LSTM model to forecast students' academic achievement. Here, the suggested system has been put into practice using a publicly accessible UCI education dataset. The unique CHGS optimization method is then used to undertake feature optimization and selection. After obtaining the optimized feature set, the hybrid Bo-LSTM technique is applied to predict student performance with high accuracy and little system complexity. The proposed COB-LSTM is distinct from other current prediction systems in that it is easy to create, easy to comprehend, fast to train and quick to evaluate the features, and predicts outcomes with a high degree of accuracy. The final step is to evaluate and test the overall results and efficacy of the suggested strategy for predicting students' performance. In this work, the COB-LSTM results were validated using the publicly accessible student database from UCI. Age, gender, residence, father's job, father's education, mother's employment, mother's education, and other characteristics are all included. Additionally, the prediction outcomes of the recommended technique are validated using the training data (75%) and testing (25%) data. Overall results show that the COB-LSTM model outperforms other learning methods with accuracy and precision of 98.9% and 98.5%, respectively. Future improvements to the work can be made by incorporating a real-time dataset for system analysis and implementation using cutting-edge deep learning techniques.

Funding Details

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

Conflict of Interest

The authors declare that they have no conflict of interest.

Competing Interests

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] N. R. Yadav and S. S. Deshmukh, "Prediction of Student Performance Using Machine Learning Techniques: A Review," in International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022), 2023, pp. 735-741.

[2] G. Phatai and T. Luangrungruang, "A Comparative Study of Hybrid Neural Network with Metaheuristics for Student Performance Classification," in 2023 11th International Conference on Information and Education Technology (ICIET), 2023, pp. 448-452.

[3] I. Babu, R. MathuSoothana, and S. Kumar, "Evolutionary Algorithm Based Feature Subset Selection for Students Academic Performance Analysis," Intelligent Automation & Soft Computing, vol. 36, 2023.

[4] D. Yewale, S. Vijayaragavan, and V. Bairagi, "An Effective Heart Disease Prediction Framework based on Ensemble Techniques in Machine Learning," International Journal of Advanced Computer Science and Applications, vol. 14, 2023.

[5] E. Evangelista, "An Optimized Bagging Ensemble Learning Approach Using BESTrees for Predicting Students' Performance," International Journal of Emerging Technologies in Learning (Online), vol. 18, p. 150, 2023.

[6] I. Issah, O. Appiah, P. Appiahene, and F. Inusah, "A systematic review of the literature on machine learning application of determining the attributes influencing academic performance," Decision Analytics Journal, p. 100204, 2023.

[7] A. Kukkar, R. Mohana, A. Sharma, and A. Nayyar, "Prediction of student academic performance based on their emotional wellbeing and interaction on various e-learning platforms," Education and Information Technologies, pp. 1-30, 2023.

[8] Z. Wang, W. Yan, C. Zeng, Y. Tian, and S. Dong, "A unified interpretable intelligent learning diagnosis framework for learning performance prediction in intelligent tutoring systems," International Journal of Intelligent Systems, vol. 2023, 2023.

[9] S. Batool, J. Rashid, M. W. Nisar, J. Kim, H.-Y. Kwon, and A. Hussain, "Educational data mining to predict students' academic performance: A survey study," Education and Information Technologies, vol. 28, pp. 905-971, 2023.

[10] P. Guleria and M. Sood, "Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling," Education and Information Technologies, vol. 28, pp. 1081-1116, 2023.

[11] T. Cardona, E. A. Cudney, R. Hoerl, and J. Snyder, "Data mining and machine learning retention models in higher education," Journal of College Student Retention: Research, Theory & Practice, vol. 25, pp. 51-75, 2023.

[12] G. Shen, S. Yang, Z. Huang, Y. Yu, and X. Li, "The prediction of programming performance using student profiles," Education and Information Technologies, vol. 28, pp. 725-740, 2023.

[13] A. Emerson, W. Min, R. Azevedo, and J. Lester, "Early prediction of student knowledge in game-based learning with distributed representations of assessment questions," British Journal of Educational Technology, vol. 54, pp. 40-57, 2023.

[14] M. Alshurideh, B. Al Kurdi, S. A. Salloum, I. Arpaci, and M. Al-Emran, "Predicting the actual use of m-learning systems: a comparative approach using PLS-SEM and machine learning algorithms," Interactive Learning Environments, vol. 31, pp. 1214-1228, 2023.

[15] M. H. B. Roslan and C. J. Chen, "Predicting students' performance in English and Mathematics using data mining techniques," Education and Information Technologies, vol. 28, pp. 1427-1453, 2023.

[16] K. Karataş, I. Arpaci, and Y. Yildirim, "Predicting the culturally responsive teacher roles with cultural intelligence and self-efficacy using machine learning classification algorithms," Education and Urban Society, vol. 55, pp. 674-697, 2023.

[17] R. Mehdi and M. Nachouki, "A neuro-fuzzy model for predicting and analyzing student graduation performance in computing programs," Education and Information Technologies, vol. 28, pp. 2455-2484, 2023.

[18] Y. Teng, J. Zhang, and T. Sun, "Data-driven decision-making model based on artificial intelligence in higher education system of colleges and universities," Expert Systems, vol. 40, p. e12820, 2023.

[19] A. B. Bhatia, K. Mittal, R. Mahajan, and P. Whig, "Computational Psychometrics Social Analysis of Learners in Their Learning Behaviour Using AI Algorithms," in AI-Enabled Social Robotics in Human Care Services, ed: IGI Global, 2023, pp. 1-32.

[20] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," Smart Learning Environments, vol. 9, p. 11, 2022.

[21] S. Hussain and M. Q. Khan, "Student-Performulator: Predicting Students' Academic Performance at Secondary and Intermediate Level Using Machine Learning," Annals of Data Science, vol. 10, pp. 637-655, 2023/06/01 2023.

[22] D. T. Ha, P. T. T. Loan, C. N. Giap, and N. T. L. Huong, "An empirical study for student academic performance prediction using machine learning techniques," International Journal of Computer Science and Information Security (IJCSIS), vol. 18, pp. 75-82, 2020.

[23] M. Imran, S. Latif, D. Mehmood, and M. S. Shah, "Student academic performance prediction using supervised learning techniques," International Journal of Emerging Technologies in Learning, vol. 14, 2019.

[24] A. Abu Saa, M. Al-Emran, and K. Shaalan, "Factors affecting students' performance in higher education: a systematic review of predictive data mining techniques," Technology, Knowledge and Learning, vol. 24, pp. 567-598, 2019.

[25] N. Tomasevic, N. Gvozdenovic, and S. Vranes, "An overview and comparison of supervised data mining techniques for student exam performance prediction," Computers & education, vol. 143, p. 103676, 2020.

[26] H. A. Mengash, "Using data mining techniques to predict student performance to support decision

making in university admission systems," Ieee Access, vol. 8, pp. 55462-55470, 2020.

[27] C. Jalota and R. Agrawal, "Analysis of educational data mining using classification," in 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 2019, pp. 243-247.

[28] M. Injadat, A. Moubayed, A. B. Nassif, and A. Shami, "Systematic ensemble model selection approach for educational data mining," Knowledge-Based Systems, vol. 200, p. 105992, 2020.

[29] B. Albreiki, N. Zaki, and H. Alashwal, "A systematic literature review of student'performance prediction using machine learning techniques," Education Sciences, vol. 11, p. 552, 2021.

[30] H. Waheed, S.-U. Hassan, N. R. Aljohani, J. Hardman, S. Alelyani, and R. Nawaz, "Predicting academic performance of students from VLE big data using deep learning models," Computers in Human behavior, vol. 104, p. 106189, 2020.

[31] S. Shreem and H. Turabieh, "Student's Performance Prediction using Hybrid Machine Learning Classifiers," Int. J. Comput. Sci. Inf. Secur.(IJCSIS), vol. 19, pp. 87-103, 2021.

[32] Dineshkumar, R., Kalimuthu, M., Deepika, K., & Gopalakrishnan, S. (2022). Engineering education with tool based technical activity (TBTA). Journal of Engineering Education Transformations, 36(2, October), 185–191. https://doi.org/10.16920/jeet/2022/v36i2/22166

[33] Dineshkumar, R., Kalimuthu, M., Deepika, K., & Gopalakrishnan, S. (2022). Engineering education with tool based technical activity (TBTA). Journal of Engineering Education Transformations, 36(2, October), 185–191. https://doi.org/10.16920/jeet/2022/v36i2/22166

[34] Vaqur, M. ., Kumar, R. ., Singh, R. ., Umang, U., Gehlot, A. ., Vaseem Akram, S. ., & Joshi, K. . (2023). Role of Digitalization in Election Voting Through Industry 4.0 Enabling Technologies. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2), 123–130. https://doi.org/10.17762/ijritcc.v11i2.6136

[35] Elena Petrova, Predictive Analytics for Customer Churn in Telecommunications , Machine Learning Applications Conference Proceedings, Vol 1 2021.

[36] Soundararajan, R., Stanislaus, P. M., Ramasamy, S. G., Dhabliya, D., Deshpande, V., Sehar, S., Bavirisetti, D. P. Multi-Channel Assessment Policies for Energy-Efficient Data Transmission in Wireless Underground Sensor Networks (2023) Energies, 16 (5), art. no. 2285, .