Exploring the Impact of Time Management Skills on Academic Achievement with an XGBC Model and Metaheuristic Algorithm

Songyang Li
Police Skills and Tactics Training Department, Criminal Investigation Police University of China,
Shenyang 110035, Liaoning, China

***Abstract*—Estimating a student's academic performance is a crucial aspect of learning preparation. In order to predict understudy academic performance, this consideration uses a few Machine Learning (ML) models and Time Administration Aptitudes data from the Time Structure Questionnaire (TSQ). While a number of other useful characteristics have been used to forecast academic achievement, TSQ findings, which directly evaluate students' time management skills, have never been included. This oversight is surprising, as time management skills likely play a significant role in academic success. Time administration may be an ability that may impact the student's academic accomplishment. The purpose of this research is to look at the connection between college students' academic success and their ability to manage their time well. The Extreme Gradient Boosting Classification (XGBC) model has been utilized in this study to forecast academic student performance. To enhance the prediction accuracy of the XGBC model, this study employed three optimizers: Giant Trevally Optimizer (GTO), Bald Eagle Search Optimization (BESO), and Seagull Optimization Algorithm (SOA). Impartial performance evaluators were employed in this study to assess the models' predictions, minimizing potential biases. The findings showcase the success of this approach in developing an accurate predictive model for student academic performance. Notably, the XGBE surpassed other models, achieving impressive accuracy and precision values of 0.920 and 0.923 during the training phase.**

***Keywords*—Student academic performance; time management; machine learning; extreme gradient boosting classification; metaheuristic algorithm**

I. Introduction

A. Background

Student academic execution could be an exceptionally vital perspective for colleges and schools since it speaks to how healthy they teach their understudies [1]. As early as possible, it is important to predict a student's academic success so that universities may take appropriate action [2]. For illustration, if an understudy is demonstrated to have a terrible review, the teacher can give extra fabric or a session for the understudy [3]. A few things have been conducted to anticipate student academic execution utilizing different highlights. Highlight choice is a critical viewpoint in making forecasts [4]. Numerous highlights have been utilized, extending from social, statistical, behavioral, individual, and academic information. Among those elements, the academic elements continue to be the most persuasive factor when choosing how to implement academically [5].

However, another consideration [6] appeared curious: that time administration aptitude is related to academic execution. A student with a tall review point normally tends to have solid time administration aptitude [7]. Nevertheless, few tests have been conducted utilizing time administration aptitudes, including anticipating understudy execution. In this way, a ponder must be approved out by attempting to use time administration ability information as a highlight to form forecasts [8]. ML points to forming machines that can do their jobs skillfully by utilizing clever computer programs [9], [10]. ML has the potential to shed light on many topics, including categorization problems. The process of dividing input vectors into a finite number of distinct, specified categories or classes is known as classification [11], [12].

B. Literature Review

A ponder in 2008 [13] employments understudy foundation (sexual orientation, age, family, etc.), understudy social exercises (week after week think about time, free time after school, extra−curricular exercises, etc.), and coursework result (to begin with a period review and moment period review) as highlights for doing classification. Five diverse calculations are utilized to decide the most excellent calculations. They are NB, SVM, NN, DT, and RF. As a result, 93% exactness is accomplished for twofold classification and 78.5% precision for $five$ −level classification. The most important aspect of categorization, it was discovered, is coursework. Utilizing the same dataset as [13], another study in 2018 [14] appears comparative comes about. Even Nevertheless, the most notable aspect of classifying coursework is its outcome. Another consideration [15] conducted in Jordanian utilized $three$ distinctive include categories: individual data (sex, family status, age, etc.), instructive data (tall school stream, tall school review, college sort, etc.), and geographic information (travel time and transportation sort). By utilizing NN, 97% precision can be obtained for $four$ −level classification. At the same time, DT, as it were, gets approximately 66% precision. Instructive data is the foremost important and noteworthy highlight within the classification. Focusing on scholastic variables, another thinks about [16] connecting straight regression (LR), numerous regression (MR), and NN to anticipate students' GPA. This ponders centered on scholastic components. As the result, 83% precision is gotten through NN.

Curious discoveries were made in study [17]. 43 diverse highlights have been created to anticipate understudy scholarly execution in arithmetic. They consider employments seven distinctive calculations: forward-thinking Auxiliary Condition Modeling ($pSEM$), Multilayer Perceptron (MLP) $NN, C5.0$ of DT, Calculated Relapse (LoR, Successful Negligible Optimization (SMO) of SVM, and RF. The accuracy may go up to 93.52%. This analysis has also made an effort to identify the most compelling features when classifying. The three most insightful points, therefore, have to do with managing time. These include the repetition of thinking through and preparing for assessments and examinations, the quantity of time dedicated to independent study, and the repetition of completing homework and clarifying mathematical problems [18].

Claessens et al. [18] describe time management practices as actions that lead to making effective use of time while carrying out certain goal-directed tasks. Since time management is a talent, it can be quantified. Using a Time Management Questionnaire or Time Structure Questionnaire (TSQ) is one method of measuring time management skills (TMS). Quill and Bond introduced TSQ in 1983 [19], including seventeen items about time management. A 1 to 7 scale with the labels "Yes, continuously" and "No, never" was used to rank each item. The answers to the preexisting items were included to get the total score. An individual's ability to manage their time improves with increasing score. As seen in the study [20], by deleting one item and adding ten more, TSQ improved upon Quill and Bond's 1988 efforts, making a total of 26 unique items available. According to studies [20] and [21], there was a sign that understudies with great TMS tend to have great scholastic execution as well. Other than that, it found that understudies who had a great TMS score essentially detailed more prominent work and life fulfilment [20]. Moreover, individuals with organized and Intentional time management are associated with high levels of confidence and depressive symptoms [22].

Malykh et al.'s study [23] complements existing literature by demonstrating that the format of non-symbolic comparison tasks significantly affects children's numerosity estimation, with homogeneous formats enhancing the congruency effect and heterogeneous formats reducing it. This aligns with previous research indicating that visual properties can either aid or hinder numerical processing, depending on the context. Their findings suggest that younger children, in particular, are prone to relying on visual cues, which can skew their numerical estimations. As children age, their ability to process numerical information independently of these cues improves, highlighting the importance of developmental considerations in educational assessments.

C. Objective

This study utilizes ML, specifically Extreme Gradient Boosting (XGB), to predict student performance based on their time management skills. In an effort to improve the single model's performance, three metaheuristic algorithms are employed: Giant Trevally Optimizer (GTO), Bald Eagle Search Optimization (BESO), and Seagull Optimization Algorithm (SOA). This study assessed the performance of ML models in evaluating student performance, ensuring fairness through performance evaluators like accuracy and precision. Pre-processing techniques helped create a clean and effective training dataset, while feature selection identified the most relevant input factors. This approach provides a novel overview of ML in student performance evaluation, potentially aiding institutions in optimizing their strategies and analyzing their students more effectively. However, it is important to acknowledge that even with these measures, the potential for bias in the data or chosen algorithms cannot be eliminated. This study makes several key contributions: it integrates TSQ data, which evaluates students' time management skills, into predictive models—a novel approach not previously utilized in predicting academic performance. The study also ensures fairness and minimizes potential biases through the use of impartial performance evaluators. Finally, the findings offer valuable insights that can help educational institutions improve their strategies and analyze student performance more effectively, potentially leading to more effective academic interventions and support systems. Acknowledging the potential for bias in data and algorithms, the study raises awareness about the limitations and ethical considerations of using ML for educational purposes. This awareness can lead to more cautious and responsible application of these technologies in educational psychology.

D. Research Organization

The introduction of this study is divided into four key sections: background, related work, objectives, and research organization. The subsequent structure of the paper is as follows: Section II provides detailed overviews of various machine learning techniques, including models and optimization algorithms, along with a brief description of the evaluation metrics used. Section III examines the dataset, highlighting the correlation between input and output variables and the feature selection processes. Section IV presents comparative results based on metric values to assess the performance of the models. Section V, titled "Discussion," discusses the study's limitations and potential directions for future research. Finally, Section VI summarizes the key findings and conclusions derived from the study.

II. METHODOLOGY

A. Extreme Gradient Boosting Classification (XGBC)

Comparable with angle boosting, XGBoost [24] combines a frail base classifier into a more grounded classifier. At each emphasis of the preparing handle, the remaining base classifier is utilized within the following classifier for optimizing the objective work. Assume the base classifiers are trees with a number of K [25], [26]. For an input test xi, the yield is calculated by:

$$\hat{y}_i = \sum_{i=1}^{k} f_k(x_i), f_k \in F \tag{1}$$

where, $f_k(x_i)$ is the yield of the k_{th} trees and F is the space of all relapse trees. Based on angle boosting, XGBoost makes a few enhancements by regularizing the objective work [27]:

$$L = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \tag{2}$$

Where the previous term could be a misfortune work that measures the contrast between the forecast $\hat{y}_i$ and label y_i. The last-mentioned term could be a regularization term that measures the complexity of the trees [28].

The total objective work cannot be optimized straightforwardly. Instep, added substance way is considered [29]. Let $\hat{y}_i(t)$ be the expectation of the i_{th} test at the t_{th} emphasis, the objective work is composed as:

$$L^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \simeq \sum_{i=1}^{n}\left[l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \Omega(f_t) \quad (3)$$

Where, gi is the primary arranged fractional subsidiary of the misfortune work and hi is the moment arranged halfway subsidiary of the misfortune work. Subsequently, the misfortune work must be twice differentiable. The steady terms of Eq. (3) are evacuated, and the objective work is disentangled as follows:

$$\tilde{L}^{(t)} = \sum_{i=1}^{n}\left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \Omega(f_t) \quad (4)$$

The regularized term is characterized by

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \quad (5)$$

where, T is the number of takes off within the tree. ω is the T-dimension vector of scores on take off. γ and λ» are steady coefficients speaking to the complexity of clears out and scale of punishment. The space of trees is characterized as $F = \{f(x) = \omega_{q(x)}.\, q(x)$ could be an outline relegating the test to the comparing leaf. The occurrence set of leaf j is I_j. Thus, Eq. (4) can be expanded as follows:

$$\tilde{L}^{(t)} = \sum_{i=1}^{n}\left[g_i f_i(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 = \sum_{j=1}^{T}\left[\left(\sum_{i\in I_j} g_i\right)\omega_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right)\omega_j^2\right] + \gamma T \quad (6)$$

Eq. (6) can be encouraged compressed by characterizing $G_j = \sum_{i\in I_j} g_j$ and $H_j = \sum_{i\in I_j} h_i$. Expecting the structure of the tree to be settled, the ideal esteem of all left can be calculated by Eq. (7). And the comparing esteem of objective work can be obtained utilizing Eq. (8).

$$\omega_j^* = -\frac{G_j}{H_i + \lambda} \quad (7)$$

$$\tilde{L}^{(t)}(q) = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_i + \lambda} + \gamma T \quad (8)$$

As the structures of trees can be assessed, an estimation for the part hubs is characterized in Eq. (9). Characterize IL and IR as the occurrence sets after the part.

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L+\lambda} + \frac{G_R^2}{H_R+\lambda} - \frac{(G_L+G_r)^2}{H_L+H_R+\lambda}\right] - \gamma \quad (9)$$

B. Giant Trevally Optimizer (GTO)

The GTO approach was chosen to address the $P-OPF$ issue over different scenarios. The GTO strategy could be a metaheuristic calculation that determines motivation from the chasing behaviors of the monster trevally [30], [31]. The giant trevally employments procedures that include designed scavenging developments, choice of an ideal chasing locale, and jumping out of the water to capture prey [32]. The GTO calculation duplicates these procedures into a $3-$ step preparation: broad look, determination of zone, and assault.

1) Extensive search: The GTO method recreates the long separations mammoth trevallies travel to find nourishment employing a numerical show based on Exact flights, a shape of arbitrary walk. This stage progresses the algorithm's investigation capability and helps in maintaining a strategic distance from nearby optima. The condition utilized in this stage can be outlined as delineated below:

$$X(t+1) = Best_p \times R + (Maximum - Minimum) \times R + Minimum \times Levy(Dim) \quad (10)$$

where, the location vector of the enormous trevally in the subsequent iteration is denoted by $X(t + 1)$, $Best_p$ speaks to the leading position gotten, R speaks to an arbitrary number extending from 1, and $Levy\ (Dim)$ speaks to the Require flight.

2) Choosing area: In this stage, the calculation finds the ideal chasing locale based on nourishment accessibility interior of the look space. The taking after condition is utilized to reproduce this behavior numerically:

$$X(t+1) = Best_p \times \mathcal{A} \times R + Mean_{Info} - Xi(t) \times R \quad (11)$$

where, A may be a parameter that controls position alters, $Xi(t)$ denotes the current position, and R may be an arbitrary number. The successful utilization of all information gotten from earlier areas is suggested by the term Mean Info for these giant trevallies.

3) Attacking: The last arrangement of the algorithm mimics the trevally's attack on its victim. The trevally's behavior is affected by light refraction, which affects its ability to see. In order to replicate this behavior, the computation first uses Snell's equation to compute the visual twisting V. Next, it uses the following to simulate the trevally attack:

$$X(t+1) = \mathcal{L} + \mathcal{V} + \mathcal{H} \quad (12)$$

where, $X(t + 1)$ signifies another position, $\mathcal{L}$ is the dispatch speed, $\mathcal{V}$ is the visual mutilation, and $\mathcal{H}$ is the jumping incline work, in this manner permitting the calculation to move from the stage of investigation to the stage of misuse. Fig. 1 presents the process of the GTO.

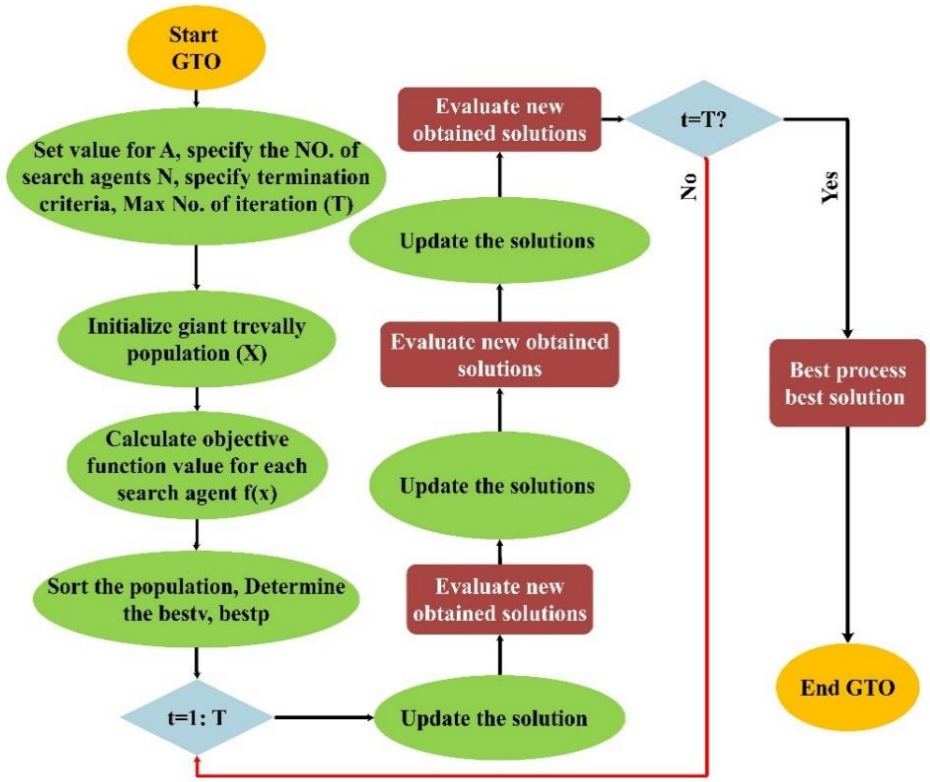


Fig. 1. The process of the GTO.

C. Bald Eagle Search Optimization (BESO)

It's possible that the BES computation is a subsequent meta-heuristic optimization computation that was suggested in 2020. Bald eagles (BE) are ranked highest in the food chain according to their measurement. They are unintentional hunters. They can eat any straightforward, easily available food that is high in protein [33], [34]. They choose an angle, especially salmon, dead or alive, as their primary food source. Because eagles have extraordinary eyesight and can simultaneously look in two different orientations, they can locate angles from a great distance. The main source of inspiration for BES was their cunning social conduct with regard to their pursuing device. Three phases make up the chasing component of BE [35], [36]. These phases include swooping, gazing in space, and choosing a spot. The eagle selects the area with the greatest concentration of prey during the selecting-the-space phase. The eagle starts hunting for prey within the selected space during the searching-in-the-space phase.

Finally, during the swooping phase, the falcon starts to swoop from its optimal position from the previous phase. At that moment, it is chosen which point is optimum to pursue. All of the eagle's subsequent developments are directed toward this goal.

1) Mathematical Model: The numerical definition of the chasing component of BE is characterized by the taking after:

Selecting−space stage. The BE, in this stage, decides the ideal zone based on the sum of nourishment. This behavior is numerically characterized as:

$$X_{new} = X_{best} + \alpha \times r(X_{mean} - X_i) \tag{13}$$

Where X_{best} is the chosen look space based on the finest eagle's position, X_{mean} is the harsh division between each of the bare hawks' postures (cruel of the populace), Xi is the present hawk position, r could be an arbitrary parameter produced in [0-1], and α may be a consistent parameter.

Phase of searching in space. In this step, the BE moves entirely various headings within the selected spiral zone from the previous stage in search of prey. Additionally, a decision is made on who would lead the pursuit and swooping of prey. The following numerical description of this behavior:

$$\begin{gathered}
X_{new} = X_i + z(i) \times (X_i - X_{i+1}) + p(i) \\
\times (X_i - X_{mean}) \\
p(i) = \frac{pr(i)}{max|pr|}, z(i) = \frac{zr(i)}{max|zr|} \\
pr(i) = r(i) \times \cos\big(\theta(i)\big), zr(i) \\
= r(i) \times \sin\big(\theta(i)\big) \\
\theta(i) = \alpha \times \pi \times r1 \\
r(i) = \theta(i) + R \times r2
\end{gathered} \tag{14}$$

where, $r1$ and $r2$ are two random parameters, R is another constant parameter with a value between 0.5 and 2, and α is a constant parameter with a value in the range $[0.5, 2]$. Fig. 2 presents the flowchart of the $BESO$.

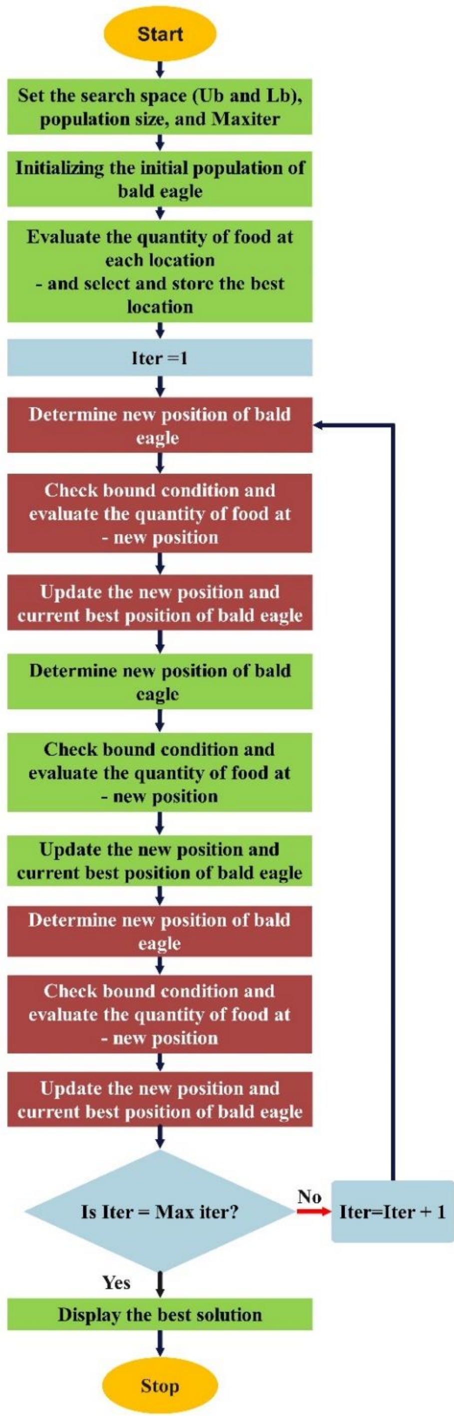


Fig. 2. The flowchart of the BESO.

D. Seagull Optimization Algorithm (SOA)

Around the world, gulls, or more accurately, the seagull family, are seabirds. Seagulls come in a variety of varieties, each with unique bulk and length. Being omnivores, seagulls provide support to squirrels, angels, reptiles, terrestrial and aquatic animals, and night crawlers. The majority of gulls are protected by white plumes [37], [38], [39]. Gulls are exceptionally intelligent feathered creatures. They utilize breadcrumbs to pull in angle and make the sound of rain on their feet to draw in night crawlers covering up underground. Numerical models of predator relocation and assault are examined. The computation recreated the movement of a group of gulls from one area to another during the relocation. The requirements a seagull must fulfill are as follows:

An extra variable, A, is used to calculate the unused look operator area, thereby preventing collisions between adjoining look specialists.

$$C_s = A \times P_s \tag{15}$$

Where Cs speaks to the position of look specialist, which does not collide with other look specialists, Ps speaks to the current position of look operator, x demonstrates the current emphasis, and A speaks to the development behavior of look specialist in a given look space.

$$A = f_c - (x \times ({}^{f_c}/_{Max_{iteration}})) \tag{16}$$

Where fc is presented to control the recurrence of utilizing variable A, which is straightly diminished from fc to 0.

After maintaining a strategic distance from the collision between neighbors, the look specialists move toward the heading of best neighbor.

$$M_s = B \times (P_{bs}(x) - P_s(x)) \tag{17}$$

Where MS speaks to the positions of look operator Ps towards the leading fit look specialist P_{bs}. Because B behaves randomly, it may be trusted to balance proper amounts of abuse and investigation. The calculation for B is:

$$B = 2 \times A^2 \times rd \tag{18}$$

Where rd could be an arbitrary number that lies within the extent of [0,1]. Finally, the look specialist can overhaul its position with regard to the best look specialist by:

$$D_s = |C_s + M_s| \tag{19}$$

Where Ds speaks to the removal between the look operator and best-fit look specialist.

The purpose of this enhancement is to capitalize on the engagement and history of the look preparation. Attacking prey causes the spiraling activity to occur inside the discussion. This behavior is represented as follows in the x, y, and z planes:

$$\acute{x} = r \times \cos(k) \tag{20}$$

$$\acute{y} = r \times \sin(k) \tag{21}$$

$$\acute{z} = r \times k \tag{22}$$

$$r = u \times e^{kv} \tag{23}$$

Where r is the span of each turn of the winding, k may be an irregular number in extend $[\,0 \leq k \leq 2\pi\,]$. u and v are constants to characterize the winding shape, and e is the base of the common logarithm. The overhauled position of the look operator is calculated utilizing Eq. (19) to Eq. (22).

$$P_s(x) = (D_s \times \acute{x} \times \acute{y} \times \acute{z}) + P_{bs}(x) \tag{24}$$

Where Ps spares the leading arrangement and overhauls the position of other look operators.

E. Evaluation Criteria

To see on the off chance that a classifier is sweet, distinctive ways of judging it are utilized. Precision may be a common way to determine how numerous forecasts are right. Accuracy, review, and precision are vital measures that are regularly utilized together. Accuracy measures how exact a test is at finding positive cases, while recall looks at finding all the genuine positive cases. The f1-score could be a combined degree of exactness and review.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{25}$$

$$Precision = \frac{TP}{TP+FP} \tag{26}$$

$$Recall = TPR = \frac{TP}{P} = \frac{TP}{TP+FN} \tag{27}$$

$$F1\ score\ = \frac{2\times Recall \times Precision}{Recall+Precision} \tag{28}$$

These conditions utilize TP for accurately distinguishing a positive case, FP for wrongly anticipating a positive case, TN for accurately foreseeing a negative case, and FN for wrongly foreseeing a negative case.

III. Description of Dataset

As part of continuous research at Nottingham Trent International College, data was collected using a questionnaire [40]. The objective of this research is to elucidate the many elements that impact the time management skills of 125 students. The dataset contains various demographic data on students, including age, gender, nationality, study programs (Foundation, International Year One, Pre-Master's, and Language Only), academic performance indicators, language course achievements, and attendance records. The dataset further includes the responses provided by students on their proficiency in managing their time. Fig. 3 illustrates the connection between the input and output variables, with the respective magnitudes shown on the right side using distinct color coding. The questions are as follows:

Questionnaire:
1. Do you often have a sense of aimlessness in your life without a clear and specific goal? 2. Do you always find it easy to manage your tasks? 3. Once you begin an activity, do you persist with it until you finish it? 4. Do you occasionally feel a sense of insignificance towards the tasks you must do during the day? 5. Do you organize your activities on a daily basis? 6. Do you tend to procrastinate? 7. Do you have a tendency to transition haphazardly from one task to another during the day? 8. Do you abandon your planned activities just because your friend refuses? 9. Do you believe that you adequately use your time? 10. Do you have a tendency to get easily bored with your daily tasks? 11. Do your significant interests and activities in life have a tendency to shift often? 12. Are you aware of the exact amount of time you dedicate to each of your homework assignments?

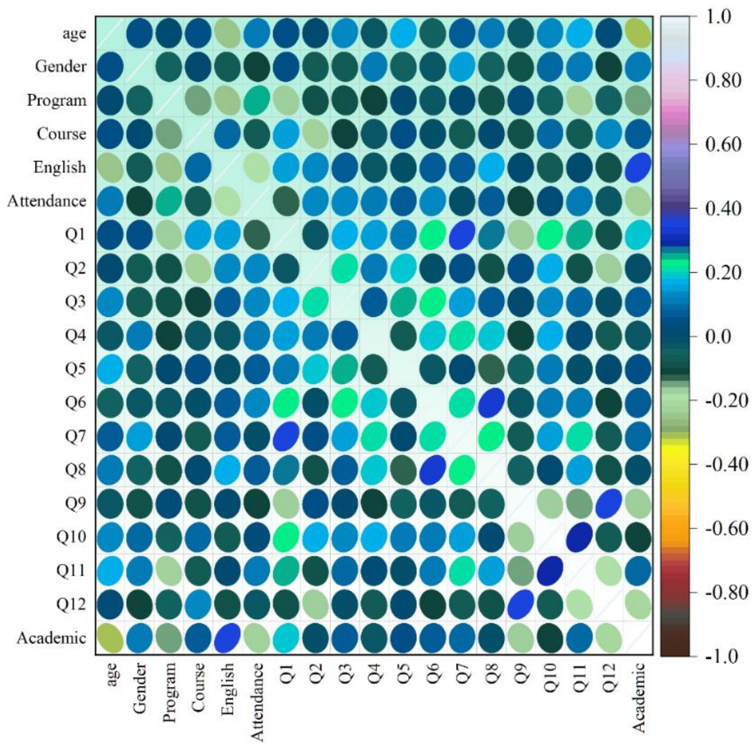


Fig. 3. The relationships between input and output variables by utilizing the correlation matrix.

A. Feature Selection

In ML, feature selection plays a critical role in building efficient and accurate models. Among various techniques, f-classification, a supervised technique utilizing the F-statistic from analysis of variance (ANOVA), offers valuable insights into feature relevance. Fig. 4 displays a bar chart summarizing the results of the feature selection process using f-classification for the input variables. Notably, the average score achieved by this method is 0.35. Features exceeding this average score, highlighted in the chart, were selected for inclusion in the model training process.

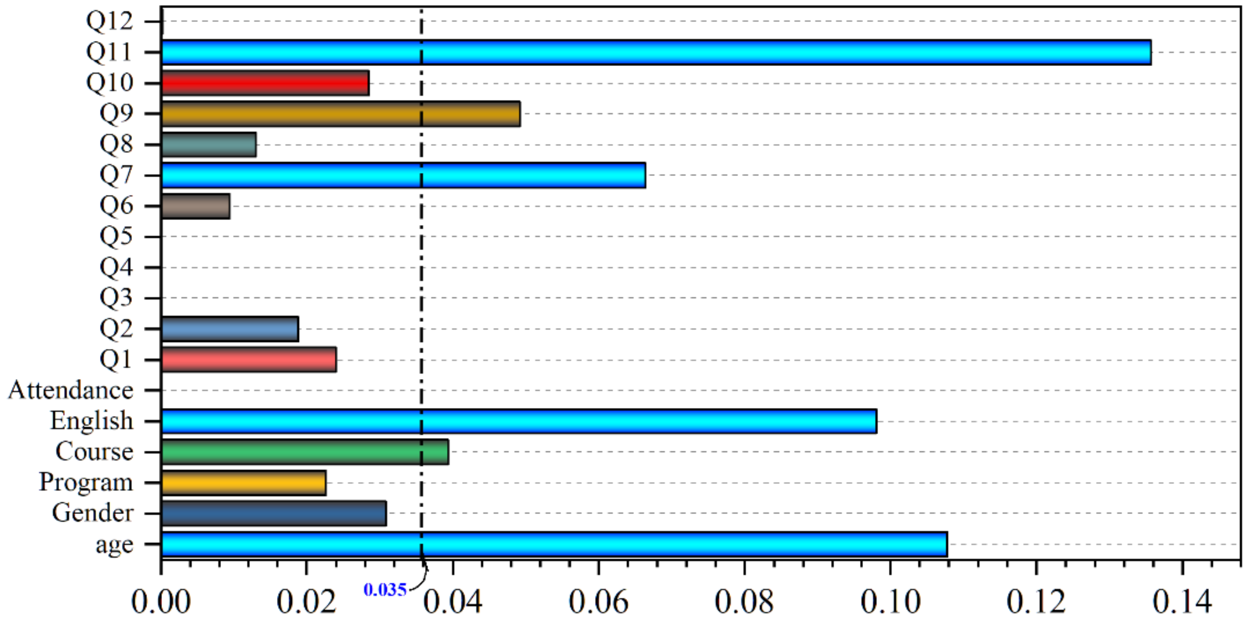


Fig. 4. The bar plot for the result of the feature selection method.

IV. Results

A. Convergence Curve

The headway of an iterative optimization strategy over time is spoken to graphically by the meeting bend that appeared in Fig. 5. It shows the changes within the objective work esteem of the calculation with each emphasis, showing that the calculation is getting near a perfect arrangement. When an algorithm approach joining within the setting of optimization issues, it implies that it consistently minimizes or maximizes the objective work until it comes to a point at which extra emphasis results in minor advancements.

The plot shows that the XGBE model reliably beats the XGGT and XGSO models, accomplishing a top exactness of 0.86 after 120 iterations, showing its predominant execution. Be that as it may, even though the XGGT and XGSO model's precision appeared to be a discernible increment after 110 iterations, it still might not outperform the execution of the XGBE model.

B. Comparative Analysis for Predicted Models Based on Metrics' Results

Table I shows the outcomes of both the single and hybrid models and Fig. 6 visualizes these differences. There are four models, and each model has been compared in four different metric values and two different Sections. The metrics are (Accuracy, Recall, Precision, and F1-score) and the Sections are Train and Test. The model performs best if the number of values gets close to one. Among the evaluated models, XGBE exhibited the highest accuracy in the training section. In terms of precision, XGSO ranked second, followed by XGGT. Conversely, XGBC demonstrated the weakest recall performance during training.

During testing, XGSO achieved the highest F1 score, demonstrating overall solid performance. XGBE and XGSO excelled in the recall, showcasing their ability to identify true positives, with XGGT following closely behind in second place. XGBE took the lead for precision, with XGSO again securing the second-best position.

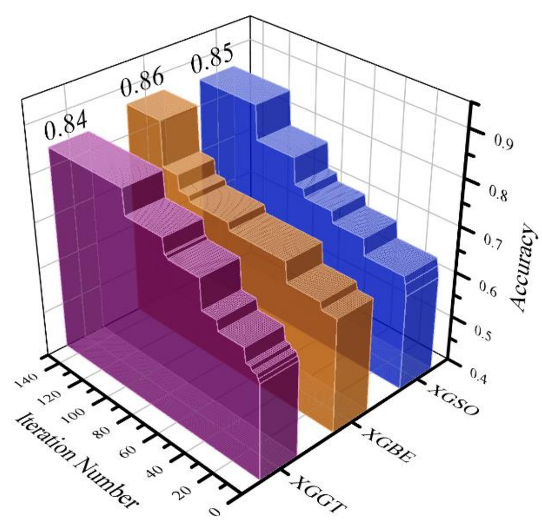
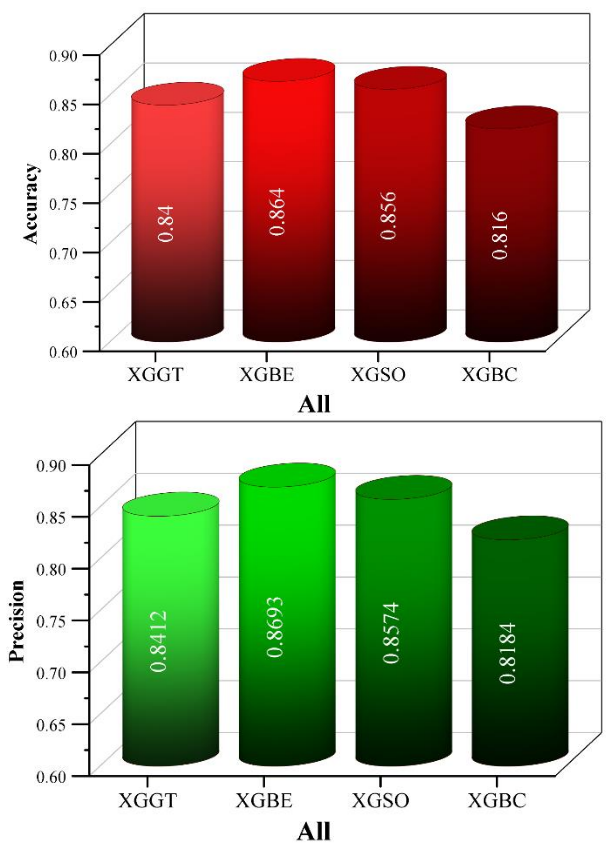


Fig. 5. Convergence curve for hybrid models.

TABLE I. THE OUTCOMES OF BOTH THE SINGLE AND HYBRID MODELS ARE SHOWCASED IN THE PRESENTATION

Section	Metric	Model			
		XGGT	***XGBE***	***XGSO***	***XGBC***
Train	*Accuracy*	0.897	0.920	0.908	0.885
	Precision	0.896	0.923	0.909	0.886
	Recall	0.897	0.920	0.908	0.885
	F1_Score	0.896	0.920	0.908	0.885
Test	*Accuracy*	0.711	0.737	0.737	0.658
	Precision	0.710	0.749	0.737	0.653
	Recall	0.711	0.737	0.737	0.658
	F1_Score	0.703	0.733	0.735	0.634
All	*Accuracy*	0.840	0.864	0.856	0.816
	Precision	0.841	0.869	0.857	0.818
	Recall	0.840	0.864	0.856	0.816
	F1_Score	0.840	0.864	0.856	0.815

Table II presents a comprehensive performance evaluation of four individual and hybrid models across different conditions and metrics. The evaluation leverages four grading criteria (Poor, Acceptable, Good, Excellent) and three key metrics (Precision, Recall, F1-score) to compare model performance comprehensively. Interestingly, both XGBE and XGBC models exhibit peak precision in the "Poor" condition, while both XGGT and XGSO models share the second-best performance. However, under the "Acceptable" condition, XGBE takes the lead in precision, with XGBC exhibiting the weakest performance.



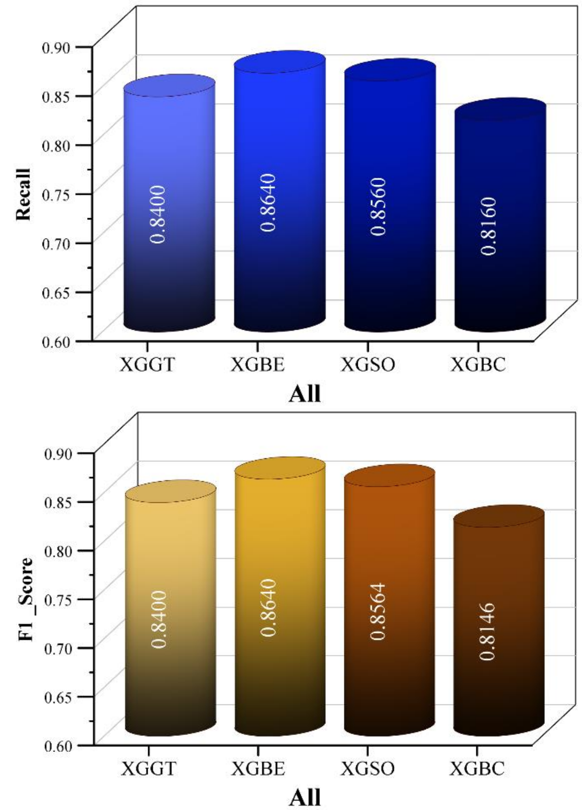


Fig. 6. A 3D bar plot indicating the difference between the measured and predicted values.

This detailed analysis, presented in Table II, allows for a thorough understanding of how different models perform under varying conditions and based on crucial evaluation metrics.

In Recall value at Good condition, three of the models have the same and the best performance. The XGGT, XGBE, XGSO, and XGBC models have the lowest and weakest performance. In excellent condition at the Recall value, the XGBE, XGSO, and XGBC have the highest performance, and the XGGT model has the weakest performance.

TABLE II. MODELS ACHIEVED RESULTS IN THE DIFFERENT PRESENTED CONDITIONS

Metric	Condition	Model			
		XGGT	***XGBE***	***XGSO***	***XGBC***
precision	Poor	0.857	0.923	0.857	0.923
	Acceptable	0.857	0.896	0.840	0.774
	Good	0.850	0.872	0.810	0.811
	Excellent	0.773	0.760	1.000	0.864
recall	Poor	0.800	0.800	0.800	0.800
	Acceptable	0.875	0.896	0.875	0.854
	Good	0.810	0.810	0.810	0.714
	Excellent	0.850	0.950	0.950	0.950
f1-Score	Poor	0.828	0.857	0.828	0.857
	Acceptable	0.866	0.896	0.857	0.812
	Good	0.829	0.840	0.810	0.760
	Excellent	0.810	0.844	0.974	0.905

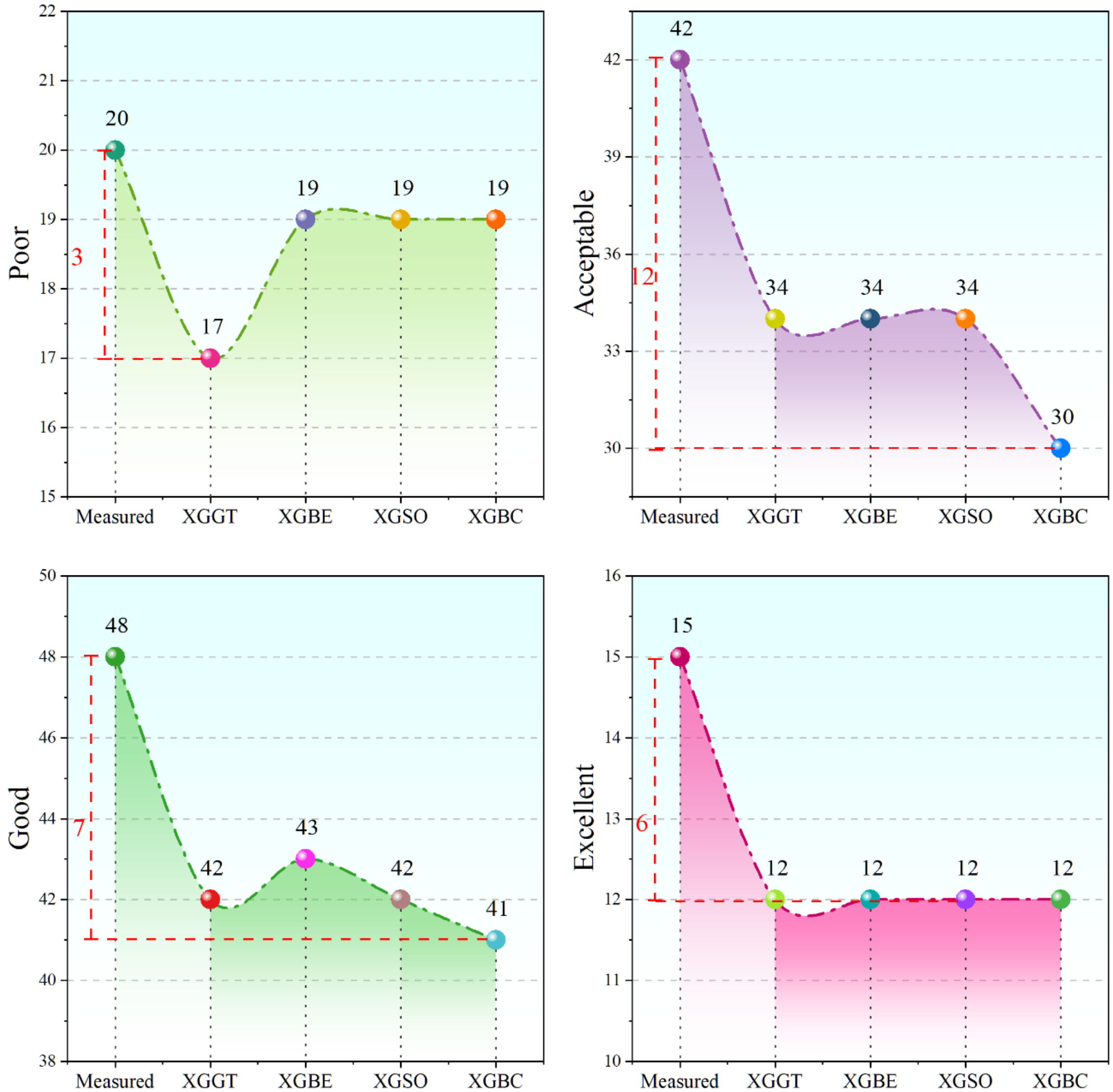


Fig. 7. Line-symbol plot for the visual evaluation of the models' performance.

Fig. 7 presents a line-symbol plot visualizing the performance of different models under varying conditions. In this plot, smaller differences between a model's prediction and the actual measured value indicate better performance for that specific condition.

Acceptable Condition: With a measured value of 42, three models – XGGT, XGBE, and XGSO – exhibited predictions close to the actual value, indicating strong performance in this condition.

Good Condition: For the measured value of 48, XGBE outperformed the others with the closest prediction. Both XGGT and XGSO also predicted close to the measured value, demonstrating good performance.

Excellent Condition: All models achieved the same prediction and performance for the measured value of 15 in this condition.

Fig. 8 presents confusion matrices to evaluate the accuracy of each model across various conditions. For example, in the Poor condition, the XGGT model accurately predicted 17 out of 20 samples, achieving an 85% accuracy rate. Notably, all misclassified samples belonged to the Acceptable category (3 samples). Similarly, in the Acceptable condition, the XGGT model maintained good performance, correctly classifying 34 out of 42 samples (81% accuracy). However, misclassifications occurred in both the Good (5 samples) and Poor (3 samples) categories.

In the evaluation of predictive models, the XGBE model exhibited commendable performance. Among the 48 items categorized as being in Good condition, the XGBE model accurately predicted 43 of them while misclassifying five. Notably, two of the misclassified items were erroneously labeled as Poor condition, two as Acceptable condition, and one as Excellent condition. Moreover, when assessing the 15 items categorized as Excellent condition, the XGBE model demonstrated substantial predictive accuracy by correctly identifying 12 of them. However, it did misclassify three items, with two categorized as Acceptable condition and one as Poor condition. Furthermore, in the assessment of 48 items initially classified as Good condition, the XGSO model accurately predicted 42 of them. However, it misclassified six items, with five categorized as Acceptable condition and one as Excellent condition.

Fig. 9 depicts the Receiver Operating Characteristic (ROC) curves generated to assess the effectiveness of the most proficient hybrid models. The ROC curve serves as a widely utilized visual aid for evaluating a model's performance and illustrating the balance between sensitivity and specificity in binary classification tasks. Sensitivity, also known as the true positive rate or recall, gauges a model's ability to detect positive cases accurately. Conversely, specificity denotes the true negative rate and indicates how well the model can identify negative cases. The ROC curve plots the true positive rate against the false positive rate at various thresholds for the assessed probabilities of the model.

Fig. 9 unmistakably illustrates that the XGBE model exhibits the highest and most consistent performance in identifying cases classified as Poor. It consistently achieves the highest true positive rate (TPR) while maintaining the lowest false positive rate (FPR), underscoring its reliability.

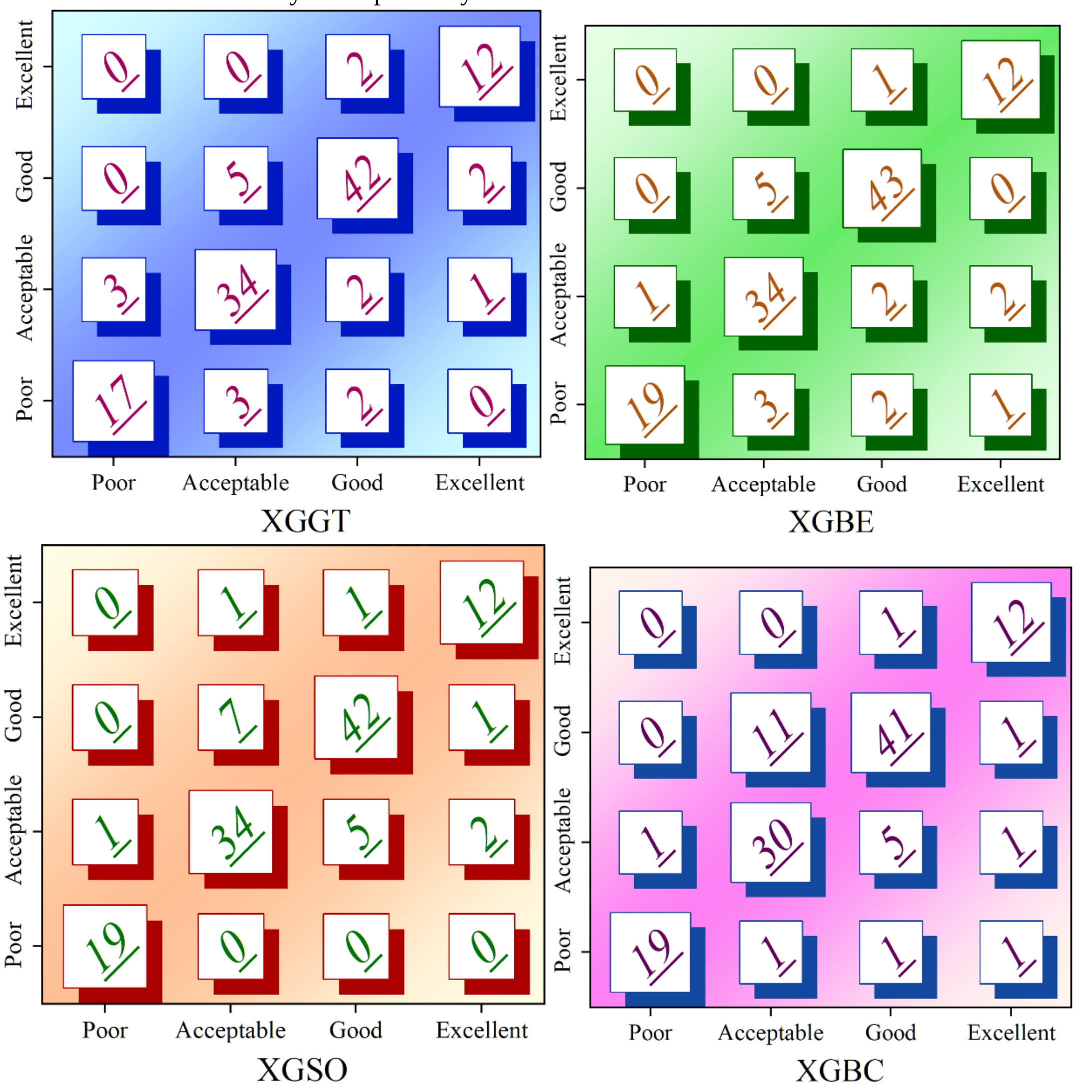


Fig. 8. Confusion matrix for the correctly classified and misclassified values of the models.

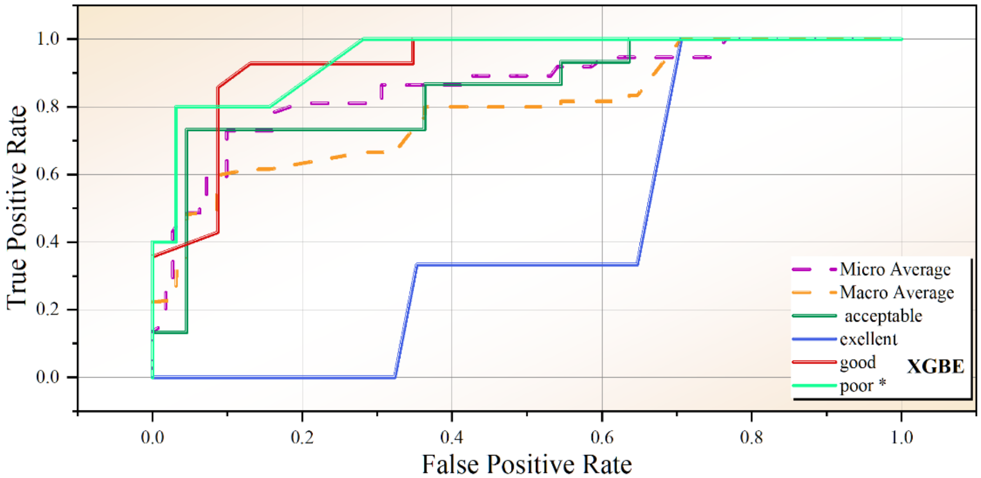


Fig. 9. Line plot for the ROC curve of the best-performed hybrid model.

V. Discussion

A. Limitations of the Study

The study acknowledges several limitations, including the possibility that the dataset used may not fully capture all relevant variables influencing academic performance, suggesting the presence of unmeasured factors. While the predictive models showed promising accuracy, there remains room for improvement, as alternative optimization techniques not explored could enhance performance. Additionally, the generalizability of the findings may be limited to the specific context and population studied, meaning the results might not apply to different educational settings or student groups. These limitations highlight areas for further research and potential enhancements in future studies.

B. Implications of the Study

The study's implications and significances are notable in several areas. Firstly, it underscores the importance of time management skills in predicting academic performance, highlighting a previously underexplored factor in educational outcomes. By integrating ML models with data from the TSQ, the research offers a novel approach to understanding and forecasting student success. The findings suggest that educational institutions can leverage these models to identify at-risk students early, allowing for timely interventions that could reduce dropout rates and foster a more supportive learning environment. Furthermore, the study demonstrates the effectiveness of using advanced optimization algorithms, such as the Bald Eagle Optimizer, in enhancing model accuracy, thereby contributing to the broader field of educational data mining and predictive analytics. These insights pave the way for future research and the development of more comprehensive and accurate predictive tools in education.

C. Future Works

Future work in predicting academic performance using ML and time management skills could explore incorporating a broader range of variables, such as psychological factors and extracurricular activities, to create more comprehensive models. Researchers could experiment with additional optimization techniques to improve model accuracy and conduct longitudinal studies to track performance over time. Validating these models across diverse educational settings and student populations would help assess their generalizability. Developing and evaluating intervention strategies based on these models could enhance their practical utility, while integrating them into real-world educational systems could support timely interventions and reduce dropout rates. Additionally, creating user-friendly tools for educators and addressing ethical considerations related to privacy and data security will be crucial for the responsible implementation of these technologies.

VI. Conclusion

The investigation into predicting academic students' performance has yielded valuable new insights into the intricate interplay among the myriad factors influencing educational outcomes. These factors encompass socioeconomic backgrounds, study habits, and prior academic achievements. A comprehensive examination of these factors has helped elucidate these relationships. Furthermore, this study has concentrated on analyzing a dataset that encompasses students' time management skills and their influence on academic performance. Despite the valuable insights gained, it is important to acknowledge some limitations of this study. Firstly, the dataset used may not fully capture all relevant variables influencing academic performance, and there may be other unmeasured factors at play. Additionally, the predictive models developed in this study, while demonstrating promising accuracy in forecasting academic performance, may still have room for improvement. The integration of three optimizers (BESO, SAO, and GTO) with the base model XGBoost Classifier (XGBC) aimed to enhance performance, but there could be alternative optimization techniques that were not explored. Moreover, the generalizability of the findings may be limited to the specific context and population studied. Despite these limitations, as the education sector evolves, there is a clear opportunity to integrate these models into school systems to identify at-risk students early and provide them with timely support. Such proactive measures can significantly reduce dropout rates and foster a more supportive learning environment. The models underwent differentiation across various stages, scenarios, and metrics. Among them, the XGBoost with Bald Eagle Optimizer (XGBE) emerges as particularly robust in this analysis compared to the XGBoost with Giant trevally Optimizer (XGGT), XGBoost with Seagull Optimization Algorithm (XGSO), and XGBoost Classifier (XGBC) models. The XGBE model demonstrated superior performance, boasting high precision and accuracy values of 0.920 and 0.923, respectively, during the training phase, surpassing both the XGSO, XGGT, and XGBC models.

ACKNOWLEDGMENT

This work was supported by. 2020 Liaoning Province Education Science "13th Five-Year Plan" project “Research on the Course Construction of Police Command and Tactics in Police Colleges and Universities” project number: JG20DB452.

This work was supported by. Plan to cultivate major projects of basic scientific Research funds for central universities in 2020“Research on the application of self-actuated intelligent shooting gloves in improving the shooting technique of police” project number:3242020016.

This work was supported by. 2022 Social Science Planning Fund Project of Liaoning Province “Research on the Legal Responsibility of Police Weapon Use under the background of anti-terrorism response” project number: L22BFX011.

REFERENCES

[1] C. Basila, “Good time management and motivation level predict student academic success in college on-line courses,” International Journal of Cyber Behavior, Psychology and Learning (IJCBPL), vol. 4, no. 3, pp. 45–52, 2014.

[2] B. K. Britton and A. Tesser, “Effects of time-management practices on college grades.,” J Educ Psychol, vol. 83, no. 3, p. 405, 1991.

[3] M. Rahmat, S. Shahrani, R. Latih, N. F. M. Yatim, N. F. A. Zainal, and R. Ab Rahman, “Major problems in basic programming that influence student performance,” Procedia-Social and Behavioral Sciences, vol. 59, pp. 287–296, 2012.

[4] R. V Adams and E. Blair, “Impact of time management behaviors on undergraduate engineering students’ performance,” Sage Open, vol. 9, no. 1, p. 2158244018824506, 2019.

[5] K. R. Wentzel and A. Wigfield, “Academic and social motivational influences on students’ academic performance,” Educ Psychol Rev, vol. 10, pp. 155–175, 1998.

[6] S. L. Miertschin, C. E. Goodson, and B. L. Stewart, “Time management skills and student performance in online courses,” in 2015 ASEE Annual Conference & Exposition, 2015, pp. 26–1585.

[7] T. H. Macan, C. Shahani, R. L. Dipboye, and A. P. Phillips, “College students’ time management: Correlations with academic performance and stress.,” J Educ Psychol, vol. 82, no. 4, p. 760, 1990.

[8] S. Nasrullah_PhD and M. S. Khan_PhD, “The impact of time management on the students’ academic achievements,” Journal of Literature, Languages and Linguistics, vol. 11, pp. 66–71, 2015.

[9] M. Mohammed, M. B. Khan, and E. B. M. Bashier, Machine learning: algorithms and applications. Crc Press, 2016.

[10] B. J. C. Claessens, W. Van Eerde, C. G. Rutte, and R. A. Roe, “A review of the time management literature,” Personnel review, vol. 36, no. 2, pp. 255–276, 2007.

[11] C. Bishop, “Pattern recognition and machine learning,” Springer google schola, vol. 2, pp. 5–43, 2006.

[12] O. Z. Maimon and L. Rokach, Data mining with decision trees: theory and applications, vol. 81. World scientific, 2014.

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

[14] C.-C. Kiu, “Data mining analysis on student’s academic performance through exploration of student’s background and social activities,” in 2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), IEEE, 2018, pp. 1–5.

[15] Y. S. Alsalman, N. K. A. Halemah, E. S. AlNagi, and W. Salameh, “Using decision tree and artificial neural network to predict students academic performance,” in 2019 10th international conference on information and communication systems (ICICS), IEEE, 2019, pp. 104–109.

[16] L. Mutanu and P. Machoka, “Enhancing computer students’ academic performance through predictive modelling-a proactive approach,” in 2019 14th International Conference on Computer Science & Education (ICCSE), IEEE, 2019, pp. 97–102.

[17] P. Sokkhey and T. Okazaki, “Comparative study of prediction models on high school student performance in mathematics,” in 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), IEEE, 2019, pp. 1–4.

[18] B. J. C. Claessens, W. Van Eerde, C. G. Rutte, and R. A. Roe, “A review of the time management literature,” Personnel review, vol. 36, no. 2, pp. 255–276, 2007.

[19] N. T. Feather and M. J. Bond, “Time structure and purposeful activity among employed and unemployed university graduates,” Journal of Occupational Psychology, vol. 56, no. 3, pp. 241–254, 1983.

[20] T. H. Macan, C. Shahani, R. L. Dipboye, and A. P. Phillips, “College students’ time management: Correlations with academic performance and stress.,” J Educ Psychol, vol. 82, no. 4, p. 760, 1990.

[21] B. K. Britton and A. Tesser, “Effects of time-management practices on college grades.,” J Educ Psychol, vol. 83, no. 3, p. 405, 1991.

[22] N. T. Feather and M. J. Bond, “Time structure and purposeful activity among employed and unemployed university graduates,” Journal of Occupational Psychology, vol. 56, no. 3, pp. 241–254, 1983.

[23] S. Malykh et al., “Large-scale study of the precision of the approximate number system: Differences between formats, heterogeneity and congruency effects,” Heliyon, vol. 9, no. 4, 2023.

[24] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016, pp. 785–794.

[25] Y. Qiu, J. Zhou, M. Khandelwal, H. Yang, P. Yang, and C. Li, “Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration,” Eng Comput, vol. 38, no. Suppl 5, pp. 4145–4162, 2022.

[26] W. Li, Y. Yin, X. Quan, and H. Zhang, “Gene expression value prediction based on XGBoost algorithm,” Front Genet, vol. 10, p. 1077, 2019.

[27] S. Ramraj, N. Uzir, R. Sunil, and S. Banerjee, “Experimenting XGBoost algorithm for prediction and classification of different datasets,” International Journal of Control Theory and Applications, vol. 9, no. 40, pp. 651–662, 2016.

[28] Ogunleye and Q.-G. Wang, “XGBoost model for chronic kidney disease diagnosis,” IEEE/ACM Trans Comput Biol Bioinform, vol. 17, no. 6, pp. 2131–2140, 2019.

[29] O. Sagi and L. Rokach, “Approximating XGBoost with an interpretable decision tree,” Inf Sci (N Y), vol. 572, pp. 522–542, 2021.

[30] H. T. Sadeeq and A. M. Abdulazeez, “Giant trevally optimizer (GTO): A novel metaheuristic algorithm for global optimization and challenging engineering problems,” IEEE Access, vol. 10, pp. 121615–121640, 2022.

[31] Z. Hai-yu, “Virtual Machine Allocation in Cloud Computing Environments using Giant Trevally Optimizer,” International Journal of Advanced Computer Science and Applications, vol. 14, no. 9, 2023.

[32] D. M. Utama and C. Febrita, “Low-carbon no-idle permutation flow shop schedulling problem: giant trevally optimizer vs African vultures optimization algorithm,” Int J Adv Appl Sci, vol. 12, no. 3, pp. 195–204, 2023.

[33] P. Ashwini, N. Suguna, and N. Vadivelan, “Improved bald eagle search optimization with entropy-based deep feature fusion model for breast cancer diagnosis on digital mammograms,” Multimed Tools Appl, pp. 1–19, 2023.

[34] Al Mazroa, M. K. Ishak, A. Aljarbouh, and S. M. Mostafa, “Improved Bald Eagle Search Optimization With Deep Learning-Based Cervical

Cancer Detection and Classification," IEEE Access, vol. 11, pp. 135175–135184, 2023.

[35] Gokul Karthik, R. Saravanakumar, and P. Vijayakumar, "Bald eagle search optimization on dual fueled reactivity controlled combustion ignition based engine characteristics by altering low reactive fuels," Environ Prog Sustain Energy, vol. 40, no. 6, p. e13683, 2021.

[36] L. A. Maghrabi et al., "Enhancing Cybersecurity in the Internet of Things Environment Using Bald Eagle Search Optimization with Hybrid Deep Learning," IEEE Access, 2024.

[37] G. Dhiman et al., "EMoSOA: a new evolutionary multi-objective seagull optimization algorithm for global optimization," International Journal of Machine Learning and Cybernetics, vol. 12, pp. 571–596, 2021.

[38] S. Sankar, R. Somula, B. Parvathala, S. Kolli, and S. Pulipati, "SOA-EACR: Seagull optimization algorithm based energy aware cluster routing protocol for wireless sensor networks in the livestock industry," Sustainable Computing: Informatics and Systems, vol. 33, p. 100645, 2022.

[39] V. Kumar, D. Kumar, M. Kaur, D. Singh, S. A. Idris, and H. Alshazly, "A novel binary seagull optimizer and its application to feature selection problem," IEEE Access, vol. 9, pp. 103481–103496, 2021.

[40] M. R. Rimadana, S. S. Kusumawardani, P. I. Santosa, and M. S. F. Erwianda, "Predicting student academic performance using machine learning and time management skill data," in 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), IEEE, 2019, pp. 511–515.