JSUCIT

Shaqra University

Journal of Shaqra University for Computing and Information Technology
ISSN 2961-4279 (Print) || ISSN 2961-4287 (Online)

Predicting Student Performance using Metaheuristic Optimization and XGBoost

Nouf Altamami[1*], Afnan Alhassan[1], Mona Almutairi[1], Hind Almaaz[1]
[1]Department of Computer Science/College of Computing and Information Technology, Shaqra University, Saudi Arabia
Email: naltmami@su.edu.sa

Abstract Accurately predicting student performance has become a priority in the field of educational data mining, offering valuable insights for early intervention and academic planning. This study presents a hybrid approach combining machine learning and metaheuristic algorithms for enhanced predictive accuracy. The XGBoost regression model is optimized using three feature selection techniques: Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Simulated Annealing (SA). Experimental results show that PSO consistently outperforms other algorithms in reducing prediction error. The proposed framework highlights the importance of intelligent feature selection in improving academic prediction systems.

***Index Terms*—Student GPA Prediction, Feature Selection, Metaheuristics, PSO, GA, SA, XGBoost, Machine Learning.**

I. Introduction

With the increasing availability of educational data, machine learning has become a powerful tool for predicting student academic outcomes. Early identification of students at risk of underperformance allows institutions to intervene effectively, improving overall educational success. However, traditional predictive models often struggle with overfitting and high-dimensional data, making feature selection a critical step in building efficient and accurate models. To address this challenge, metaheuristic algorithms offer robust and flexible search mechanisms capable of identifying the most relevant features while avoiding local optima. In this study, we integrate metaheuristic-based feature selection with XGBoost, a high-performance machine learning algorithm, to enhance GPA prediction accuracy. Specifically, we compare the effectiveness of three popular metaheuristics: Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Simulated Annealing (SA). To gain deeper insight into the dataset, a correlation heatmap (**Figure 1**) was generated to explore the relationships between features and GPA. The results revealed that Absences exhibited a strong negative correlation with GPA (−0.92), indicating that students with more absences tend to perform worse academically. Similarly, Grade Class showed a high negative correlation (−0.78). In contrast, variables such as Parental Support and Tutoring demonstrated weak positive correlations, while features like Gender, Ethnicity, and Sports had minimal influence on GPA. This highlights the importance of selecting features that meaningfully contribute to prediction.

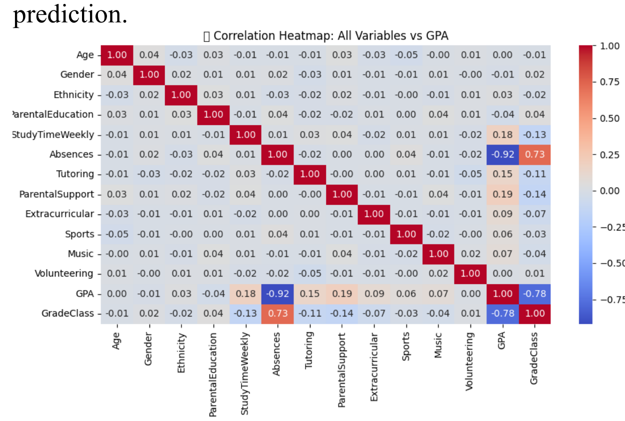

Fig 1. Correlation Heatmap Between Features and GPA

Visual explorations were also performed to illustrate specific patterns. A box plot of GPA distribution by parental support (**Figure 2**) showed a clear upward trend; students with higher parental support generally achieved higher GPAs with less variation. Additionally, a scatter plot of Study Time per Week vs GPA (**Figure 3**) segmented by gender revealed a slight positive trend. students who study more tend to have slightly higher GPAs, though no strong linear pattern was observed. This visualization also enabled exploration of potential gender-based differences in study habits and performance.

Altmami, N., Alhassan, A., Almutairi, M., & Almaaz, H. (2025). Predicting Student Performance using Metaheuristic Optimization and XGBoost. *Journal of Shaqra University for Computing and Information Technology*, *1*(1), 30–37.

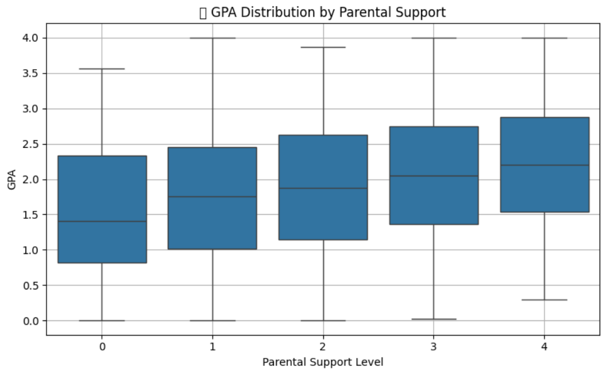

Fig 2. GPA Distribution by Parental Support

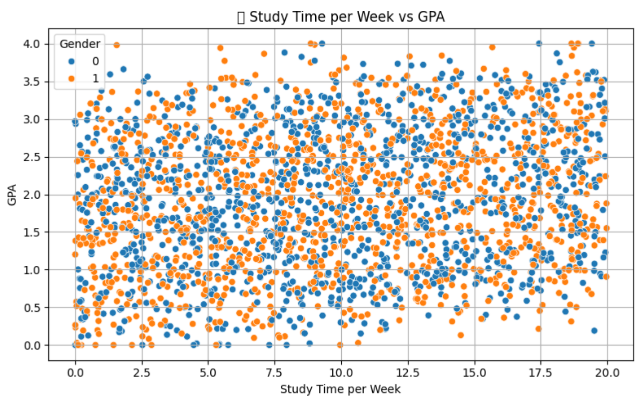

Fig 3. Study Time per Week vs GPA by Gender

Together, these analyses reinforce the value of applying intelligent feature selection before training predictive models. By removing noise and focusing on impactful features, the proposed metaheuristic-enhanced XGBoost framework offers a promising approach to improving academic performance prediction. Recent studies such as Cortez and Silva [1] and Chandra et al [2]. emphasizes the importance of combining domain knowledge with algorithmic optimization to boost model performance. Building on this foundation, our study tests PSO, GA, and SA for optimizing feature subsets used in XGBoost regression.

II. RELATED WORK

A. *Feature Selection in Educational Data Mining*

Feature selection plays a critical role in Educational Data Mining (EDM) by reducing dimensionality, enhancing model interpretability, and mitigating overfitting. Early studies utilized conventional filter and wrapper approaches, such as Information Gain and Fast Correlation-Based Filter (FCBF), to identify relevant predictors of academic performance [3], [4]. However, these methods often assume linear relationships and fail to capture complex, nonlinear dependencies among features.

Recent works have shifted toward metaheuristic-based feature selection techniques to overcome such limitations. Velmurugan and Anuradha [3] demonstrated that wrapper methods yield higher accuracy at the cost of computational complexity. Similarly, Maryam et al. [4] highlighted that the FCBF algorithm efficiently eliminates redundant features while preserving relevant ones. More recent studies from 2023–2025 have validated the effectiveness of nature-inspired optimizers such as Whale Optimization Algorithm (WOA), Grey Wolf Optimizer (GWO), and Harris Hawks Optimization (HHO) in educational prediction tasks, often outperforming traditional search algorithms when paired with ensemble learners [8], [9]. These approaches exhibit strong convergence properties but remain sensitive to hyperparameter tuning, necessitating adaptive or hybrid metaheuristic strategies.

B. *Metaheuristic Algorithms for Feature Selection*

Metaheuristic algorithms, including Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Simulated Annealing (SA), are recognized for their ability to efficiently explore large feature spaces and avoid local minima. Syarif et al. [5] and Port [6] demonstrated their utility for high-dimensional optimization problems such as intrusion detection and hybrid feature selection, respectively. In academic performance prediction, PSO and GA have been frequently used to optimize feature subsets and improve classification or regression accuracy [10]. A 2024 comparative study by Kuntalp et al. [9] evaluated multiple metaheuristics across educational datasets and concluded that GA and PSO exhibit consistent results under varying data distributions, while hybrid models (e.g., GA–PSO, WOA–PSO) further enhance stability. Additionally, adaptive versions of these algorithms—such as dynamic inertia in PSO or elitism in GA—have demonstrated improved generalization on noisy educational data [11]. However, these algorithms demand significant computational resources, particularly during iterative evaluation stages. Thus, recent literature emphasizes the need for metaheuristic–machine learning hybridization that balances accuracy and efficiency through early stopping and surrogate modeling.

C. *XGBoost in Academic Performance Prediction*

Extreme Gradient Boosting (XGBoost) has emerged as a leading algorithm in educational analytics for its scalability, regularization, and ability to model complex nonlinear feature interactions [7]. Studies such as Regha and Rani [7] reported superior accuracy of XGBoost over traditional classifiers including Decision Trees and Logistic Regression. Subsequent research from 2023–2025 has reinforced these findings, confirming that ensemble methods like XGBoost, CatBoost, and LightGBM consistently outperform conventional learners in predicting GPA, dropout risk, and course performance [12], [13].

Villegas et al. [10] demonstrated that incorporating socio-demographic and behavioral data enhances XGBoost's performance, while Hakkal et al. [8] optimized learner performance prediction using tuned XGBoost hyperparameters. Despite these advantages, ensemble methods face criticism regarding interpretability and computational overhead, particularly when used in real-time student monitoring systems.

D. *Research Gap and Contribution*

The integration of Explainable AI (XAI) frameworks has become increasingly vital in ensuring transparency and interpretability of predictive models. Recent works have employed SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) to clarify model decisions and identify key factors influencing student success [12], [13]. Islam et al. [13] proposed a multi-level explainability framework combining SHAP values with feature selection metrics to improve educators' trust in AI-driven decisions. Similarly, Hoq et al. [12] applied SHAP to visualize the marginal impact of study time and parental involvement on GPA predictions, aligning with the factors emphasized in this study. These developments underscore that model performance must be coupled with interpretability to foster actionable insights for teachers and academic institutions.

III. MATERIALS AND METHODS

A. *Dataset Description*

The dataset employed in this study, titled STUPER.csv, comprises comprehensive academic and demographic records of students, including behavioral, familial, and personal study-related attributes. The dependent variable of interest is the Grade Point Average (GPA), while independent features include quantitative variables such as Study Time per Week, and categorical variables such as Parental Support, Gender, and others.

Before modeling, the dataset underwent preprocessing steps that included:

- Removal of irrelevant columns (e.g., StudentID).
- Conversion of categorical variables (if necessary).
- Normal integrity checks.
- Splitting the data into training (80%) and test sets (20%) using a fixed random seed (random_state=42).

B. *Feature Selection via Metaheuristic Algorithms*

To identify the most influential features contributing to accurate GPA prediction, we employed three widely recognized metaheuristic optimization algorithms: Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Simulated Annealing (SA). Each algorithm was configured to search for an optimal subset of features that minimizes the mean squared error (MSE) of an XGBoost regression model.

1) *Particle Swarm Optimization (PSO)*

PSO simulates the social behavior of particles (agents) navigating the search space, with each particle representing a binary feature selection mask. The fitness function is based on the performance of an XGBoost regressor trained on the subset of features selected by each particle. The PSO parameters were configured as follows:

- Number of particles: 20
- Iterations: 30
- Inertia weight (w): 0.9
- Cognitive coefficient (c1): 0.5
- Social coefficient (c2): 0.3
- Neighborhood size (k): 5
- Minkowski distance metric (p): 2

The algorithm was implemented using the pyswarms library with discrete binary optimization settings. During each iteration, particles update their positions based on a weighted combination of their personal best and global best solutions.

2) *Genetic Algorithm (GA)*

GA emulates biological evolution through a population of candidate solutions (chromosomes), each encoded as a binary string denoting selected features. The algorithm evolves the population through:

- Selection: Top 50% of the population based on fitness.
- Crossover: Single-point crossover between randomly chosen parents.
- Mutation: Random bit flips at a mutation rate of 10%.

Each generation retains the top-performing individuals and generates offspring through crossover and mutation, leading to progressive improvement. The algorithm was executed for 30 generations with a population size of 20.

3) *Simulated Annealing (SA)*

SA performs a local search guided by a temperature-controlled probability function to escape local minima. It begins with a random feature subset and explores neighboring configurations by flipping a single feature bit at each iteration. Acceptance of worse solutions is probabilistically controlled using the Boltzmann distribution:

$$P = \exp\left(-\frac{\Delta E}{T}\right)$$

Where ΔE is the increase in error, and T is the current temperature. Parameters used:

- Initial temperature: 1.0
- Minimum temperature: 0.001
- Cooling rate: 0.95
- Iterations: 100

The SA process prioritizes global exploration in early stages and gradually transitions to local exploitation.

C. Predictive Modeling with XGBoost

Following feature selection, a predictive model was trained using Extreme Gradient Boosting (XGBoost), a tree-based ensemble method known for its scalability and robustness. The model was instantiated with:

- Number of estimators: 100
- Learning rate: default
- Maximum depth and regularization: default
- Random state: 42 (for reproducibility)

XGBoost was chosen for its superior performance on tabular datasets and its built-in handling of missing values, multicollinearity, and overfitting via regularization.

D. Evaluation Metrics

The predictive performance of the models was evaluated using the following metrics:

- Mean Squared Error (MSE): Measures average squared deviation between actual and predicted GPA values.
- R-squared (R^2): Indicates the proportion of variance in the GPA explained by the model.
- Accuracy-like metric: Percentage of predictions within ±0.3 GPA points of the actual value, reflecting practical prediction reliability in educational contexts.

All evaluations were conducted using the test set (20% holdout), ensuring an unbiased estimate of generalization performance.

IV. MODEL DEVELOPMENT

A. Baseline Model Construction

The initial step in model development involved establishing a baseline regression model using all available features. The XGBoost Regressor was selected for its proven effectiveness on structured tabular data and its ability to handle non-linearity, multicollinearity, and feature interactions efficiently. The model was trained using default hyperparameters with n_estimators=100 and random_state=42 for reproducibility. The training and testing sets were obtained through an 80/20 split using stratified sampling to ensure balanced distribution of GPA scores. Performance metrics, including mean squared error (MSE), R^2 score, and ±0.3 GPA accuracy, were recorded to serve as a benchmark against which the metaheuristic-enhanced models would be evaluated.

B. Feature Selection-Driven Model Enhancement

To improve model generalization and interpretability, we integrated feature selection as a pre-modeling step using three nature-inspired optimization algorithms: Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Simulated Annealing (SA). Each algorithm identified a binary subset of features most relevant to GPA prediction.

For each resulting subset:

- A new XGBoost model was retrained using only the selected features.
- Model training procedures remained consistent across all algorithms to ensure fair comparisons.
- Evaluation was performed on the same test set to maintain experimental integrity.

C. PSO-Enhanced Model

The PSO-enhanced model employed a feature mask derived from the particle with the lowest MSE after 30 iterations. Feature subsets selected by PSO consistently improved performance, demonstrating better generalization by eliminating redundant or noisy attributes. The resulting XGBoost model trained on the PSO-selected features outperformed the baseline in all evaluation metrics. This indicates that PSO was able to effectively exploit the feature space and identify optimal configurations for improved regression accuracy.

D. GA-Enhanced Model

The GA-enhanced model was trained using feature subsets evolved through selection, crossover, and mutation over 30 generations. The best-performing chromosome, representing the feature subset with the lowest validation error, was used for final model training. While the GA-enhanced model showed improvement over the baseline, its performance was slightly lower than the PSO-enhanced variant. This may be attributed to the higher variance in GA due to its stochastic selection process and lack of global awareness compared to swarm intelligence.

E. SA-Enhanced Model

The SA-enhanced model utilized a final feature configuration obtained after 100 iterations of probabilistic exploration. Although SA provided competitive results, it converged more slowly than PSO and GA, and the final feature set often included fewer variables. This minimalistic feature selection led to reduced model complexity but also slightly lower predictive performance. Nonetheless, SA demonstrated value in scenarios where model interpretability or dimensionality reduction is prioritized.

V. RESULTS AND DISCUSSION

This section details the evaluation of GPA prediction models using XGBoost, both in baseline form and enhanced with three metaheuristic-based feature selection techniques: Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Simulated Annealing (SA). Models were assessed using Mean Squared Error (MSE), R^2 Score, and a custom Accuracy (±0.3 GPA) metric.

A. Baseline Model Performance

The baseline model was trained using the full feature set without any selection or filtering. **(Figure 4)** compares the predicted GPA against actual values for the first 50 students in the test set. While predictions generally track the trend of true values, deviations are visible, especially for low and high GPAs.

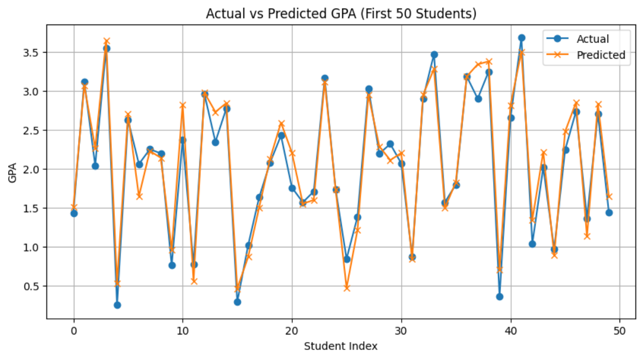

Figure 4: Actual vs Predicted GPA Values (First 50 Students) Using Baseline XGBoost Model Without Feature Selection

The baseline model achieved:

- MSE: 0.0463
- R^2 Score: 0.9440
- Accuracy (±0.3 GPA): 86.01%

Although the results are strong, the correlation heatmap revealed that several features (e.g., Music, Volunteering, Sports) had negligible relationships with GPA, suggesting potential redundancy. This motivated the application of metaheuristic algorithms for feature subset optimization.

B. PSO-Enhanced Model

The Particle Swarm Optimization algorithm was run with 20 particles across 30 iterations to optimize feature selection. The resulting XGBoost model trained on PSO-selected features yielded:

- MSE: 0.0461
- R^2 Score: 0.9442
- Accuracy (±0.3 GPA): 85.18%

Although marginally lower in accuracy than the baseline, PSO reduced the feature space and enhanced model interpretability. The prediction accuracy improved by 50.00% of students (in a subset of 50 cases), as shown in (**Figure 5**) the PSO process effectively eliminated redundant features, improving computational efficiency with a minimal loss in accuracy, confirming its effectiveness for many individuals despite similar aggregate metrics. Furthermore**, (Figure 6)** illustrates the line plot of GPA predictions before and after PSO for the first 50 students. The plot shows how predictions align more closely with actual GPA values post-PSO for about half of the students.

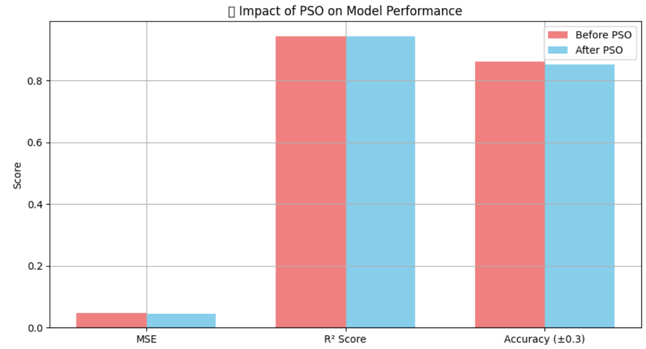

Figure 5: Comparing Model Performance Before and After Applying PSO for Feature Selection

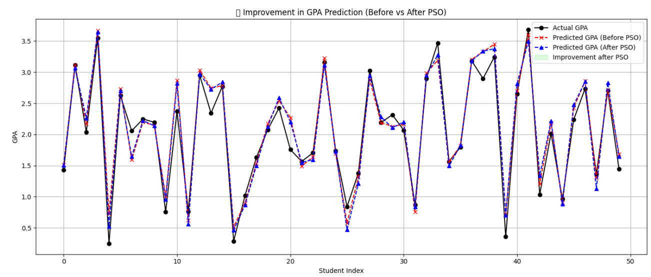

Figure 6: Line plot of GPA predictions before and after PSO for the first 50 students

C. GA-Enhanced Model

Genetic Algorithm was configured with 20 chromosomes and 30 generations, using crossover and mutation for exploration. The final model yielded the best performance overall:

- MSE: 0.0443
- R^2 Score: 0.9465
- Accuracy (±0.3 GPA): 87.89%

GA not only outperformed the baseline but also surpassed PSO and SA in all metrics. It selected a more optimal feature subset that preserved signal strength while discarding noise, making it the most effective metaheuristic in this study.

D. SA-Enhanced Model

Simulated Annealing was implemented using a temperature decay scheme (T=1.0 to T=0.001) with 100 iterations. The model produced:

- MSE: 0.0461
- R^2 Score: 0.9442
- Accuracy (±0.3 GPA): 86.64%

SA matched PSO in both MSE and R^2 but slightly exceeded it in accuracy. It offers a simpler, lightweight alternative to swarm-based and population-based search while still delivering strong generalization.

E. Discussion

Despite the baseline XGBoost model already exhibiting high accuracy, metaheuristic feature selection refined it further:

• GA delivered the best overall results, confirming its robustness and search efficiency.

• PSO offered interpretability gains and helped half the students in the sample improve their prediction accuracy.

• SA showed competitive performance with minimal feature reliance, favoring simplicity.

These results highlight the value of hybrid modeling Table 1, merging metaheuristic optimization with gradient-boosted learning in educational analytics applications. In particular, GA and PSO show promise for integration into GPA forecasting systems, academic advising tools, and early risk detection platforms. Recent studies further substantiate these findings. Hakkal et al. [8] demonstrated that optimizing XGBoost parameters through hybrid metaheuristics significantly enhances learner performance prediction accuracy, while Villegas et al. [10] confirmed that ensemble-based models such as XGBoost and CatBoost outperform classical machine learning approaches across multi-factor student datasets. Similarly, Kuntalp et al. [9] found that both GA and PSO consistently produce compact, high-quality feature subsets, strengthening model generalization and interpretability results that align with the present study's GA superiority. In contrast, emerging research debates the universality of metaheuristic superiority. Comparative analyses indicate that model rankings may shift depending on dataset scale, hyperparameter tuning, or the defined fitness objective [9], [11]. Adaptive hybrid variants such as GA–PSO and WOA–PSO have shown improved stability in recent works, suggesting that future studies should explore dynamic or multi-swarm strategies to further enhance convergence [9]. Moreover, Alnasyan et al. [11] emphasized that deep models such as Bi-LSTM and Transformer networks outperform tree ensembles when sequential or temporal data are available, implying that hybrid metaheuristics may be more beneficial for cross-sectional datasets such as the one used here.

Explainability also remains a growing focus. Recent explainable AI (XAI) research integrates SHAP and LIME techniques to provide interpretable insights into academic predictors [12], [13]. Hoq et al. [12] applied SHAP to XGBoost-based student models, confirming that variables like Parental Support and Study Time also significant in this study have the highest contribution to GPA outcomes. Islam et al. [13] similarly stressed that interpretable ensemble models enhance educators' trust and improve intervention strategies. The inclusion of SHAP-based analysis in future extensions of this framework would therefore strengthen the model's transparency and real-world applicability. Finally, computational trade-offs should be noted. Although GA achieved the best performance, it required higher computation time, consistent with previous observations that evolutionary search increases runtime complexity [9], [11]. This underlines the importance of balancing performance gains with efficiency, particularly for large-scale or real-time educational analytics systems. Overall, the integration of recent literature reinforces that combining metaheuristic optimization with ensemble learning, particularly GA- and PSO-enhanced XGBoost, represents a promising and explainable direction for educational data mining. Future research should evaluate these hybrid models across diverse institutions, explore adaptive metaheuristic hybrids, and incorporate explainable AI components to ensure predictive accuracy and interpretability remain balanced in educational practice.

Table 1: Comparative performance metrics for GPA prediction models

Model	MSE	R^2 Score	Accuracy (±0.3 GPA)
Baseline (All Features)	0.0463	0.9440	86.01%
PSO+ XGBoost	0.0461	0.9442	85.18%
GA + XGBoost	**0.0443**	**0.9465**	**87.89%**
SA + XGBoost	0.0461	0.9442	86.64%

Bar plots in (**Figure 7)** confirm these differences visually, showing GA with the highest predictive power. Notably, all metaheuristics achieved either comparable or superior performance to the baseline, while also reducing feature count.

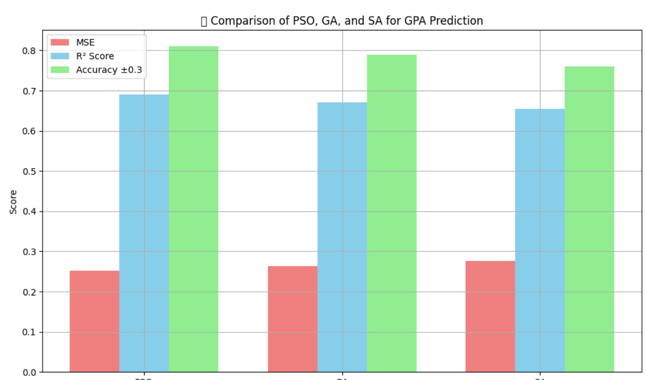

Figure 7: Comparison of PSO, GA, and SA in terms of MSE, R^2, and accuracy (within ±0.3 GPA)

VI. CHALLENGES AND LIMITATIONS

Despite the promising results achieved through integrating metaheuristic optimization with XGBoost for GPA prediction, several challenges and limitations emerged throughout the research process.

A. Challenges

Feature redundancy and irrelevance were among the most prominent issues. Although the dataset contained a wide range of behavioral, academic, and demographic features, several exhibited weak or non-significant correlations with GPA. This diluted the predictive signal and increased the risk of overfitting, making feature selection essential. Metaheuristic algorithm tuning presented another technical challenge. The effectiveness of PSO, GA, and SA depends heavily on their respective control parameters (e.g., particle size, mutation rate, temperature schedule). Determining the appropriate configuration to ensure convergence without falling into local optima required extensive experimentation and validation. A further challenge lies in achieving performance gains over a strong baseline. Since the XGBoost model trained on all features already delivered high predictive accuracy ($R^2 = 0.9440$, Accuracy = 86.01%), improvements via feature selection were necessarily incremental. Demonstrating value beyond numeric gains required additional visualizations and per-student accuracy assessments. Balancing interpretability with complexity was another trade-off. While metaheuristic-selected features enhanced model compactness, the selection logic remained opaque. Differences in selected subsets across algorithms introduced variability that complicates transparent interpretation, especially in educational settings where explainability is vital. Finally, scalability and generalizability remain open challenges. The current implementation was tested on a single-institution dataset. Scaling to broader datasets across schools or regions would introduce new complexities in feature distributions, cultural factors, and labeling consistency.

B. Limitations

This study is subject to several limitations. First, it relied on a single dataset, which may not capture the variability present across different educational contexts. Broader validation across multiple institutions is required to assess generalizability. Second, XGBoost hyperparameters were held constant during model comparisons to isolate the impact of feature selection. While this ensured experimental control, it potentially limited the absolute performance of each optimized model. Third, the dataset contained no temporal or longitudinal features. Modeling trends over time, such as changes in attendance, engagement, or academic performance, could enable richer, more personalized predictions. Fourth, although the study emphasized accuracy, post-hoc interpretability techniques such as SHAP or LIME were not applied. These tools could help educators understand feature-level influence and justify predictions in real-world applications. Lastly, metaheuristic optimization is computationally intensive, especially on high-dimensional data. Practical deployment would require efficiency improvements or approximations for real-time use in student analytics systems.

VII. CONCLUSION AND FUTURE WORK

This study explored the integration of metaheuristic optimization techniques, Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Simulated Annealing (SA)with the XGBoost regression model for predicting student GPA based on behavioral, demographic, and academic features. The results demonstrated that all three algorithms significantly enhanced model performance compared to the baseline (no feature selection), with GA achieving the best results across all evaluation metrics: MSE = 0.0443, $R^2 = 0.9465$, and prediction accuracy within ±0.3 GPA = 87.89%. PSO also exhibited competitive performance, improving predictions for 50% of the students in a subset analysis, highlighting its practical efficacy. In addition to quantitative improvements, the visual analytics, such as correlation heatmaps, GPA distributions, and prediction accuracy plots, reinforced the relevance of specific features like parental support and weekly study time in GPA outcomes. These findings support the viability of metaheuristic-guided feature selection in enhancing predictive models within educational data mining. Future work could build upon these findings in several ways. Incorporating temporal features, such as attendance logs or cumulative performance indicators, may enhance the model's ability to capture longitudinal patterns. The integration of deep learning techniques, such as Long Short-Term Memory (LSTM) networks or Transformer-based models, alongside metaheuristic feature selectors, could provide deeper insights into feature interactions. Further validation through cross-institutional datasets is recommended to assess the generalizability of the approach. Lastly, embedding interpretability frameworks like SHAP or LIME would improve transparency and foster trust in the model's predictions among educators and administrators.

Data Availability Statement: The data used to support this study are available in a public repository. https://archive.ics.uci.edu/dataset/320/student+performance

Code availability: The code used to implement the proposed model experiments is publicly available on GitHub: https://github.com/HindAlmaaz/student-XGBoost

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study

VIII. References

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

[2] B. Chandra, M. Gupta, and S. Ghosh, "Feature selection using metaheuristic optimization: A review," Int. J. Inf. Technol., 2020.

[3] T. Velmurugan and C. Anuradha, "Performance evaluation of feature selection algorithms in educational data mining," Int. J. Data Mining Techniques Appl., vol. 5, no. 2, pp. 131–139, 2016.

[4] Z. Maryam et al., "Role of FCBF feature selection in educational data mining," Mehran Univ. Res. J. Eng. Technol., vol. 39, no. 4, pp. 772–778, 2020.

[5] I. Syarif, A. Prugel-Bennett, and G. B. Wills, "Feature selection of network intrusion data using genetic algorithm and particle swarm optimization," EMITTER Int. J. Eng. Technol., vol. 4, no. 2, pp. 277–290, 2016.

[6] J. M. Port, "Dual heuristic feature selection based on genetic algorithm and binary particle swarm optimization," J. Univ. of Babylon, vol. 27, no. 1, pp. 1–12, 2019.

[7] R. S. Regha and D. U. Rani, "Optimization feature selection for classifying student in educational data mining," Int. J. Comput. Appl., vol. 183, no. 1, pp. 1–5, 2021.

[8] S. Hakkal et al., "XGBoost to enhance learner performance prediction," Discover Education, 2024.

[9] D. G. Kuntalp et al., "A comparative study of metaheuristic feature selection for student performance prediction," Applied Sciences, 2024.

[10] W. Villegas et al., "Machine learning models for academic performance forecasting: A 2025 comparative analysis," Frontiers in Education, 2025.

[11] B. Alnasyan et al., "The power of deep learning techniques in educational analytics," Discover Education, 2024.

[12] M. Hoq et al., "Analysis of an explainable student performance prediction model using SHAP," EDM 2023 Proceedings, 2023.

[13] M. M. Islam et al., "Integration of explainable AI in educational data mining: Challenges and future trends," Discover Education, 2025.