International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No.4,2025
https://theaspd.com/index.php

Smart Education Analytics: Student Academic Performance Prediction Using Ensemble Of Deep Learning Models With Improved Pelican Optimization Algorithm

Praveena Chakrapani[1], **A Anthonisan**[2]
*School of Computing Sciences Hindustan Institute of Technology and Science,*Chennai, India
rp.18772016@student.hindustanuniv.ac.in
School of Computing Sciences Hindustan Institute of Technology and Science, Chennai, India
dyregistrar@hindustanuniv.ac.in

Abstract
Educational institutions, in this era of Technological Advancements require advanced, precise innovations in order to deliverstudent-centric and quality education to learners. In spite of the vast amount of data accessible to educational institutions, there is still ashortage for accurate methods, which observes and evaluates academic performance of students to actively target helpful activities that would translate the attempts of students and educational institutions, into precise analytical methodsby utilizing Student Intervention Strategies in a proactive manner. Educational Data Mining (EDM) is that domain of Artificial Intelligence (AI) that utilizes an integration of Data Mining (DM) and Machine Learning (ML) models for making forecasts on aspects particularly related to teachers and students, and, educational institutions, in general. In this manuscript, we offer a solution for Student Academic Performance Prediction Using an Ensemble of Deep Learning Models and Improved Pelican Optimization Algorithm (SAPPEDL-IPOA). The main objective of the SAPPEDL-IPOA technique is to provide a precise and strong predictive framework for enhancing the monitoring of students' academicperformance assessments. At a primary stage, the SAPPEDL-IPOA model applies the min-max scaling approach for data pre-processing to ensure that all features are scaled within a specific range. Next, an ensemble of DL approaches, namely, Bi-directional Long Short-Term Memory (BiLSTM), Double Deep Q Networks (DDQN), and stacked AutoEncoder (SAE) model is utilized for classification of Students' Academic Performance. Finally, the Improved PPelican Optimization Algorithm (IPOA) is used for hyper-parameter tuning, in orderto optimize the parameters of the Ensemble models to enhance classification performances. To validate the improved performance of SAPPEDL-IPOA model, an extensive range of simulations is implemented and the occurances and outcomes are studied under various aspects. The comparison of the investigations delineates the improvement of the novel SAPPEDL-IPOA system under various metrics.
Keywords: *Educational Data Mining; Academic Performance; Deep Learning; Pelican Optimization Algorithm; Data Normalization*

INTRODUCTION

Currently, due to the rapid, tremendous growth of competitive academic markets, most universities are facing complications in captivating and enhancing possible learners [1]. The investigation of academic activities of students is consequently of greater importance in assisting the improvement of students and enhancing the superiority of higher education that ultimately enriches reputation of institutions [2]. The outcomes attained from academic performance prediction might be utilized to classify learners, permitting the university to offer them additional provisions like tutoring sources and timely assistance. Mentors can also utilize the results of prediction to recognize the most suitable learning behaviour for every student group and offer them extra support intended for their needs [3]. Additionally, the outcome of prediction may support learners to create knowledge about their performance and then progress in suitable learning methods. A precise forecast of student success is unique method to enhance academic quality and offer better educational services. The data gathered from students are generally utilized to take simple queries for making decisions. However, most of the data remain unused owing to complications and huge amount of datasets [4]. Thus, to examine this large number of educational data requires considerable attention and is the primary focus in order to forecast performance of students. Data mining is the method of identifying valuable data from large datasets [5]. It is implemented effectively in various fields comprising medical, business, and banking now utilized for purpose of education named Educational Data Mining (EDM). In general, DM offers numerous models for the process of analysis that contain clustering, association rules, and classification [6].EDM is the area of specialization in mining educational data to determine interesting patterns and expertise in educational organizations. EDM is utilized in emerging models in order to assess educational content for optimum understanding of student performances [7]. These substantial student data sets are generally heavy to handle, greatly unbalanced, and can lead to complications [8]. Valuable data can be explored systematically to forecast the performance of students and aid educators in giving an effective teaching approach. Over recent years, multiple investigators have projected diverse models to project the finest student performance model employing versatile student data, algorithms, tools, and techniques [9]. Nowadays, Deep Learning (DL) and Machine Learning (ML) methodologies can forecast performance of students depending on the experience of students and their academic performance scores at university [10]. Particularly, the usage of DL and ML forms predictions of students' failure likelihood; thus administrators and teachers can give an early remedy to their students. This manuscript offers a Student Academic Performance Prediction Using an Ensemble of Deep Learning Models and Improved Pelican Optimization Algorithm (SAPPEDL-IPOA). At the primary stage, the SAPPEDL-IPOA model applies the min-max scaling approach for data pre-processing. Next, an ensemble of DL methods namely, Bi-directional Long Short-Term Memory (BiLSTM), Double Deep Q Networks (DDQN), and Stacked Auto-Encoder (SAE) model is utilized for classification of student academic performances. Finally, the Improved Pelican Optimization Algorithm (IPOA) is used for hyperparameter tuning to optimize the parameters of the ensemble models to enhance classification performances. To validate the improved performance of SAPPEDL-IPOA model, an extensive range of simulations take place and the outcomes are inspected under various measures.

Related Works

Zhao et al. [11] present a novel model to forecast performance of students by changing 1D student online learning behavior data into 2D images employing 4 different text-to-image encoding models: Gramian Angular Field (GAF), Pixel Representation (PR), Recurrence Plot (RP), and Sine Wave Transformation (SWT). Moreover, conventional ML models like SVM, Extra Trees, RF, Stochastic Gradient Descent, KNN, NB, LR, Gradient Boosting, AdaBoost, and DT were utilized on the untransformed and raw data with SMOTE model for comparison. Abdasalam et al. [12] projected an Optimized Ensemble DNN to improve the precision of forecasting grades of students. This method resolves the concern of diverse and complex performance of student data by utilizing DNN, ensemble learning, and some optimization models like SGD, RMS Prop, and Adam. This model utilizes a strong framework that combines these technologies to effectually discern complicated patterns, so upgrading predictive precision. In [13], a novel hybrid DL approach with enhanced entropy rough set theory is introduced. A hybrid DL-based CRN is applied to forecast and the classifier solution is upgraded by the GRSO model. Where the hyper-parameters of the RNN and CNN are enhanced by the GRSO model.Al-Ameri et al. [14] proposed method utilizing features extracted from CNN together with ML techniques to improve predictive precision. This model removes the requirement for manual feature extractor and generates higher consequences related to utilizing DL and ML approaches separately. In the beginning, 9 ML approaches are implemented to either the convoluted or original features. The authors [15] developed the GNN-TINet method to overwhelm the limitations of preceding methods which fail to effectually take complex connections in multi-label contexts, here students can demonstrate several performance classes simultaneously. The GNN-TINet employs GNN, transformer architectures, and InceptionNet to upgrade accuracy in multi-label student performance prediction. Cutting-edge pre-processing methods namely CAI and CFI are employed.Kaur et al. [16] projected a DL method for the identification of strong and weak students utilizing multi-parametric analysis and ensemble learning model. It integrates multiple ML models, comprising SVM, NB, LR, and Multi-Layer Perceptron utilizing an ensemble learning model to develop the performance of model. Moreover, a custom 1D-CNN is intended for classification. It employs multi-parametric study to recognize strong and weak students deliberating several parameters like academic performance, age, behaviour of online learning, and location. In [17], a novel MSPP which is derived from progressive feature engineering and data pre-processing methods utilizing DL is projected.

This paper advanced a model that targets the standard concerns related to educational databases through temporal and imbalanced surroundings that are also explainable across features of AI. In addition, adaptive hyper-parameter tuning and progressive GNN layers in the MSPP method permit to creation of output of more dense representation for forecasts ensuing in more precise classification.

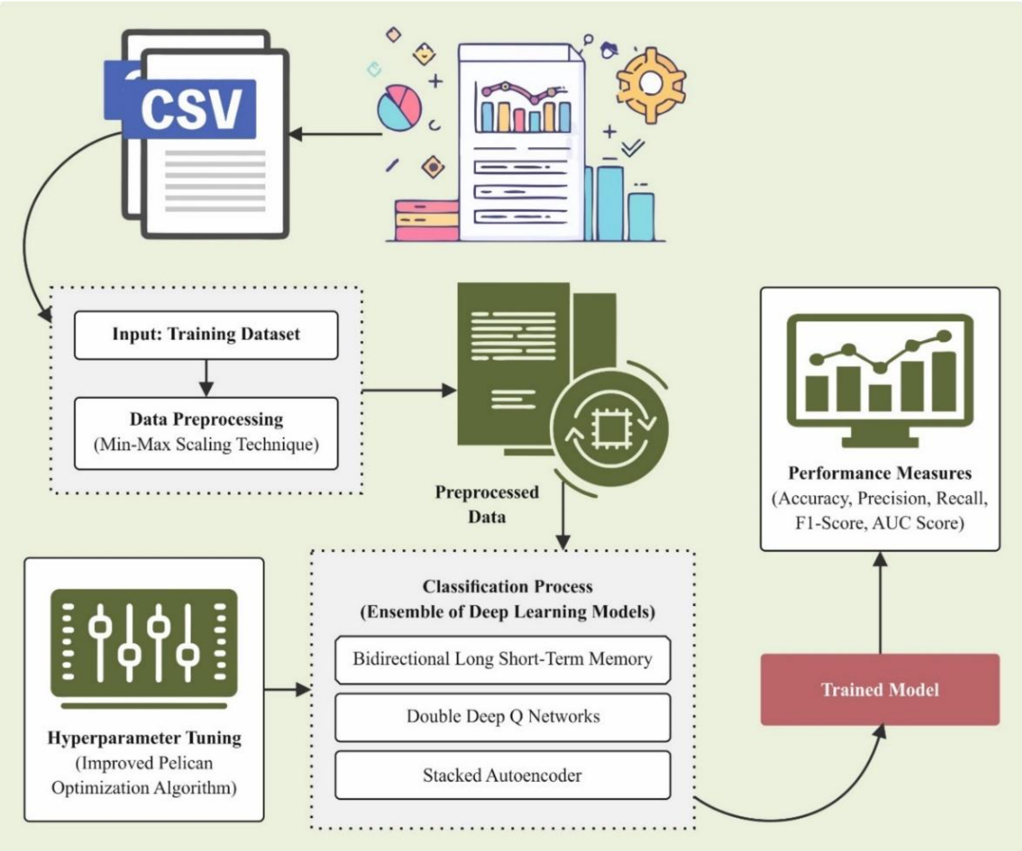

Fig. 1. Overall flow of the SAPPEDL-IPOA method

THE PROPOSED METHOD

In this manuscript, we present a novel SAPPEDL-IPOA methodology. The main objective of the SAPPEDL-IPOA technique is to provide a precise and strong predictive framework for enhancing the monitoring of student performance assessments. Fig. 1 depicts the entire flow of the SAPPEDL-IPOA system.

3.1. Min-Max scaling

At the primary stage, the SAPPEDL-IPOA model applies the min-max scaling approach for data pre-processing to ensure all features are scaled within a specific range. In the data cleansing procedure, normalization is extremely significant to connect the decision matrix in the method it is associated particularly as the conditions are not only non-compatible but also non-commensurable [18].

Define the minimal and maximal values: For all condition C_μ through each alternative $A\lambda$, calculate the minimal ($\min \mu$) and maximal ($\max \mu$) values:

$$\min\mu = \min_{\lambda}\left(tb_{\lambda\mu}\right) \quad (1)$$

$$\max\mu = \max_{\lambda}\left(tb_{\lambda\mu}\right) \quad (2)$$

Convert the data: Every value in the decision matrix is then converted utilizing the Min-Max scaling Eq. (3):

$$z_{\lambda\mu} = \frac{tb_{\lambda\mu} - \min\mu}{\max\mu - \min\mu} \quad (3)$$

This transformation results in the present outcome of the decision matrix $= [z_{\lambda\mu}]_{m*n}$, whereas all entries are scaled to a range from $(0, 1)$.

3.2. Ensemble of DL Models

Next, an ensemble of DL methods namely BiLSTM, DDQN, and SAE models are utilized for classification of student academic performances.

3.2.1. BiLSTM Method

LSTM is an expansion of the RNN structure, combining memory cells for storing long-range information that assistances in overcoming the problem of vanishing gradient in RNNs after processing longer sequential data [19]. This model has been broadly applied in different domains, namely natural language processing (NLP), sequence modeling, and speech recognition. LSTM involves 4 main gates, such as the cell state, input, output, and forget gates. The forget gate proceeds in the inputs h_{t-1} and x_t to make a value among $(0,$ 1) for c_{t-1}. Once the forget gate value is adjacent to 1, the cell layer information must be preserved, while as it is near 0, the information is rejected. The input gate is the 2nd gate, which controls what novel information should be added to the cell layer over incorporation of the $\tanh$ and sigmoid functions. The output gate is the last gate, which gives the hidden state h_t according to the upgraded state of the cell. The LSTM equations are characterized by the succeeding representations

$$f_t = \sigma\left(W_{xf}.x_t + W_{hf}.h_{t-1} + b_f\right) \quad (4)$$

$$i_t = \sigma(W_{xi}.x_t + W_{hi}.h_{t-1} + b_i) \quad (5)$$

$$\tilde{c}_t = \tanh(.x_t + W_{hc}.h_{t-1} + b_c) \quad (6)$$

$$c_t = f_t c_{t-1} + i_t \odot \tilde{c}_t \quad (7)$$

$$o_t = \sigma(W_{xo}.x_t + W_{ho}.h_{t-1} + b_o) \quad (8)$$

$$h_t = O_t \odot tanh(c_t) \quad (9)$$

Whereas, $W_{xf}, W_{xi}, W_{xc}, W_{xo}$ characterize the trained weights, b_f, b_i, b_c, b_o denotes trained biases, σ signifies the activation function of the sigmoid, x_t refers to input at time t, h_{t-1} is HL from preceding time step, c_t symbolize the cell layer at time t, h_t is output at time t, f_t stands for forgetting gate at time t, i_t symbolize input gate at time t, o_t signifies output gate at time t and designates the element-wise product.

Nevertheless, LSTM handles sequences in a single direction (forward) that bounds its capability to completely seize the setting of longer sequences. Bi-LSTM deals with the problem by processing information in either direction (forward or backward), allowing coordinated modeling of the context from either following or preceding words. The Bi-LSTM framework comprises dual layers: a forward layer, which manages series from the start to the end, and a backward layer which manages them from the end to the start. The dual HLs, $h_t^{forward}$ and $h_t^{backward}$ from the LSTM are combined to produce the last HLs, h_t^{BiLSTM} as delineated in Eq. (10).

$$h_t^{BiLSTM} = h_t^{forward} \oplus h_t^{backward} \quad (10)$$

using this bi-directional structure, Bi-LSTM allows improved contextual information capture and models long dependences in classification text. Fig. 2 portrays the infrastructure of BiLSTM.

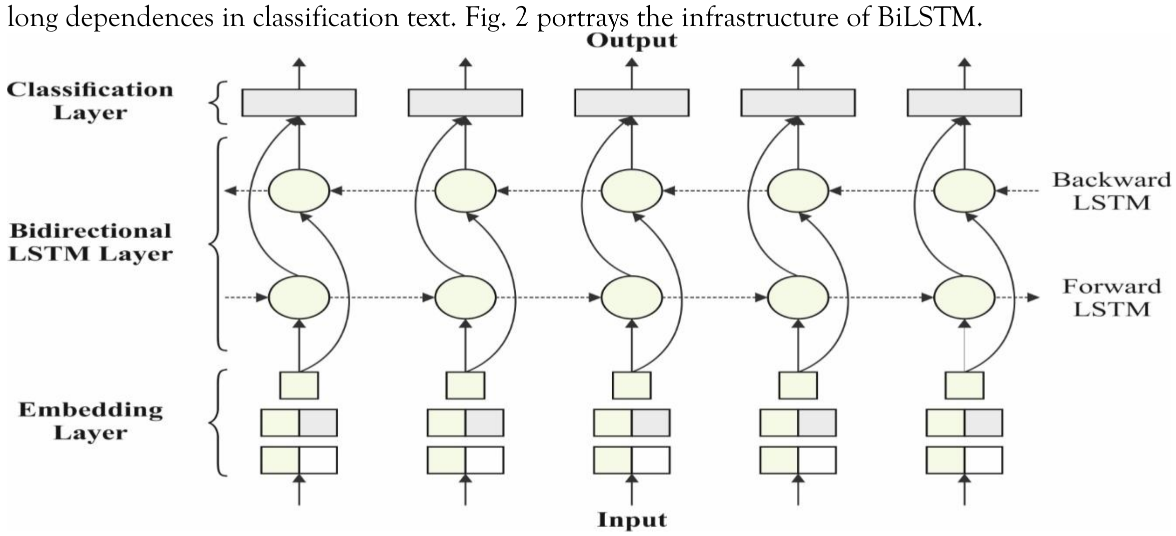

Fig. 2. Architecture of BiLSTM

3.2.2. DDQN Classifier

DQN incorporates principles of Qlearning using deep neural networks (DNNs) for processing composite atmospheres with higher-dimensional state spaces [20]. The DQN model has inspired the area of reinforcement learning (RL) by utilizing DNNs to estimate best action value functions that are important to make better decisions in the collection of states. The neural network (NN) is trained to reduce a loss function based on the error of time difference, incorporated with Q-learning upgrades and gradient descent optimization. The main problem of DQN is the update policy for the function of action-value that iteratively increases the rule. This method uses the Q-learning structure to obtain an optimizable loss function to train the NN. The update equation is described as shown (11):

$$Q(S_t, A_t) \leftarrow Q(S_t, A - t\ ) + \alpha[R_{t+1} + \gamma \max_{a}, Q(S_{t+1}, a') - Q(S_t, A_t)] \quad (11)$$

The training process includes decreasing a loss function, which calculates the difference between forecast and targeted Q-values. In Eq. (11), the loss function for the DQN model is expressed as shown in Eq. (12):

$$L(\theta_t) = \mathbb{E}[(Y_t^{DDQN} - Q(S_t, A_t; \theta_t))^2] \quad (12)$$

The targeted Q-value characterizes the predictable upcoming rewards, reduced by the feature γ, and it is important for the consistency of the learning procedure:

$$TargetQ\ = r + \gamma \max_{a}, Q(S_{t+1}, a'; \theta^-) \quad (13)$$

The DQN uses a distinct target network for stabilizing the learning upgrades. The target network's parameters are occasionally upgraded with the parameters from the major network (θ), avoiding the fast oscillation of target values that can undermine the learning procedure.

DDQN improves the unique DQN by dealing with the problem of Q-value overestimate. Either DQN or DDQN uses DNNs to estimate the Q-value function in environments with higher-dimensional state areas, marking a considerable development in RL. The DDQN method characterizes a significant iterative development above DQN. By splitting the action selection procedure from the evaluation of Q-value, DDQN decreases the over-optimistic value evaluations that can rise in DQN. The outline of this subtle but effective modification has been exposed to give more reliable and stable learning, and it prevents the agent from overestimating activities in policy growth. This allows, frequently results in improved performance on different benchmark tasks in RL.

3.2.3. SAE Model

The SAE network is a broadly employed DL structure which has dual core module: recognizer network (stacked encoders) and generative network (stacked decoders) [21]. It describes a symmetric framework with an equivalent amount of concealed neurons in the decoder and encoder layers. The encoder layers outline the novel input signals into deeper feature code, which calculates a count of high-level representations of multi-variant signals. The decoder layers use, still, deeper feature code to make the output signals such that they are as same as certain to the architecture of the input signals. Hence, abnormal instances are directly differentiated based on higher reconstruction errors than background instances. These stages are accompanied by supervised adjustment, thus networking-wide weights from the uppermost layer start to become enhanced over the backpropagation (BP) method to attain an experienced system. The non-linear functions of encoding and decoding for a new vector of feature x are designated as:

$$h_k = f_k(x) = \phi_{k,e}(W_{k,e}x_{k,i} + \beta_{k,e}) \quad (14)$$

$$\hat{h}_l = g_1(x) = \phi_{l,d}(W_{l,d}f(x_{l,i}) + \beta_{l,d})) \quad (15)$$

Here, $g_l(x)$ and $f_k(x)$ specify decoder and encoder functions in $k\ th$ hidden layer (HL) (h) and lth HL $(\hat{h}_l)$, individually. $W_{l,d}$ and $W_{k,e}$ represent weighting matrices in $\hat{h}_l$ andh_k, while $\beta_{1,d}$ and $\beta_{k,e}$ display biased vectors in $\hat{h}_l$ and h_k, correspondingly. The formulas$\phi_d(\cdot)$and$\phi_e(\cdot)$ denotes activation functions of $\hat{h}_l$ and h_k, individually, and they can be attuned by functions namely hyperbolic tangent, ramp, sigmoid, step, linear, maxout unit, and $ReLU$, to name a few.

3.3. IPOA-based Parameter Tuning

Finally, the IPOA is used for hyperparameter tuning to optimize the parameters of the ensemble models to enhance classification performances. POA serves as a population-based method, with Pelicans acting as its individuals [22]. Primarily, the population is stochastically initialized in the problem's explained upper and lower limits, stated as:

$$x_{i,j} = l_j + rand \cdot (u_j - l_j), i, j = 1,2, \dots, m \quad (16)$$

Whereas $x_{i,j}$ characterizes the $j\ th$ variable value of the $i\ th$ candidate performance, N signifies the population member counts, m refers to problem variables dimensionality, $rand$ designates a randomly generated number in the range $[0,1]$, and l_j and u_j represents upper and lower limits of the $j\ th$ problem variable, correspondingly.

Subsequently, the Pelicans start their searching stage which is separated into dual phases: Pelicans move toward prey (exploration stage) and Pelicans skimming across the surface of the water (exploitation stage).

Exploration Stage

For all Pelicans, its novel location is upgraded according to its present location and the location of its prey. When the value of the objective function at the novel location is lower than that of the prey the Pelican approaches the target; or else, it diverges from the prey. The location updated equation is provided by:

$$x_{i,j}^{P_1} = \begin{cases} x_{i,j} + rand \cdot (p_j - I \cdot x_{i,j}), F_p < F_j \\ x_{i,j} + rand \cdot (x_{i,j} - p_j), F_p \geq F_i \end{cases} \quad (17)$$

Whereas, $x_{i,j}^{P}{}^{1}$ characterizes the novel state of the ith Pelican in the jth size in the exploration stage. p_j signifies the prey's location in the jth size, and F_p is its value of the objective function. The parameter I stands for randomly generated integer that is both 1 and 2. These parameters are arbitrarily selected for all iterations and all members. If $I = 2$, it improves the displacement of all individuals, allowing them to discover novel areas in the searching region.

When the value of the objective function at the novel place is lower than that at the new location, then the novel place is established; or else, the new location remains consistent. This acceptance condition is stated as:

$$x_i = \begin{cases} x_i^{P_1}, F_i^{P_1} < F_j \\ x_i, \ \ F_i^{P_1} \geq \mathrm{F_i} \end{cases} \quad (18)$$

Exploitation Stage

After accomplishment the water surfaces, Pelicans extend their wings and take the air near the fish before picking them into their throat bags. This approach results in taking more fish in the aimed region. For all Pelicans, a novel location is arbitrarily generated near, and its objective function value is calculated. The equation to update the novel location is:

$$x_{i,j}^{P_2} = x_{i,j} + R \cdot \left(1 - \frac{t}{T}\right) \cdot (2 \cdot rand - 1) \cdot x_{i,j} \quad (19)$$

Whereas $x_{i,j}^{p_2}$ characterizes the novel condition of the $i\ th$ Pelican in the $j\ th$ size in the exploitation stage. Now, t denotes iteration counter, and T refers to maximal iteration counts. The term $\left(1 - \frac{t}{T}\right)$ symbolizes the neighborhood radius of $x_{i,j}$, considering local searches nearby all members to meet near best solutions inside the population. R stands for parameter that is set among $(0,\ 1)$, and $rand$ symbolize randomly generated number among $(0,\ 1)$.

When the objective function value at the novel location is lower than that at the new location, then the novel location is acknowledged; then, the new location stays the same. This is categorized as:

$$\chi_i = \begin{cases} x_i^{P_2}, F_i^{P_2} < F_i \\ x_{j,}\ F_i^{P_1} \geq \mathrm{F_i} \end{cases} \quad (20)$$

1. Iterative Chaotic Mapping and Refracted Opposition-Based Learning

Chaotic mapping is a kind of non-linear dynamical behavior. The aperiodicity and unpredictability of chaos might efficiently improve the optimizer efficacy of methods. Then performer optimizer pursues according to the randomness and ergodicity of chaos and, lastly, linearly converts the gained solutions to the optimizer variable area. The iterative chaotic mapping description is as shown:

$$z_{k+1} = \sin\left(\frac{a\pi}{z_k}\right) \tag{21}$$

Whereas $a \in (0,1)$ signifies controller parameter.

Opposition-Based Learning (OBL), is an optimization approach focused on expanding the search area of the population. The basic concept is to calculate the opposed solution of the present solution and then choose the improved solution for iteration to discover a further best solution for the provided problems. Nevertheless, in subsequent steps, the opposed solution might pass into the local optimal area.

To deal with this problem, Refracted OBL (ROBL) which incorporates the light refraction beliefs, was presented. It may increase the performance of the model and increase the search area to different amounts. The χ-axis characterizes the search optimizer range $[l, u]$, and the y-axis characterizes the standard line. β and α epitomize the refraction and incident angles, individually whereas $h^*$and h characterize the lengths of the refracted and incident rays. O refers to mid-point of the searching optimizer variety.

$$n = \frac{h^*((l+u)/2 - x')}{h(x^* - (l+u)/2)} \tag{22}$$

Let the scaling feature $k = h/h^*$ and replace it with Eq. (22) to get:

$$x^* = \frac{l+u}{2} + \frac{l+u}{2kn} - \frac{\chi'}{kn} \tag{23}$$

During this study, the early population of the Pelican is made utilizing an iterative chaotic mapping model incorporated by the ROBL. This model utilizes the features of the approach to guarantee uniform distributions of searching agents, permitting for the additional complete exploration of the searching area in particular ranges, and improving the speed of convergence. The mathematic expressions are outlined as:

$$\begin{cases} x_{i,j} = chaotic\_map(x_{init}, a) \\ x_{i,j}^{\star} = \dfrac{l_i + u_i}{2} + \dfrac{l_i + u_i}{2kn} - \dfrac{x_{i,i}}{kn} \end{cases} \tag{24}$$

Here, x_{init} denotes primary population. The IPOA system originates a fitness function (FF) for reaching an enhanced efficiency of classifier. It determines a positive number for denoting the better performance of candidate solution.

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{no.of\ misclassified\ samples}{Total\ no.of\ samples} * 100 \tag{25}$$

RESULT ANALYSIS AND DISCUSSION

The performance evaluation of the SAPPEDL-IPOA approach is studied under Students Performance dataset [23]. This dataset contains 2392 samples under 0-4 grades. The complete details of this dataset is specified in Table 1.

Table 1 Details of dataset

Grade	Description	No. of Samples
0	“'A' (GPA >= 3.5)”	107
1	“'B' (3.0 <= GPA < 3.5)”	269
2	“'C' (2.5 <= GPA < 3.0)”	391
3	“'D' (2.0 <= GPA < 2.5)”	414
4	“'F' (GPA < 2.0)”	1211
Total Samples		2392

Fig. 3 displays the classifier performances of the SAPPEDL-IPOA approach below 70% TRPH and 30% TSPH. Figs. 3(a) - 3(b) represent the confusion matrices through precise classification and identification of each class labels. Fig. 3(c) showcases the PR examination, which revealed higher performance over all classes. Eventually, Fig. 3(d) demonstrates the ROC study, which signifies skillful solutions with great ROC values for different classes.

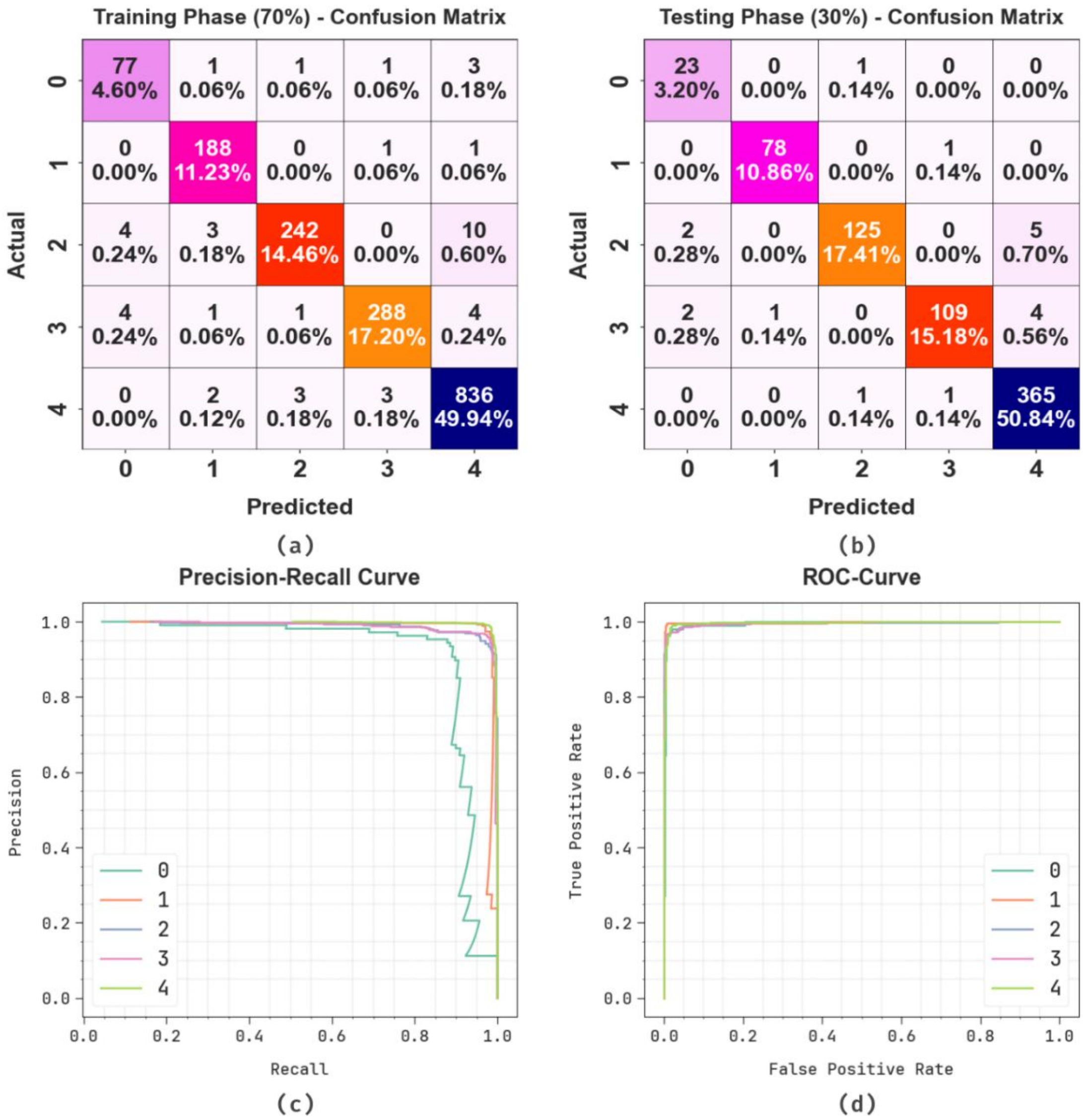

Fig. 3. Classifier outcomes of (a-b) 70%TRPH and 30%TSPH of confusion matrix and (c-d) curves of PR and ROC

Table 2 and Fig. 4 exemplifies the students' performance detection of SAPPEDL-IPOA system below 70% TRPH and 30% TSPH. The performances suggest that the SAPPEDL-IPOA model accurately recognized the samples. Using 70% TRPH, the SAPPEDL-IPOA method delivers average $accu_y$, $prec_n$, , $reca_l$, $F1_{score}$, and AUC_{score} of 98.97%, 96.23%, 96.17%, 96.18%, and 97.70%, respectively. Moreover, using 30%TSPH, the SAPPEDL-IPOA approach provides average $accu_y$, $prec_n$, , $reca_l$, $F1_{score}$, and AUC_{score} of 99.00%, 95.63%, 96.54%, 96.00%, and 97.87%, respectively.

Table 2 Students Performance detection of SAPPEDL-IPOA model under 70% TRPH and 30% TSPH

Class Labels	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$	AUC_{score}
TRPH (70%)					
0	99.16	90.59	92.77	91.67	96.13

1	99.46	96.41	98.95	97.66	99.24
2	98.69	97.98	93.44	95.65	96.54
3	99.10	98.29	96.64	97.46	98.14
4	98.45	97.89	99.05	98.47	98.44
Average	98.97	96.23	96.17	96.18	97.70
TSPH (30%)					
0	99.30	85.19	95.83	90.20	97.63
1	99.72	98.73	98.73	98.73	99.29
2	98.75	98.43	94.70	96.53	97.18
3	98.75	98.20	93.97	96.04	96.82
4	98.47	97.59	99.46	98.52	98.45
Average	99.00	95.63	96.54	96.00	97.87

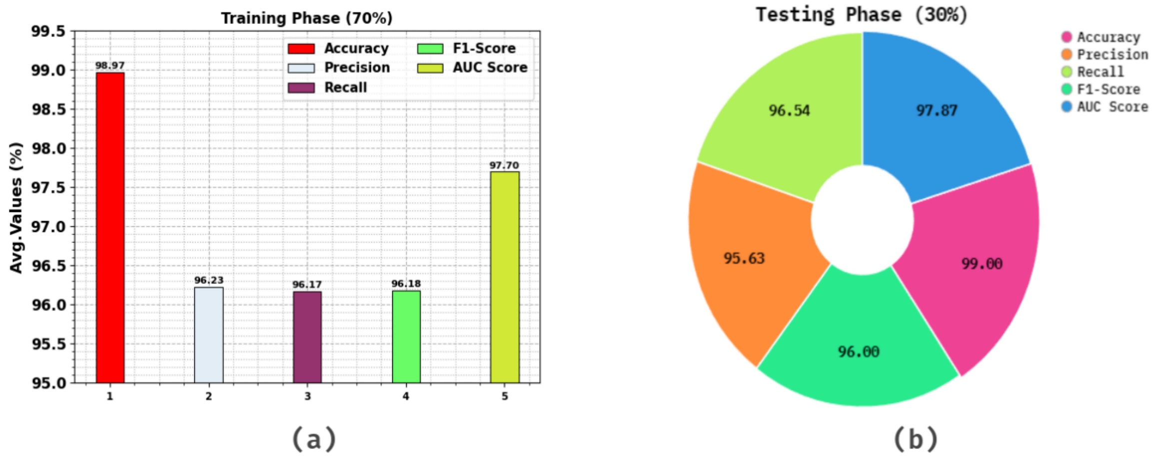

Fig. 4. Average of SAPPEDL-IPOA model under 70%TRPH and 30%TSPH

In Fig. 5, the training (TRA) $accu_y$ and validation (VAL) $accu_y$ performances of the I SAPPEDL-IPOA technique is depicted. The values of $accu_y$ are computed across a time period of 0-25 epochs. The figure underscores that the values of TRA and VAL $accu_y$ presents an increasing trend indicating the proficiency of the SAPPEDL-IPOA method with enhanced performance through multiple repetitions. Moreover, the TRA and VAL $accu_y$ values remain close through the epochs, notifying diminish overfitting and expresses superior performance of the SAPPEDL-IPOA algorithm, which guarantees reliable calculation even with unseen samples.

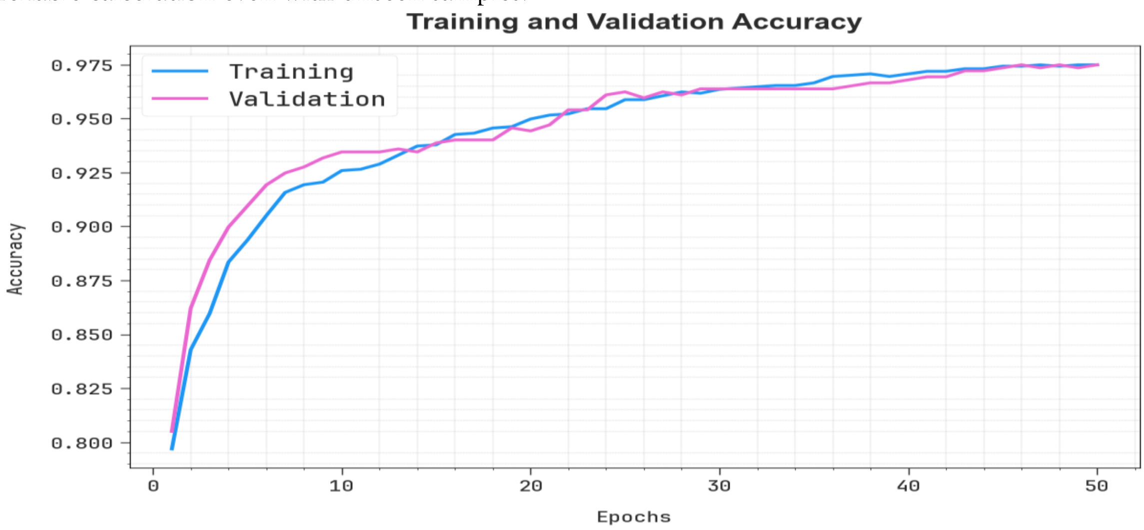

Fig. 5. $Accu_y$ curve of SAPPEDL-IPOA model

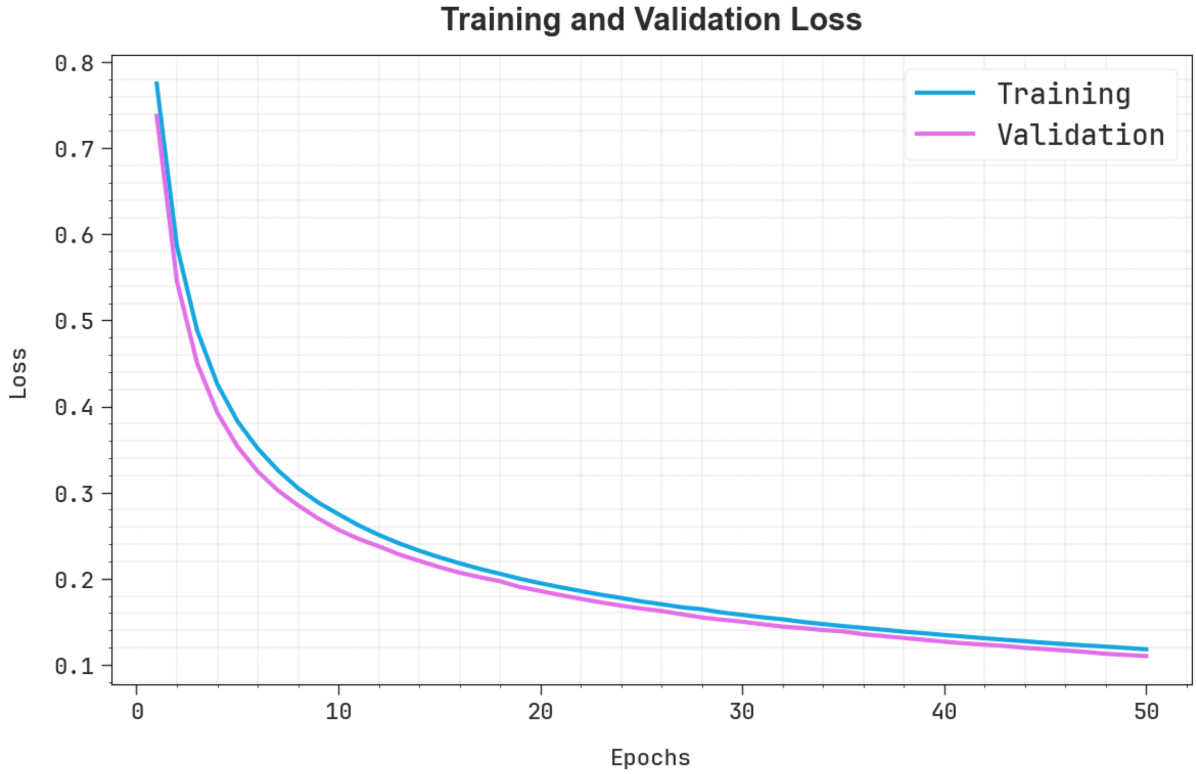

Fig. 6. Loss curve of SAPPEDL-IPOA model

In Fig. 6, the TRA loss (TRALOS) and VAL loss (VALLOS) graph of the SAPPEDL-IPOA technique is shown. The values of loss are computed through a time period of 0-25 epochs. It is exemplified that the values of TRALOS and VALLOS demonstrate a diminishing tendency, which highlights the competency of the SAPPEDL-IPOA approach in equalizing an equilibrium among generalization and data fitting. The successive dilution in values of loss as well securities the improved performance of the SAPPEDL-IPOA methodology and tune the calculation results after a while.

The comparative study of SAPPEDL-IPOA technique with existing algorithms are illustrated in Table 3 and Fig. 7 [24-26]. The model performances indicated that the SAPPEDL-IPOA system outperformed great performances. According to $accu_y$, the SAPPEDL-IPOA model has maximum $accu_y$ of 99.00% whereas the SMOTE + Knn, Wrapper-Based + DT, Forward + NB, SVM, GNB, NB, and AdaBoost techniques have decreased $accu_y$ of 97.06%, 94.87%, 85.67%, 98.11%, 98.01%, 93.55%, and 91.06%, correspondingly. Moreover, according to $Prec_n$, the SAPPEDL-IPOA technique has superior $accu_y$ of 95.63% whereas the SMOTE + Knn, Wrapper-Based + DT, Forward + NB, SVM, GNB, NB, and AdaBoost methodologies have diminish $accu_y$ of 90.28%, 92.84%, 88.50%, 94.71%, 88.98%, 86.90%, and 88.53%, correspondingly. Besides, according to $F1_{score}$, the SAPPEDL-IPOA method has enhanced $accu_y$ of 96.00% whereas the SMOTE + Knn, Wrapper-Based + DT, Forward + NB, SVM, GNB, NB, and AdaBoost approaches have lower $accu_y$ of 87.61%, 88.73%, 91.27%, 87.02%, 90.95%, 87.32%, and 96.00%, correspondingly.

Table 3 Comparative analysis of SAPPEDL-IPOA model with existing classifiers

Classifier	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$
SMOTE + Knn	97.06	90.28	88.89	87.61
Wrapper-Based + DT	94.87	92.84	87.50	88.73
Forward + NB	85.67	88.50	91.64	93.26
SVM Method	98.11	94.71	94.08	91.27

GNB Model	98.01	88.98	93.09	87.02
Naïve Bayes	93.55	86.90	87.56	90.95
AdaBoost	91.06	88.53	91.47	87.32
SAPPEDL-IPOA	99.00	95.63	96.54	96.00

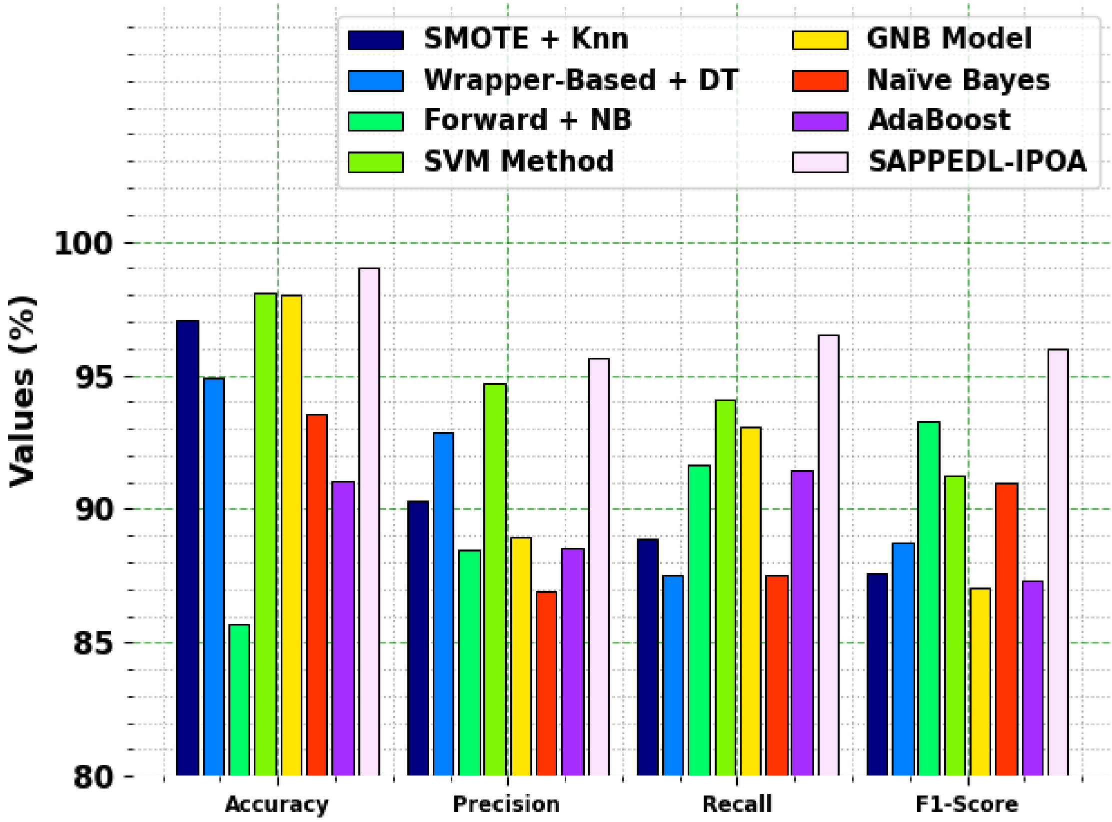

Fig. 7. Comparative analysis of SAPPEDL-IPOA model with existing classifiers

The processing time (PT) outcome of SAPPEDL-IPOA algorithm with existing methodologies are displayed in Table 4 and Fig. 8. According to PT, the SAPPEDL-IPOA system gets minimal PT of 2.29 min, while the SMOTE + Knn, Wrapper-Based + DT, Forward + NB, SVM, GNB, NB, and AdaBoost methodologies gain better PT values of 8.47 min, 6.81 min, 7.16 min, 8.38 min, 5.42 min, 4.37 min, and 5.80 min, respectively.

Table 4 PT outcome of SAPPEDL-IPOA model with existing classifiers

Classifier	Processing Time (min)
SMOTE + Knn	8.47
Wrapper-Based + DT	6.81
Forward + NB	7.16
SVM Method	8.38
GNB Model	5.42
Naïve Bayes	4.37
AdaBoost	5.80
SAPPEDL-IPOA	2.29

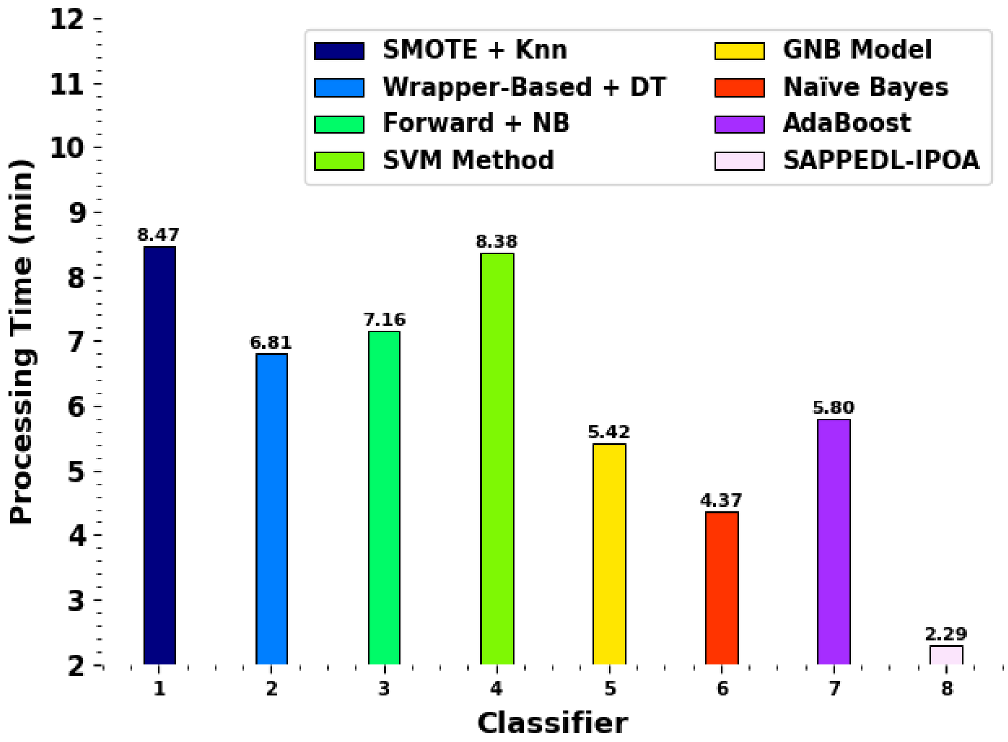

Fig. 8. PT outcome of SAPPEDL-IPOA model with existing classifiers

CONCLUSION

In this manuscript, we offer a SAPPEDL-IPOA methodology. The main objective of the SAPPEDL-IPOA technique is to provide a precise and strong predictive framework for enhancing the monitoring of student performance assessments. At the primary stage, the SAPPEDL-IPOA model applies the min-max scaling approach for data preprocessing to ensure all features are scaled within a specific range. Next, an ensemble of DL methods namely BiLSTM, DDQN, and SAE models are utilized for the classification of student academic performances. Finally, the IPOA is used for hyperparameter tuning in order to optimize the parameters of the ensemble models to enhance classification performances. To validate the improvement in the solution offered by the SAPPEDL-IPOA algorithm, an extensive range of simulations are conducted and the outcomes are studied under several facets. The comparison investigation reported the improvement of the SAPPEDL-IPOA system under diverse metrics.

REFERENCES

1. Waheed, H., Hassan, S.U., Aljohani, N.R., Hardman, J., Alelyani, S. and Nawaz, R., 2020. Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human behavior*, *104*, p.106189.
2. Nabil, A., Seyam, M. and Abou-Elfetouh, A., 2021. Prediction of students' academic performance based on courses' grades using deep neural networks. *IEEE Access*, *9*, pp.140731-140746.
3. Albreiki, B., Zaki, N. and Alashwal, H., 2021. A systematic literature review of student'performance prediction using machine learning techniques. *Education Sciences*, *11*(9), p.552.
4. Yağcı, M., 2022. Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1), p.11.
5. H. Alamri, L., S. Almuslim, R., S. Alotibi, M., K. Alkadi, D., Ullah Khan, I. and Aslam, N., 2020, December. Predicting student academic performance using support vector machine and random forest. In *Proceedings of the 2020 3rd International Conference on Education Technology Management* (pp. 100-107).
6. Rastrollo-Guerrero, J.L., Gómez-Pulido, J.A. and Durán-Domínguez, A., 2020. Analyzing and predicting students' performance by means of machine learning: A review. *Applied sciences*, *10*(3), p.1042.
7. Tsiakmaki, M., Kostopoulos, G., Kotsiantis, S. and Ragos, O., 2020. Transfer learning from deep neural networks for predicting student performance. *Applied Sciences*, *10*(6), p.2145.

8. Francis, B.K. and Babu, S.S., 2019. Predicting academic performance of students using a hybrid data mining approach. *Journal of medical systems*, *43*(6), p.162.
9. Beckham, N.R., Akeh, L.J., Mitaart, G.N.P. and Moniaga, J.V., 2023. Determining factors that affect student performance using various machine learning methods. *Procedia Computer Science*, *216*, pp.597-603.
10. Ijaz, M., Batool, R. and Asim, S.M., 2022. Factors Influencing Educational Achievements of the Students: A case study of University of Peshawar, Pakistan. *Full Length Article*, *6*(1), pp.48-8.
11. Zhao, S., Zhou, D., Wang, H., Chen, D. and Yu, L., 2025. Enhancing Student Academic Success Prediction Through Ensemble Learning and Image-Based Behavioral Data Transformation. *Applied Sciences*, *15*(3), p.1231.
12. Abdasalam, M., Alzubi, A. and Iyiola, K., 2024. Student grade prediction for effective learning approaches using the optimized ensemble deep neural network. *Education and Information Technologies*, pp.1-25.
13. Nayani, S. and P, S.R., 2025. Combination of deep learning models for student's performance prediction with a development of entropy weighted rough set feature mining. *Cybernetics and Systems*, *56*(2), pp.170-212.
14. Al-Ameri, A., Al-Shammari, W., Castiglione, A., Nappi, M., Pero, C. and Umer, M., 2024. Student Academic Success Prediction Using Learning Management Multimedia Data With Convoluted Features and Ensemble Model. *ACM Journal of Data and Information Quality*.
15. Zhang, X., Zhang, Y., Chen, A.L., Yu, M. and Zhang, L., 2025. Optimizing multi label student performance prediction with GNN-TINet: A contextual multidimensional deep learning framework. *PloS one*, *20*(1), p.e0314823.
16. Kaur, H., Kaur, T., Bhardwaj, V. and Kumar, M., 2024. An ensemble deep learning model for classification of students as weak and strong learners via multiparametric analysis. *Discover Applied Sciences*, *6*(11), p.595.
17. Balachandar, V. and Venkatesh, K., 2025. A multi-dimensional student performance prediction model (MSPP): An advanced framework for accurate academic classification and analysis. *MethodsX*, *14*, p.103148.
18. Sharma, S.K., 2024. Enhancing Stock Portfolio Selection with Trapezoidal Bipolar Fuzzy VIKOR Technique: A Multicriteria Decision-Making Approach
19. Najwa, S., Winarni, S., Pravitasari, A.A. and Hidayat, Y., 2025. Emotion classification in social media posts related to telecommunication services using bidirectional LSTM. *Commun. Math. Biol. Neurosci.*, *2025*, pp.Article-ID
20. Liu, H., Shen, Y., Zhou, W., Zou, Y., Zhou, C. and He, S., 2024. Adaptive speed planning for unmanned vehicle based on deep reinforcement learning. *arXiv preprint arXiv:2404.17379*.
21. Esmaeiloghli, S., Lima, A. and Sadeghi, B., 2024. Lithium exploration targeting through robust variable selection and deep anomaly detection: An integrated application of sparse principal component analysis and stacked autoencoders. *Geochemistry*, p.126111.
22. Qiu, S., Dai, J. and Zhao, D., 2024. Path planning of an unmanned aerial vehicle based on a multi-strategy improved Pelican optimization algorithm. *Biomimetics*, *9*(10), p.647.
23. https://www.kaggle.com/datasets/rabieelkharoua/students-performance-dataset
24. Bujang, S.D.A., Selamat, A., Krejcar, O., Mohamed, F., Cheng, L.K., Chiu, P.C. and Fujita, H., 2022. Imbalanced classification methods for student grade prediction: a systematic literature review. *IEEE Access*, *11*, pp.1970-1989.
25. Alhazmi, E. and Sheneamer, A., 2023. Early predicting of students performance in higher education. *IEEE Access*, *11*, pp.27579-27589.
26. Pek, R.Z., Özyer, S.T., Elhage, T., Özyer, T. and Alhajj, R., 2022. The role of machine learning in identifying students at-risk and minimizing failure. *IEEE Access*, *11*, pp.1224-1243.