Meta-Model Classification Based on the Naïve Bias Technique Auto-Regulated via Novel Metaheuristic Methods to Define Optimal Attributes of Student Performance

Zhen Ren[1], Mingmin He[2]
General Education Department, Sichuan University Jinjiang College, Meishan 620860, Sichuan, China

***Abstract*—Accurately assessing and predicting student performance is critical in today's educational environment. Schools are dependent on evaluating students' skills, forecasting their grades, and providing customized instruction to improve their academic performance. Early intervention is essential for pinpointing areas in need of development. By predicting students' futures in particular subjects, data mining, a potent technique for revealing hidden patterns within large datasets, helps lower failure rates. These methods are combined in the field of educational data mining, which focuses on the analysis of data from educators and students with the aim of raising academic achievement. In this study, the Naive Bayes classification (NBC) model is given the main responsibility for predicting student performance. However, two cutting-edge optimization strategies, Alibaba and the Forty Thieves (AFT) and Leader Harris Hawk's optimization (LHHO), have been used to improve the model's accuracy. The study's findings show that the NBC+AFT model performs more accurately than the other models. Accuracy, Precision, Recall, and F1-Score all display impressive performance metrics for a superior model, with values of 0.891, 0.9, 0.89, and 0.89, respectively. These metrics outperform those of competing models, highlighting how successful this strategy is. Because of the NBC+AFT model's strong performance, educational institutions are getting closer to a time when they will be able to predict students' success more precisely and help them along the way, making everyone's academic journey more promising and brighter.**

Keywords*—Student performance; machine learning; classification; Naive Bayes Classification; Alibaba and the forty thieves; Leader Harris Hawk's Optimization

I. Introduction

Educational data mining is a powerful approach that utilizes data mining techniques to analyze vast amounts of data stored by educational institutions. These data repositories contain a wealth of information, encompassing personal and academic details of students and faculty, syllabi, question papers, circulars, and more. Various educational institutions, both universities and independent organizations, have adopted educational data mining strategies to enhance the academic experiences of their students and faculty [1], [2], [3]. These strategies are seamlessly integrated into their systems to align with their extensive databases. A few examples of educational data mining applications include:

1) Student performance prediction: One of the most critical aspects for educational institutions is assessing student performance. Previous academic records can serve as a basis for predicting student success. This analysis can unveil the relationships between students' abilities and interests and their academic achievements, enabling teachers to offer tailored support to those who need it the most.

2) Teacher evaluation: The effectiveness of teachers is often measured by their students' performance, feedback, and other relevant factors. Analyzing these data helps institutions enhance the quality of instruction and support their teaching staff better.

3) Question paper analysis: Evaluation of question papers can determine their level of difficulty, aiding institutions in standardizing scores across multiple sessions of examinations.

Predicting student performance is a complex challenge, akin to having a master key that opens doors to addressing underperformance by foreseeing a student's academic trajectory. This predictive ability empowers educators and decision-makers to intervene promptly and provide the necessary support to ensure every student's academic success [4]. Moreover, it extends beyond the classroom, offering insights into a student's final exam results by considering various variables, including quiz scores, homework completion, and project achievements. These holistic assessments provide a comprehensive picture of a student's academic proficiency [5], [6].

In the realm of education, machine learning algorithms have demonstrated remarkable versatility in tackling various challenges, such as classification, web mining, clustering, association rules, and deep learning. Researchers continuously explore advanced algorithms, such as clustering and classification, to develop highly accurate educational models due to the complexity of educational data. These models hold the potential to enhance the overall educational experience of students significantly.

The application of machine learning algorithms in education has yielded positive outcomes across various domains, including classification problems, clustering, association rules, web mining, and deep learning [7], [8], [9]. Researchers in the field are actively exploring advanced algorithms like clustering and classification to build more precise educational models [10], [11], [12]. Notable examples include using machine learning techniques to correlate predictor factors with e-learning system usability, predicting students' grades, forecasting PISA test scores, and predicting adult learners' decisions to continue ESOL courses. While many prediction techniques, including regression, density estimation, and classification, are well-established, modern data science emphasizes trust and a comprehensive understanding of prediction models. Post-hoc interpretability approaches, like Local Interpretable Model-agnostic Explanations (LIME), are gaining popularity as they provide explanations for predictions made by trained black-box models, ensuring stakeholders can comprehend and rely on the insights from complex models.

In today's educational landscape, the development of robust machine-learning tools to assist educators in making well-informed decisions is not a luxury but a necessity. These tools reduce the risk of student failure, ultimately leading to improved educational outcomes. The primary objective of projects in this domain is to create dependable models for predicting student grades [13], [14]. Datasets encompassing a wide range of student performance-related factors, including personal information, educational background, and personal details, provide a comprehensive understanding of each student's situation [15], [16], [17]. By harnessing the power of machine learning, data mining, and predictive analytics, education is evolving. These technologies equip decision-makers with the tools and insights they need to identify and support underperforming students, resulting in enhanced educational outcomes for both students and institutions [18], [19].

Thammasiri et al. [20] developed a model for predicting low academic performance among freshmen. The combination of support vector machines with SMOTE yielded the greatest accuracy of 90.24%, solving class imbalance concerns. Ajay et al. [21] explored how the "CAT" social component predicts student achievement among Indians. They used four classifiers and discovered that the IB1 model had the best accuracy, at 82%. This characteristic defined people based on their social position, which had a direct influence on their educational performance. Edin Osmanbegovic et al. [22] developed a model that may predict student academic progress while addressing data dimensionality concerns. Although Naïve Bayes had the best accuracy (76.65%), it did not adequately solve the class imbalance issue. Dorina et al. [23] created a predicted model for student achievement utilizing a variety of categorization methods. While the MLP model was the most accurate at identifying successful students, it struggled to handle high-dimensional data and class imbalances. Carlos employed machine learning to develop a student failure prediction model, which achieved 92.7% accuracy using the ICRM classifier. However, due to differences in student characteristics, their study did not include testing at various educational levels.

This study is dedicated to the critical task of predicting student performance in G3 through an innovative machine-learning approach. The primary objective of this research is to optimize the performance of the Naive Bayes classification (NBC), a task made challenging by the acquisition of experimental data. The heart of this project lies in meticulous parameter optimization, which is key to enhancing the NBC model's effectiveness. To address this optimization challenge, employ a synergistic combination of two powerful algorithms: Alibaba and the Forty Thieves (AFT) and Leader Harris Hawk's optimization (LHHO). This harmonious integration of algorithms creates a cascade effect, resulting in a highly advantageous approach within the field of infrastructure, thereby elevating the intricacies involved in predicting student performance. By utilizing this novel approach, the aim extends beyond merely predicting student performance accurately. Seeks to enhance the overall effectiveness of the NBC model. Through meticulous parameter optimization and the utilization of cutting-edge algorithms, this research endeavours to provide valuable insights and solutions to the challenges faced in the education sector. This approach has the potential to revolutionize the way student performance is understood and supported on their academic journeys. Presents a promising avenue for improving the accuracy of student performance predictions and ultimately enhancing the quality of educational support. By taking these actions, it is aimed to make a positive impact on the academic prospects of all students, fostering a brighter and more promising educational future. Additionally, it strives to equip educators and decision-makers with the necessary resources to intervene effectively and guarantee the success of each student. To address the missing research summary or structure for the rest of the paper at the end of the 'Introduction' section, consider adding a brief paragraph that outlines the key components or sections to be covered in the upcoming sections of the paper.

In the following sections, a comprehensive analysis of the proposed hybrid method will delve into student performance prediction. Section I will present the experimental methodology, including the student performance data used for testing. In addition, Section II will provide an in-depth overview of the theoretical foundations, detailing the NBC coupled with AFT and LHHO. Results and comparisons with benchmark methods will be discussed in Section III, and Section IV will conclude with insights and implications drawn from the findings.

II. MATERIALS AND METHODOLOGY

A. Data Gathering

This section explores the application of a Naive Bayes classification (NBC) machine learning model to predict student performance based on various predictor variables. The table presents correlation coefficients that reveal the strength and direction of relationships between each predictor variable and the crucial student performance variable, G3, representing students' final grades. These correlation coefficients serve as valuable tools for understanding the multifaceted factors influencing student performance in an educational context. These three characteristics were chosen as model outputs (dependent variables), along with the number of absences from school. They were then split into four categories based on their grades: 0–12 = poor; 12–14 = acceptable; 14–16 = good; and 16–20 = excellent. Fig. 1 shows that students' ages exhibit a negative correlation with G3, indicating that older students tends to attain lower final grades. This can be attributed to increased responsibilities and distractions accompanying age. Parental education levels of both mothers (Medu and Fedu) and fathers (Fedu) show positive correlations with student performance, suggesting that students with parents boasting higher education levels tend to achieve superior grades. Family support (famsup) and school support (schoolsup) do not reveal a significant correlation with G3, while schoolsup (school support) presents a slightly negative correlation of -0.082. Aspirations for higher education (higher) have a robust positive correlation of 0.182, highlighting the importance of nurturing academic ambitions among students. Internet access (0.098) is positively correlated with student performance, emphasizing the role of technology and online resources in augmenting learning and research opportunities. Study time (0.098) is positively correlated with G3, indicating that students who dedicate more time to studying are more likely to secure superior final grades. Previous failures (-0.360) display a pronounced negative correlation with G3, illustrating that students with fewer past failures perform substantially better, emphasizing the urgency of addressing academic setbacks promptly. These correlation coefficients provide educators, policymakers, and parents with valuable insights into the multifaceted factors influencing student performance, enabling them to tailor interventions and strategies to support students in achieving enhanced academic outcomes [24], [25]. The potential of machine learning models, specifically NBC, in identifying these pivotal relationships and guiding evidence-based decision-making within the education sector is highlighted.

B. Naive Bayes classification (NBC)

The Naive Bayes classification (NBC), a probabilistic type, employs Bayes' theorem and assumes robust feature independence. Its key strength lies in its straightforward design, obviating the need for intricate iterative parameter estimation techniques. Additionally, it has been noted by Das et al. [26] that the NB classifier is resilient to noise and irrelevant attributes. The NB classifier is based on the following equation:

$$y = \underset{y_i=\{landslide,non-landslide\}}{\arg\max} P(y_i) \prod_{i=1}^{14} P(\frac{x_i}{y_i}) \quad (1)$$

where, $P(y_i)$is the prior probability of y_i, $P(\frac{x_i}{y_i})$is the posterior probability, and it can be calculated by:

$$P\left(\frac{x_i}{y_i}\right) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}} \quad (2)$$

where, μ is the mean and σ is the standard deviation of x_i.The flowchart of the NBC is shown in the Fig. 2.

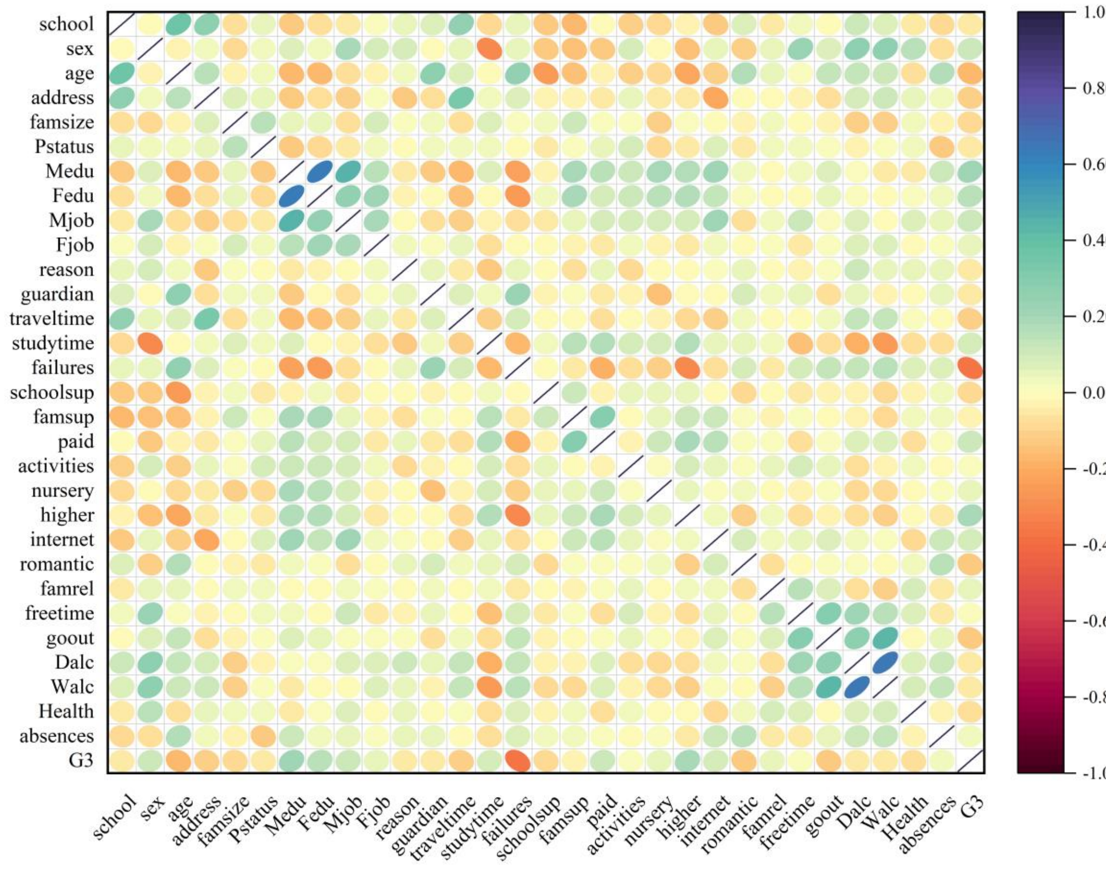

Fig. 1. Correlation matrix for the input and output variables.

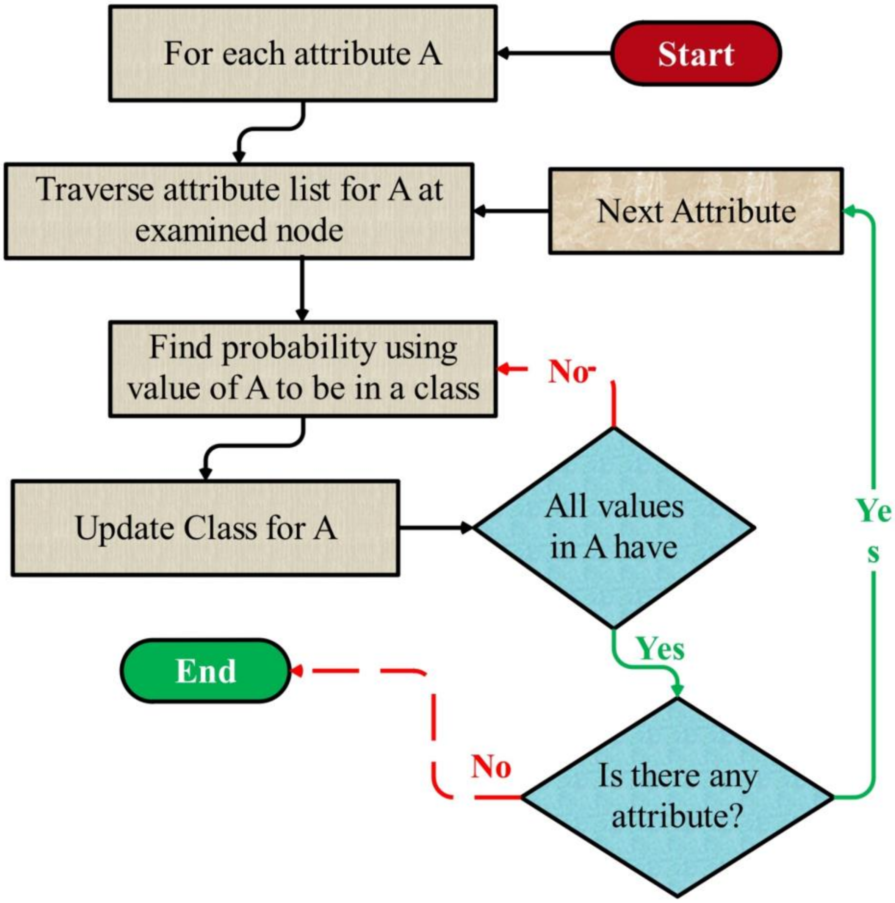

Fig. 2. The flowchart of the NBC.

C. Alibaba and the Forty Thieves (AFT)

The present study provides an explanation of the mathematical structure underlying the fundamental AFT algorithm, which is extensively detailed. Three separate states that are included in the framework can be looked at and defined as follows:

State one: Based on data gathered from a source, the thieves' chase after Ali Baba can be simulated using Eq. (3), which shows their relative positions [27], [28].

$$x_i^{t+1} = gbest^t + \left[\mathrm{Td}^{\mathrm{t}}(best_i^t - y_i^t)r_1 + \mathrm{Td}^{\mathrm{t}}\left(y_i^t - m_{a(i)}^t\right)r_2\right]sgn(rand - 0.5), \;\; p \geq 0.5, \;\; q > P_{p^t} \qquad (3)$$

x_i^{t+1} denotes the position of the $i-th$ thief at the next time step $(t+1)$.

$m_{a(i)}^t$ represents the level of Marjaneh's wit used to disguise thief i, at time t.

$best_i^t$ represents the best position achieved by thief i up to the current time step (t).

$gbest^t$ refers to the best global position achieved by any thief up to the current time step (t).

r_1, r_2, $rand$, p, and q are randomly generated values that fall within the range of $[0,1]$.

$p \geq 0.5$ indicates either a value of 0 or 1.

y_i^t depicts the position of Ali Baba concerning thief i, at time t.

a is defined by using Eq. (6).

$sgn(rand - 0.5)$ take on a value of either -1 or 1.

Td^{t} represents the tracking distance of the thieves, which is defined by Eq. (4).

P_{p^t} represents the potential perceptual ability of the thieves to detect Ali Baba, as defined in Eq. (5).

$$\mathrm{Td}^{\mathrm{t}} = \tau_0 e^{-\tau_1 (\frac{t}{T})_1^{\tau}} \qquad (4)$$

τ_0 $(\tau_0 = 1)$ serves as an initial estimate for the tracking distance.

τ_1 $(\tau_1 = 2)$ is employed to regulate the balance between exploitation and exploration.

t and T, respectively, refer to the current and maximum iteration values.

$$P_{p^t} = \lambda_0 \log(\lambda_1 (\frac{t}{T})^{\lambda_0} \quad (5)$$

λ_0 ($\lambda_0 = 1$) represents the final estimate of the probability that the thieves will succeed in achieving their goal following the search.

λ_1 ($\lambda_1 = 1$) stands for a constant that controls the ratio of exploration to exploitation.

$$a = [(n-1).rand(n,1)] \quad (6)$$

The result of creating a sequence of random numbers between 0 and 1 is $\boldsymbol{U}$ and n, 1.

$$m_{a(i)}^t = \begin{cases} x_i^t & if\ f(x_i^t) \geq f(m_{a(i)}^t) \\ m_{a(i)}^t & if\ f(x_i^t) < f(m_{a(i)}^t) \end{cases} \quad (7)$$

$f(0)$ represents the score or value of the fitness function.

State two: When the thieves realize they have been tricked, they might start venturing into unknown and unexpected areas.

$$x_i^{t+1} = Td^t[(u_j - l_j)r + l_j]; p \geq 0.5\ , q \leq P_{p^t} \quad (8)$$

The boundaries of the search space for dimension j are represented byu_j (the upper bound) and l_j (the lower bound).

r is a random variable created within the range of [0, 1].

State three: The thieves may investigate additional search positions outside of those obtained by applying equations in order to enhance the AFT algorithm's exploration and exploitation components. The following scenario can be mathematically expressed as Eq. (9):

$$x_i^{t+1} = gbest^t - [\mathrm{Td^t}(best_i^t - y_i^t)r_1 + \mathrm{Td^t}(y_i^t - m_{a(i)}^t)r_2]sgn(rand - 0.5) \quad (9)$$

Algorithm 1 provides an exact and succinct presentation of the iterative pseudo-code steps of the basic AFT algorithm.

Algorithm 1: AFT algorithm
Define and begin the control parameters.
Begin and evaluate the initial, best, and global positions of all thieves
Begin Marjane's wit level concerning all thieves
Set $t \leftarrow 1$
While ($t \leq T$) do
Update the parameter P_{p^t} using Eq. (5).
for each thief, do
if ($p \geq 0.5$) then
if ($q \geq P_{p^t}$) then
Update the thieves' position using Eq. (4).
else
Update the thieves' position using Eq. (8).
end if
else
Update the thieves' position using Eq. (9).
end if
end for
Update the new, best, and global positions of all thieves
Update Marjane's wit plans using Eq. (7).
$t = t + 1$
end while
Return the best global solution

D. Leader Harris Hawk's Optimization (LHHO)

The algorithm known as LHHO was developed using the exploratory behavior of the Harris hawk as a model. Owing to its equal chance q perching strategy, the original Harris Hawks Optimisation (HHO) algorithm has a finite exploration capacity. If q is greater than or equal to 0.5, then hawks will randomly choose a tall tree to perch on; if q is less than 0.5, then they will base their perching decisions on the locations of other family members [29]. This is in accordance with the HHO algorithm. However, this limitation can be overcome by assigning a perch probability to each hawk [30].

During the exploration phase ($|E| \geq 1$) a concept called adaptive perch probability (h_{ap}^i) can be introduced for the ith hawk. This probability value is determined by the fitness value of the current hawk with a position vector X_i denoted as $f(X_i)$, as well as the fitness values of the best-performing hawk with the position vector X_{prey}, denoted as $f(X_{prey})$, and the worst-performing hawk with the position vector X_{worst}, denoted as $f(X_{worst})$. By taking these factors into account, the adaptive perch probability (h_{ap}^i) can be formulated as:

$$h_{ap}^i = \frac{|f(X_i) - f(X_{prey})|}{|f(X_{worst}) - f(X_{prey})|} \quad , \qquad i = 1,2,3, \ldots , N \quad (10)$$

Then, the exploration phase can be modeled as:

$$X_i(new) = \begin{cases} X_{round}(t) - d_1|X_{round}(t) - 2d_2X_i(t)| & q > h_a^i \\ (X_{prey}(t) - X_w(t)) - d_3(LB + d_4(UB - LB)) & q < \end{cases} \quad (11)$$

The model incorporates $X_w(t)$, which reflects the population's average position vector during the exploration phase of N hawks. In contrast, during the exploitation phase ($|E| < 1$), four offensive techniques that are similar to those used in HHO are employed by the model.

- Soft besiege ($r \geq 0.5\ and\ |E| \geq 0.5$)

$$X_i(new) = X_{prey}(t) - E|JX_{prey}(t) - X_i(t)| \quad (12)$$

Where J is the jump strength as of $J = 2(1 - r_5)$

- Hard besiege ($r \geq 0.5\ and\ |E| < 0.5$)

$$X_i(new) = X_{prey}(t) - E|X_{prey}(t) - X_i(t)| \quad (13)$$

- Soft besiege with progressive rapid dives ($r < 0.5\ and\ |E| \geq 0.5$)

$$X_i(new) = \begin{cases} Y_i & if\ f(Y_i) < f(X_i(t)) \\ Z_i & if\ f(Z_i) < f(X_i(t)) \end{cases} \quad (14)$$

The Yi and Zi can be calculated using Eq. $Y_i = X_{prey}(t) - E\left|JX_{prey}(t) - X_i(t)\right|$ and Eq. $Z_i = Y_i + S \times LF(D)$, respectively.

- Hard besiege with progressive rapid dives ($r < 0.5$ and $|E| < 0.5$)

$$X_i(new) = \begin{cases} Y_i & if\ f(Y_i) < f(X_i(t)) \\ Z_i & if\ f(Z_i) < f(X_i(t)) \end{cases} \tag{15}$$

The equations used to calculate Y_i and Z_i are as follows: $Z_i = Y_i + S \times LF(D)$, respectively. It can be observed that the escape energy $|E|$ remains below 1 after 50% of the maximum iterations, indicating that the HHO algorithm only exploits solutions after this point. This restricted investigation increases the likelihood of discovering less-than-optimal solutions and becoming stuck in a local minimum. To help explore the end, a leader-based mutation-selection method is proposed as an addition to the HHO algorithm.

Here, to put the leader-based mutation-selection strategy into practice, first determine the position vectors of the best, second-best, and third-best hawks, denoted as X_{best}^t, X_{best-1}^t, and X_{best-2}^t, respectively. These position vectors are determined based on the fitness function value of the new position vector $X(new)$ among the N individual hawks. The study can then define the new mutation position vector for the i_th hawk, denoted as $X_i(mut)$, as follows:

$$\begin{aligned} X_i(mut) = X_i(new) + 2 * \left(1 - \frac{t}{t_{max}}\right) \\ * (2 * rand - 1)\left(2 * X_{best}^t \right. \\ \left. - \left(X_{best-1}^t + X_{best-2}^t\right)\right) \\ + (2 * rand - 1)\left(X_{best}^t - X_i(new)\right) \end{aligned} \tag{16}$$

where a rand is a random number in the range $(0, 1)$. Then, the position vector for the next generation $X_i(t+1)$, can be obtained by the selection process described in Eq. (17). Similarly, the X_{prey} is updated using Eq. (17). The flowchart of the LHHO is shown in Fig. 3.

$$X_i(t+1) = \begin{cases} X_i(mut) & f\left(X_i(mut)\right) < f\left(X_i(new)\right) \\ X_i(new) & f\left(X_i(mut)\right) > f\left(X_i(new)\right) \end{cases} \tag{17}$$

$$X_{prey} = \begin{cases} X_i(mut) & f\left(X_i(mut)\right) < f\left(X_{prey}\right) \\ X_i(new) & f\left(X_i(new)\right) > f\left(X_{prey}\right) \end{cases} \tag{18}$$

E. Performance Evaluation Methods

Numerous evaluation criteria are used to assess the classifiers' performance. The most popular criterion for assessing classification accuracy is PESTEL, which gauges a classifier's efficacy by looking at the proportion of correctly predicted samples, as shown in the equation below. Two additional popular evaluation indices are precision and recall. The ratio of values with a positive class to those that are expected to be positive is known as recall. Conversely, precision, which can be defined as the following equations, is the likelihood that a positive prediction will come true. The f1-score, which is defined as follows, is a new value that can be produced by combining Precision and Recall.

Equations contain	Equation	Assessment	
TP means that the outcome was positive and in line with this prediction. **FP** means it was not what was expected, which was a negative result. **TN** indicates that the outcome was negative, as predicted. **FN** means that although the study was expecting a negative result, the outcome was positive.	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$	***Higher is desirable***	(19)
	$Precision = \frac{TP}{TP + FP}$	***Higher is desirable***	(20)
	$Recall = TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$	***Higher is desirable***	(21)
	$F1\ score = \frac{2 \times Recall \times Precision}{Recall + Precision}$	***Higher is desirable***	(22)

III. Results and Discussion

The outcomes of the models that were given are shown in Table I. Each model was assessed using a variety of index values, such as accuracy, precision, recall, and F1-score. The first model, NBC+AFT, demonstrated its ability to predict student performance with an accuracy of 0.891 accurately. Additionally, it showed a high degree of precision (0.9), indicating that it could correctly predict positive outcomes. The model demonstrated an F1-score of 0.89, signifying a balance between precision and recall, and a recall of 0.89, indicating its efficacy in identifying pertinent instances.

In comparison to the NBC+AFT model, the second model, NBC+LHH, performed marginally worse in terms of accuracy (0.881), precision (0.88), recall (0.88), and F1-score (0.88), but it was still very capable of making predictions. The accuracy of the third model, NBC, was 0.873, indicating that it was capable of producing accurate predictions. Additionally, it displayed F1-score, precision, and recall values of 0.87, demonstrating a balanced performance in terms of true positive predictions and the model's capacity to find pertinent instances. Overall, these models' results show how well optimization methods like AFT and LHHO can be used to improve efficiency. The NBC+AFT model slightly outperformed the others, demonstrating its potential for accurate student performance prediction, even though all models achieved high accuracy and showed a trade-off between precision and recall.

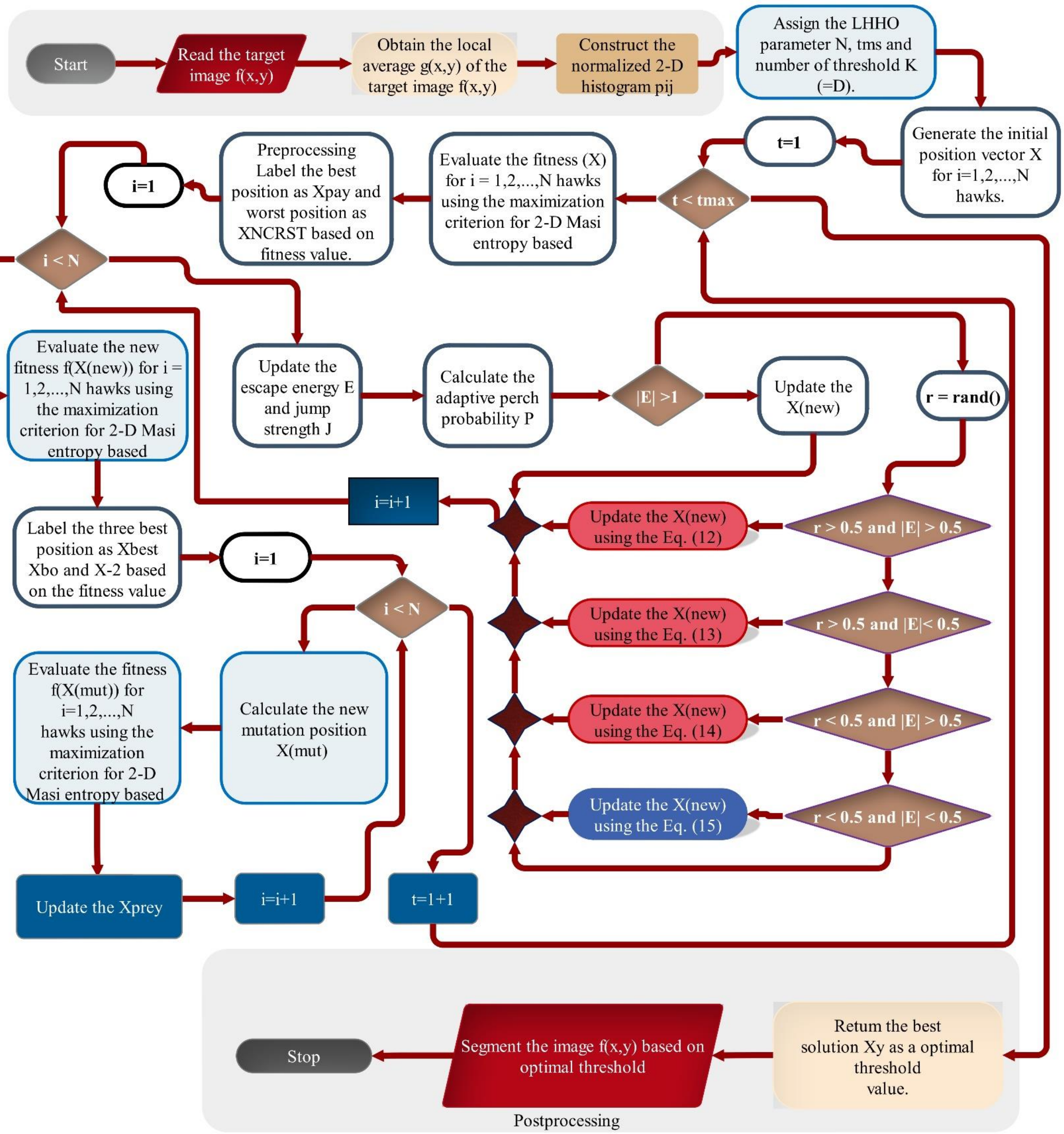

Fig. 3. Flowchart of LHHA.

The performance evaluation indices for the developed models NBC+AFT, NBC+LHH, and NBC are shown in this table. To give a complete picture of how well the models predict student performance, they are evaluated across a range of performance grades, from Excellent to Poor.

TABLE I. RESULT OF PRESENTED MODELS

Model	Index values			
	Accuracy	*Precision*	*Recall*	*F1 _score*
NBC+AFT	0.891%	0.9%	0.89%	0.89%
NBC+LHH	0.881%	0.88%	0.88%	0.88%
NBC	0.873%	0.87%	0.87%	0.87%

NBC+AFT:

- Poor: With a precision of 0.93, the model's ability to predict poor grades is impressive and suggests a strong identification of students who perform poorly. Additionally, it has a high recall of 0.97, indicating that it is capable of identifying most underperformers. The F1-score in this category is 0.95, which indicates a very balanced performance.
- Acceptable: Within the Acceptable grade, the NBC+AFT model showcases a precision of 0.84, signifying its capability to identify students with acceptable performance correctly. However, the recall is 0.74, indicating that it might miss some of these

students. The F1-score is 0.79, reflecting a reasonable balance between precision and recall.

- Good: The model consistently maintains a precision of 0.75 in the good grade category, demonstrating its dependability in identifying students who are performing well. Additionally, its recall score is 0.92, indicating that it can identify the majority of high performers. The F1-score of 0.83 indicates that recall and precision are in a healthy balance.

- Excellent: The NBC+AFT model exhibits high precision (1) for the Excellent grade, indicating a strong ability to recognize students who perform excellently. With a recall of 0.62, the model appears to account for 62% of students who perform exceptionally well. With an F1-score of 0.77, recall and precision are fairly balanced.

NBC+LHH and NBC:

- In all grade categories, the NBC+LHH model performs similarly, with F1-scores, precision, and recall matching those of the NBC+AFT model. Precision and recall for the NBC model are marginally different from those of the NBC+AFT and NBC+LHH models. It continues to perform well, nevertheless, in recognizing students in various grade levels.

These assessment indices offer insightful information about how well the developed models performed, highlighting how well they predicted student performance across a range of grade levels. The decision between NBC+AFT, NBC+LHH, and NBC may be influenced by particular educational environments as well as the intended harmony between recall and precision for various grade levels.

The line symbol plot shown in Fig. 4, as illustrated in the Table II, presents a visual representation of the measured data compared to the predictions generated by three distinct models: NBC+AFT, NBC+LHHO, and NBC. The measured values represent the actual number of students falling into each performance category, while the model predictions indicate the estimated numbers for each category.

1) Poor performance:

- NBC+AFT (226): The NBC+AFT model predicts that 226 students will perform poorly.

- NBC+LHHO (226): The NBC+LHHO model, closely aligned with NBC+AFT, also estimates 226 students to have poor performance.

- NBC (222): The standard NBC model predicts 222 students to fall into this category.

2) Acceptable performance:

- NBC+AFT (46): The NBC+AFT model predicts that 46 students will perform at an acceptable level.

- NBC+LHHO (45): The NBC+LHHO model estimates 45 students to have acceptable performance.

- NBC (45): The standard NBC model concurs with NBC+LHHO, also predicting 45 students in this category.

3) Good performance:

- NBC+AFT (55): The NBC+AFT model predicts that 55 students will achieve a good performance level.

- NBC+LHHO (51): The NBC+LHHO model estimates 51 students to fall into this category.

- NBC (49): The standard NBC model predicts 49 students in this group.

TABLE II. EVALUATION INDEXES OF THE DEVELOPED MODELS' PERFORMANCE BASED ON GRADES

Model	**Grade**	**Index values**		
		Precision	***Recall***	***F1-score***
NBC+AFT	Excellent	1	0.62	0.77
	Good	0.75	0.92	0.83
	Acceptable	0.84	0.74	0.79
	Poor	0.93	0.97	0.95
NBC+LHH	Excellent	0.93	0.65	0.76
	Good	0.73	0.85	0.78
	Acceptable	0.83	0.73	0.78
	Poor	0.93	0.97	0.95
NBC	Excellent	0.85	0.72	0.78
	Good	0.74	0.82	0.78
	Acceptable	0.78	0.73	0.75
	Poor	0.94	0.95	0.94

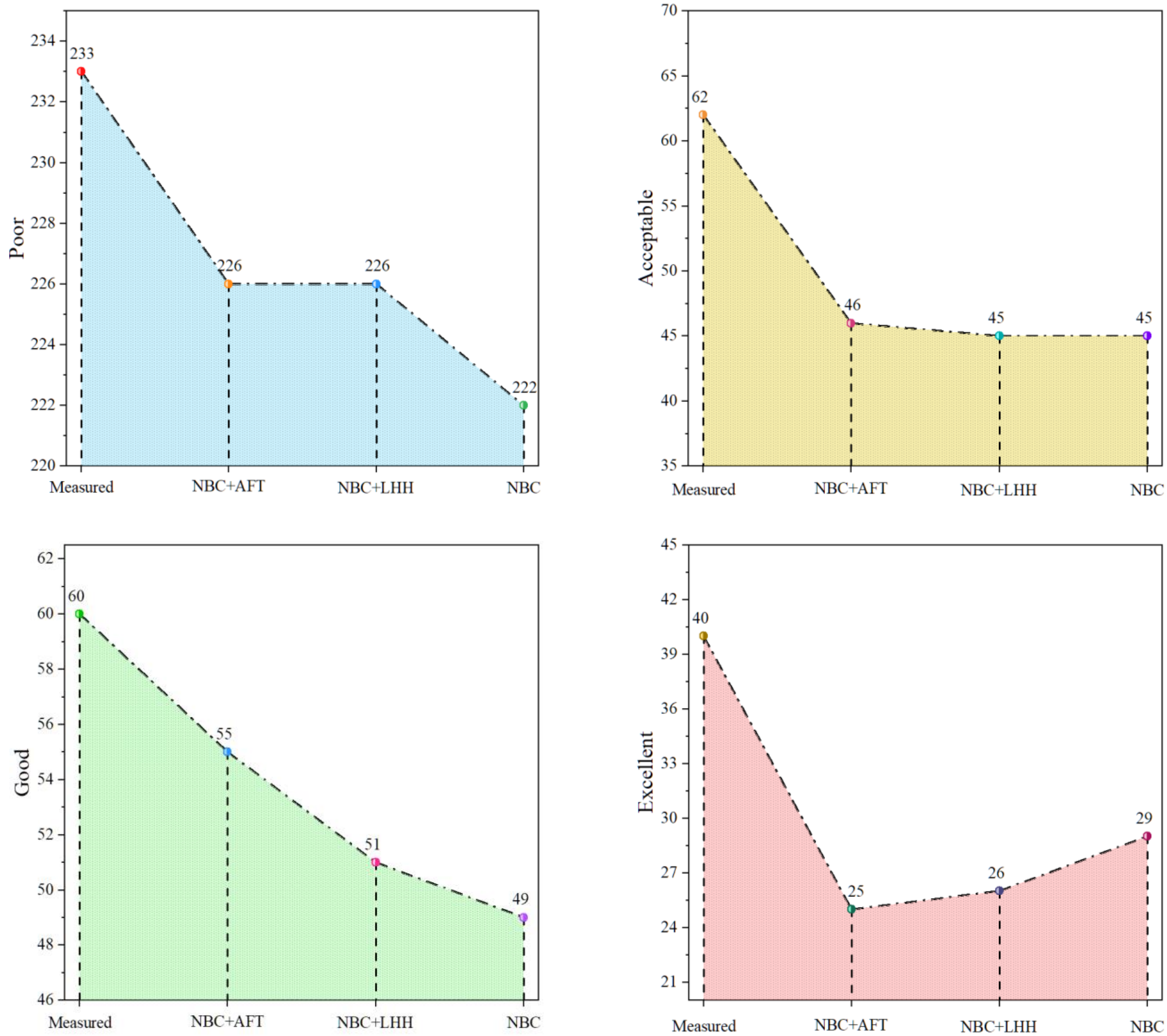

Fig. 4. Line-symbol plot for the classification accuracy of Meta-models.

4) Excellent performance:

- NBC+AFT (25): The NBC+AFT model predicts that 25 students will attain an excellent level of performance.
- NBC+LHHO (26): The NBC+LHHO model closely aligns with NBC+AFT, estimating 26 students to achieve excellence.
- NBC (29): The standard NBC model forecasts that 29 students will reach an excellent performance level.

These line symbol plot values illustrate how well the models align with the actual measured data for different performance categories. The variations in the predictions of each model offer insights into their individual capabilities and accuracy in identifying student performance levels. In this context, NBC+AFT, NBC+LHHO, and NBC exhibit similarities and differences in their predictions, highlighting the strengths and limitations of each approach in assessing student performance.

Three confusion matrices that show how the NBC, NBC+AFT, and NBC+LHHO models relate to the observed and predicted classes are shown in Fig. 5. The observed classes are plotted on the horizontal axis, and the predicted classes are plotted on the vertical axis. Interestingly, these matrices' diagonal cells—which match the precise predictions—have higher values than their off-diagonal cells.

- NBC+AFT: Specifically, the NBC+AFT hybrid model shows an impressive capacity to predict most observation classes accurately. To provide more context, let's look at the NBC+AFT plot. Out of the 40 students in the excellent class, the NBC+AFT hybrid model correctly predicts 25 of them to be in the same excellent category. The remaining three students are incorrectly assigned to the poor class, 1 to the acceptable class, and 11 to the good class.

- NBC+LHHO: In the NBC+LHHO storyline, the impoverished class comprises 233 pupils. In this bad class, the NBC+LHHO hybrid model predicts 226 students with skill; only four students are incorrectly placed in the acceptable class, and only three students are incorrectly placed in the good class.
- NBC: On the other hand, the NBC story revolves around 60 pupils in the superior class. Of these, 49 are correctly predicted by the NBC hybrid model to be in the good category; one student is mistakenly placed in the poor class, five in the acceptable class, and five in the excellent class.

These results underscore the efficacy of the NBC+AFT hybrid model in accurately predicting student performance classes, with notably fewer misclassifications compared to the other models.

The convergence curve of hybrid models with 200 iterations is shown in Fig. 6. The accuracy parameter is represented by the vertical axis in this visualization, and the horizontal axis corresponds to the number of iterations. This graph's analysis reveals that the NBC+LHHO hybrid model, which records an accuracy of 0.76 and reaches its ideal iteration at number 126, is the model with the lowest accuracy. As an illustration of this, the green NBC+AFT hybrid model achieves the highest accuracy value of all the models, 0.79.9. This model performs better than the others in terms of accuracy, reaching its optimal iteration point at 128.

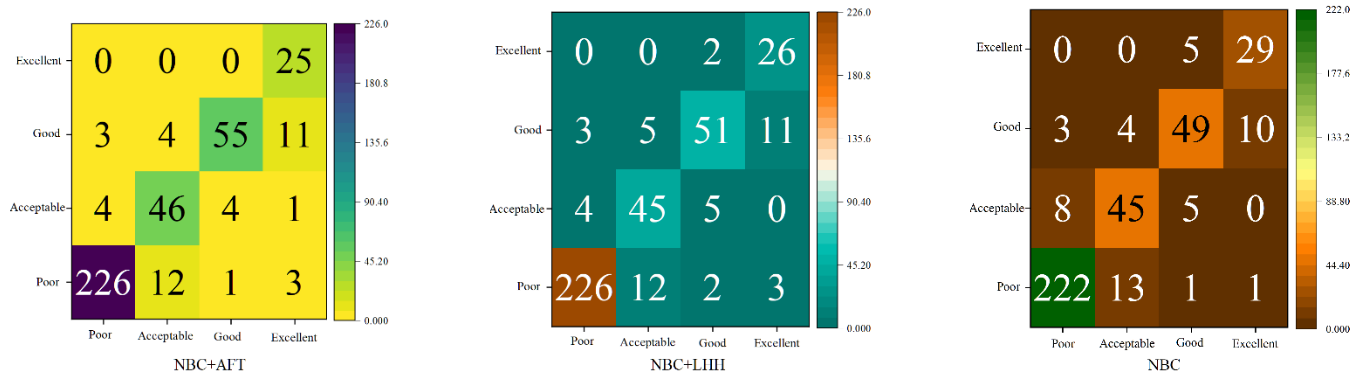

Fig. 5. Confusion matrix for each model's accuracy.

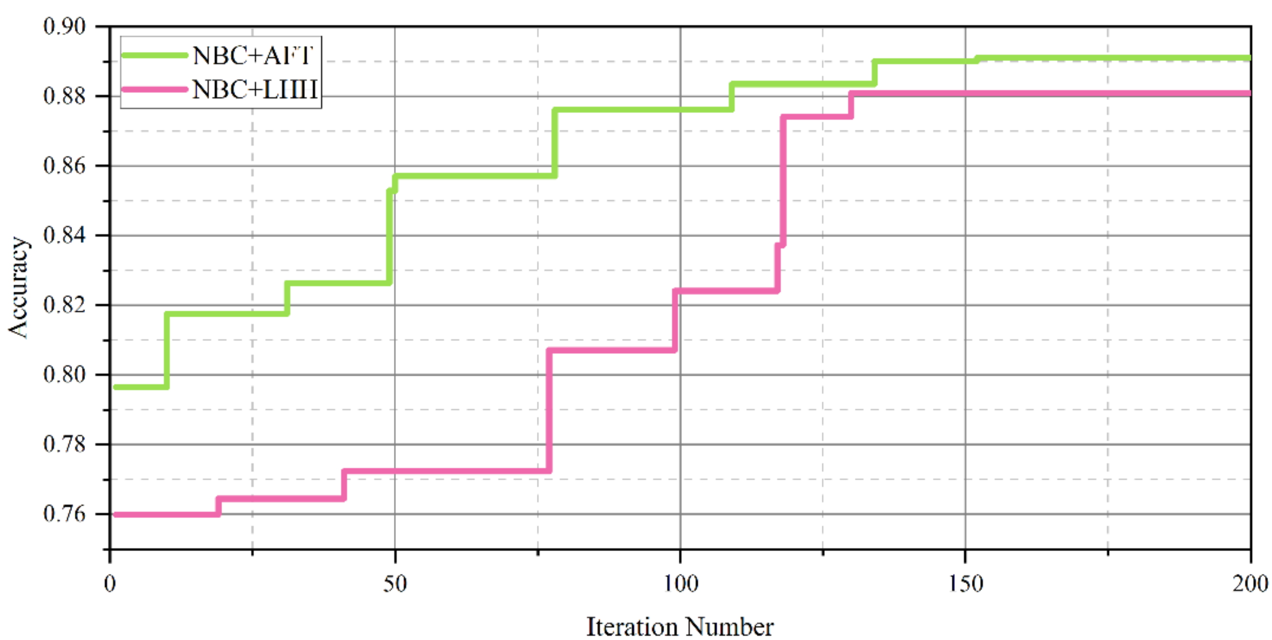

Fig. 6. Convergence curve of hybrid models.

IV. Conclusion

Forecasting student performance is still an important task in today's educational environment. Educational establishments are responsible for determining the skills of their students, projecting their academic performance, and making proactive efforts to enhance their future success. Predictive model accuracy and efficacy have significantly increased as a result of utilizing machine learning techniques, particularly the Naive Bayes classification (NBC), in conjunction with sophisticated optimization algorithms like Alibaba and the Forty Thieves (AFT) and Leader Harris Hawk's optimization (LHHO). The improvements are especially noticeable when looking at important assessment metrics like F1-Score, Accuracy, Precision, and Recall. Forecasting student performance is still an important task in today's educational environment. Educational establishments are responsible for determining the skills of their students, projecting their academic performance, and making proactive efforts to enhance their future success. Predictive model accuracy and efficacy have significantly increased as a result of utilizing machine learning techniques, particularly the Naive Bayes classification (NBC), in conjunction with sophisticated optimization algorithms like Alibaba and the Forty Thieves (AFT) and Leader Harris Hawk's optimization (LHHO). The improvements are especially noticeable when looking at important assessment metrics like F1-Score, Accuracy, Precision, and Recall. In this thorough analysis, the NBC+AFT hybrid model has proven to be the best performer, consistently outperforming other models. With its outstanding performance in terms of Accuracy, Precision, Recall, and F1-Score, this model is the best option for educational institutions committed to improving the prediction of student performance. It performs exceptionally well at predicting academic grades with the least amount of incorrect categorizations, which is an essential feature for making informed decisions. The importance of sophisticated machine learning models and optimization strategies in the field of predicting student performance is highlighted by this study. In particular, the NBC+AFT hybrid model provides educational institutions with an efficient way to assess and assist students according to their academic performance. These models have the potential to revolutionize academic guidance and support, improving student outcomes in a data-driven setting in the process. The future of education is expected to be shaped by sophisticated machine learning techniques that prioritize accuracy, precision, recall, and F1-Score as the volume and complexity of educational data continue to rise.

References

[1] 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, no. 3, pp. 637–655, 2023.

[2] Vijayalakshmi and K. Venkatachalapathy, "Comparison of predicting student's performance using machine learning algorithms," International Journal of Intelligent Systems and Applications, vol. 11, no. 12, p. 34, 2019.

[3] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," Smart Learning Environments, vol. 9, no. 1, p. 11, 2022.

[4] J. Xu, K. H. Moon, and M. Van Der Schaar, "A machine learning approach for tracking and predicting student performance in degree programs," IEEE J Sel Top Signal Process, vol. 11, no. 5, pp. 742–753, 2017.

[5] Acharya and D. Sinha, "Early prediction of students performance using machine learning techniques," Int J Comput Appl, vol. 107, no. 1, pp. 37–43, 2014.

[6] D. M. Ahmed, A. M. Abdulazeez, D. Q. Zeebaree, and F. Y. H. Ahmed, "Predicting university's students performance based on machine learning techniques," in 2021 IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS), IEEE, 2021, pp. 276–281.

[7] H. Pallathadka, A. Wenda, E. Ramirez-Asís, M. Asís-López, J. Flores-Albornoz, and K. Phasinam, "Classification and prediction of student performance data using various machine learning algorithms," Mater Today Proc, vol. 80, pp. 3782–3785, 2023.

[8] H. Turabieh, "Enhanced Binary Genetic Algorithm as a Feature Selection to Predict Student Performance," 2021.

[9] Asselman, M. Khaldi, and S. Aammou, "Enhancing the prediction of student performance based on the machine learning XGBoost algorithm," Interactive Learning Environments, vol. 31, no. 6, pp. 3360–3379, 2023.

[10] H. Altabrawee, O. A. J. Ali, and S. Q. Ajmi, "Predicting students' performance using machine learning techniques," JOURNAL OF UNIVERSITY OF BABYLON for pure and applied sciences, vol. 27, no. 1, pp. 194–205, 2019.

[11] E. S. Bhutto, I. F. Siddiqui, Q. A. Arain, and M. Anwar, "Predicting students' academic performance through supervised machine learning," in 2020 International Conference on Information Science and Communication Technology (ICISCT), IEEE, 2020, pp. 1–6.

[12] M. Koutina and K. L. Kermanidis, "Predicting postgraduate students' performance using machine learning techniques," in International Conference on Engineering Applications of Neural Networks, Springer, 2011, pp. 159–168.

[13] Y. A. Alsariera, Y. Baashar, G. Alkawsi, A. Mustafa, A. A. Alkahtani, and N. Ali, "Assessment and evaluation of different machine learning algorithms for predicting student performance," Comput Intell Neurosci, vol. 2022, 2022.

[14] S. S. Shreem, H. Turabieh, S. Al Azwari, and F. Baothman, "Enhanced binary genetic algorithm as a feature selection to predict student performance," Soft comput, vol. 26, no. 4, pp. 1811–1823, 2022.

[15] P. Cortez and A. M. G. Silva, "Using data mining to predict secondary school student performance," 2008.

[16] J. L. Rastrollo-Guerrero, J. A. Gómez-Pulido, and A. Durán-Domínguez, "Analyzing and predicting students' performance by means of machine learning: A review," Applied sciences, vol. 10, no. 3, p. 1042, 2020.

[17] P. Dabhade, R. Agarwal, K. P. Alameen, A. T. Fathima, R. Sridharan, and G. Gopakumar, "Educational data mining for predicting students' academic performance using machine learning algorithms," Mater Today Proc, vol. 47, pp. 5260–5267, 2021.

[18] S. Hashim, W. A. Awadh, and A. K. Hamoud, "Student performance prediction model based on supervised machine learning algorithms," in IOP Conference Series: Materials Science and Engineering, IOP Publishing, 2020, p. 32019.

[19] Masci, G. Johnes, and T. Agasisti, "Student and school performance across countries: A machine learning approach," Eur J Oper Res, vol. 269, no. 3, pp. 1072–1085, 2018.

[20] Thammasiri, D. Delen, P. Meesad, and N. Kasap, "A critical assessment of imbalanced class distribution problem: The case of predicting freshmen student attrition," Expert Syst Appl, vol. 41, no. 2, pp. 321–330, 2014.

[21] K. Pal and S. Pal, "Data mining techniques in EDM for predicting the performance of students," International Journal of Computer and Information Technology, vol. 2, no. 06, pp. 2279–2764, 2013.

[22] Osmanbegovic and M. Suljic, "Data mining approach for predicting student performance," Economic Review: Journal of Economics and Business, vol. 10, no. 1, pp. 3–12, 2012.

[23] Kabakchieva, "Student performance prediction by using data mining classification algorithms," International journal of computer science and management research, vol. 1, no. 4, pp. 686–690, 2012.

[24] S. B. Kotsiantis, “Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades,” Artif Intell Rev, vol. 37, pp. 331–344, 2012.

[25] M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, and S. Ali, “Using machine learning to predict student difficulties from learning session data,” Artif Intell Rev, vol. 52, pp. 381–407, 2019.

[26] Das, A. Stein, N. Kerle, and V. K. Dadhwal, “Landslide susceptibility mapping along road corridors in the Indian Himalayas using Bayesian logistic regression models,” Geomorphology, vol. 179, pp. 116–125, 2012.

[27] M. Braik, M. H. Ryalat, and H. Al-Zoubi, “A novel meta-heuristic algorithm for solving numerical optimization problems: Ali Baba and the forty thieves,” Neural Comput Appl, vol. 34, no. 1, pp. 409–455, 2022, doi: 10.1007/s00521-021-06392-x.

[28] P. Sharma, S. Thangavel, S. Raju, and B. R. Prusty, “Parameter Estimation of Solar PV Using Ali Baba and Forty Thieves Optimization Technique,” Math Probl Eng, vol. 2022, p. 5013146, 2022, doi: 10.1155/2022/5013146.

[29] Kumar, V. Sharma, and R. Naresh, “Leader Harris Hawks algorithm based optimal controller for automatic generation control in PV-hydro-wind integrated power network,” Electric Power Systems Research, vol. 214, p. 108924, 2023.

[30] M. K. Naik, R. Panda, A. Wunnava, B. Jena, and A. Abraham, “A leader Harris hawks optimization for 2-D Masi entropy-based multilevel image thresholding,” Multimed Tools Appl, pp. 1–41, 2021.