

A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction

Minakshi Sharma ^{1,✉}, Vibhakar Mansotra ¹

1. Computer Science and IT, University of Jammu, Jammu, IND

Received: July 17, 2026 | Review began: August 03, 2026 | Review ended: September 08, 2026 | Published: September 09, 2026

© Copyright 2026

This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Educational institutions face substantial challenges in improving student performance and academic quality while efficiently managing the massive volumes of data generated through learning management systems, student interactions, and assessments. The fusion of Learning Analytics with Educational Data Mining facilitates the extraction of valuable insights from these diverse educational datasets, enabling informed decision-making and the preventive identification of at-risk students. Machine Learning and Deep Learning approaches have evolved as effective tools for forecasting academic performance, student engagement, and dropout risk. A Genetic Algorithm (GA) is implemented for feature selection to determine the ideal subset of features prior to model development. The efficacy of conventional machine learning models, namely Random Forest and Extreme Gradient Boosting (XGBoost), is compared with that of a proposed hybrid Autoencoder-Multilayer Perceptron (MLP) model. The experimental findings indicate that the proposed hybrid model performs better than the baseline models in terms of predictive accuracy, confirming its potential as an effective method for predicting student performance. The proposed Hybrid Genetic Algorithm-Autoencoder-MLP (HGA-AMLMP) framework accomplished an overall prediction accuracy of 93% on the held-out test set, outperforming the baseline Random Forest (90%) and XGBoost (91%) models. The outcomes affirm the effectiveness of the proposed hybrid framework in providing useful predictive information for recognizing students who may be at academic risk. Student outcomes were formulated as a binary classification problem, where students with a final grade $G3 \geq 10$ were classified as Pass, and those with $G3 < 10$ were classified as Fail. However, the results are limited to the evaluated datasets and experimental setting, and further validation on independent educational datasets is required to establish generalizability.

Categories: Ensemble Learning, AI applications, AI/ML-based decision support systems

Keywords: educational data mining, learning analytics, random forest, xgboost, multi-layer perceptron, autoencoder, genetic algorithm

Introduction

Educational institutions generate enormous amounts of data through learning management systems, attendance records, online assessments, academic evaluations, and student interactions. Effectively evaluating these data has become increasingly important for improving student success, reducing dropout rates, and boosting academic quality [1]. Educational analytics, aided by learning analytics and educational data mining, helps academic institutions assess meaningful patterns and construct predictive models that facilitate informed decision-making and timely interventions [2].

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

Current advancements in artificial intelligence, notably machine learning and deep learning, have substantially improved the ability to predict student academic success. Traditional machine learning models, such as Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), and Extreme Gradient Boosting (XGBoost), have produced promising results in detecting students who are at academic risk. Nowadays, deep learning frameworks, including Artificial Neural Networks, Autoencoders, Transformer-based models, and Long Short-Term Memory networks, have been used to identify complex nonlinear relationships within educational datasets. Apart from these developments, the expected outcomes of such models are highly dependent on data quality, feature representation, and the existence of redundant or irrelevant features.

Feature selection plays an essential role in improving predictive accuracy by identifying the most informative attributes while reducing data dimensionality and computational complexity. Metaheuristic optimization techniques, particularly Genetic Algorithms (GAs), have proven effective in selecting optimal feature subsets by exploring a large search space and eliminating redundant features [3]. The use of optimized feature subsets not only improves model performance but also minimizes training time and enhances model generalization.

Considering the most widely used models, RF and XGBoost have consistently shown strong predictive power for educational statistics due to their robustness and competence in regulating complex feature interactions [4]. However, acquiring highly nonlinear latent representations can potentially be difficult for such models from high-dimensional educational datasets. Autoencoders offer an efficient solution by learning compact and useful representations of features, which can consequently improve the performance of downstream neural network classifiers. The integration of an Autoencoder with a Multilayer Perceptron (MLP) enables efficient feature learning while also enhancing classification performance.

Inspired by these observations, this research proposes a Hybrid Genetic Algorithm-Autoencoder-MLP (HGA-AMLMP) framework for student performance prediction. The GA is deployed to evaluate the optimal subset of features before model development, thereby reducing feature redundancy and boosting computational efficacy. The successful implementation of the proposed hybrid Autoencoder-MLP model is assessed in relation to two prominent baseline models, namely RF and XGBoost. The findings reveal that the proposed hybrid model achieves an overall prediction accuracy of 93%, outperforming XGBoost and RF, thereby showing its effectiveness for educational and learning analytics.

The major accomplishments of the present study are stated as follows:

- A GA-based feature selection approach is utilized to estimate the most informative features from the educational dataset.
- Baseline machine learning methods, namely RF and XGBoost, are implemented for comparative assessment.
- A novel hybrid Autoencoder-MLP (HGA-AMLMP) model is proposed to enhance student performance prediction by learning robust, undiscovered feature representations.
- A thorough comparative investigation signifies the excellence of the proposed hybrid model over traditional machine learning approaches.

The proposed framework offers an effective and scalable solution for identifying students who may be at academic risk at an intermediate stage of their academic progression and provides predictive information that may assist educators in identifying students who could benefit from further academic support.

In this study, the prediction setting assumes that previously available academic indicators, including G1 and G2, may be available at the time of prediction. Therefore, the proposed framework should be interpreted as an intermediate-stage academic-risk prediction approach rather than prediction before any prior academic assessment.

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

Related study

Student performance prediction has emerged as a prominent field of study with respect to learning analytics and educational data mining, aiming to facilitate the early identification of students who are at risk and support educational decision-making. The review analyzes numerous studies that demonstrate the extensive use of machine learning and deep learning techniques for academic performance prediction, prior identification of at-risk learners, and engagement forecasting.

The literature demonstrates the growing adoption of machine learning and deep learning methodologies for student performance prediction and educational analytics. Several studies have employed traditional machine learning algorithms, such as DT, SVM, RF, Logistic Regression, K-Nearest Neighbors, and XGBoost, to identify at-risk students and predict academic achievement with promising accuracy [5-7].

In one of the studies, the author examined the academic dataset of 50 master's students by considering demographic characteristics, previous academic achievements, and course-level grades. Among all the evaluated classification models, Linear Discriminant Analysis attained the highest predictive accuracy of 79% and 77%, outperforming SVM, which demonstrated sensitivity to class imbalance [8]. Berhanu and Abera [9] applied a DT algorithm to a dataset consisting of 199 preprocessed records and 49 attributes. The model, built with the RapidMiner (RM) tool, produced 27 classification rules. The study proposed by Oyediji et al. [10] assessed the performance of three different predictive models: Neural Network, Linear Regression integrated with deep learning, and supervised Linear Regression. The findings offered valuable insights into the comparative efficiency and suitability of these techniques for educational performance prediction [10]. Ismail et al. [11] explored student performance prediction using various machine learning algorithms, such as SVM, RF, DT, Artificial Neural Networks, and Naïve Bayes. This research highlighted the effectiveness of these approaches in improving predictive accuracy in education [11].

The research carried out a comparative evaluation of multiple machine learning classification models using academic data from 1,044 students across two high schools and found that XGBoost achieved superior performance in predicting student academic outcomes based on accuracy, F-measure, and execution time [12].

Dervenis et al. [13] evaluated machine learning approaches on the UCI Student Performance Dataset, where RF showed great predictive power with an area under the curve (AUC) of 90.5%, reflecting high predictive effectiveness, but exhibited lower accuracy when extended to multi-class grade prediction. Roy and Farid's [14] research study introduced Adaptive Feature Selection Algorithm to optimize student performance prediction through effective feature selection and redundancy reduction. Using SVM, K-Nearest Neighbors, and DT classifiers, the approach delivered high accuracy while requiring fewer features [14]. Kalita et al. [15] proposed a Bidirectional Long Short-Term Memory-based Grade Point Average prediction framework that achieved 88.23% accuracy, surpassing conventional machine learning approaches, while SHapley Additive exPlanations analysis highlighted the most influential factors affecting student performance. In another study, an ensemble framework combining two deep learning methodologies, namely MLP and Dual Attention Network (DANet), was evaluated on the OULAD dataset of approximately 32,000 students. The results showed superior performance, achieving 96.20% accuracy compared to standalone MLP (95.86%) and DANet (95.72%), illustrating the effectiveness of hybrid deep learning approaches in educational analytics [16]. Table 1 highlights a comparative review of previously published studies. The comparison highlights the dataset, methodology, feature selection methods, predictive models, and their reported outcomes.

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

Model	Key Features	Accuracy	Limitations
RF, SVM, LR, NB, KNN [17]	Predicted final exam grades using midterm scores of 1,854 undergraduate students; comparative analysis of six ML models	The proposed model achieved an accuracy of 70%–75%	Moderate performance; limited by dependence on midterm scores as primary predictors
Deep Artificial Neural Network [18]	Utilized virtual learning environment (VLE) clickstream dataset to identify at-risk students; compared with LR and SVM	ANN achieved an accuracy of 84%–93%	Performance highly dependent on the quality of feature engineering
DT, RF, SVM, XGBoost [19]	Indonesian national admission (SNMPTN) and GPA records to predict academic performance; comparative evaluation of four classifiers	XGBoost achieved an accuracy of 77%	Moderate predictive strength; with challenges in handling class imbalance
SVM, DT, NB, KNN [20]	Used K-means clustering for feature extraction and parameter optimization; compared four ML classifiers	SVM outperformed the other models, achieving an accuracy of 96%	NB showed weaker performance; results may depend on dataset characteristics
Heterogeneous Ensemble (RF + KNN + A1DE) [21]	Combines RF, KNN, and A1DE using majority voting; evaluated on 300 postgraduate CS students and applied 10-fold cross-validation	The ensemble model achieved an accuracy of 96.88%	The small sample size may limit the generalizability and robustness of the results across diverse educational settings
Ensemble ML Framework (LR, RF, XGBoost, MLP) with SHAP & LIME [22]	Integrates seven educational datasets (demographics, registration, assessments, and VLE engagement); employs ensemble machine learning with SHAP and LIME	XGBoost achieved an accuracy of 95.1%	The study is limited to structured online learning data and does not incorporate temporal deep learning models or multimodal educational data
LR, RF, AdaBoost, Gradient Boosting, DT, NB, Neural Network [23]	Utilized survey-based personal, socio-economic, family, and academic data (35 features) to predict early school dropout	DT with SMOTE showed the best dropout detection: Recall = 85.7% and F1-score = 37.5%	The dataset was collected from a single vocational high school and is affected by class imbalance, several zero-filled features, and limited generalizability to other educational settings
Evaluated 15 ML classifiers: XGBoost, Light GBM, SVM, AdaBoost, KNN, NB, etc. [24]	Developed an early warning and recommendation system	The Ensemble (Boosting + Grid) model showed the best	Performance was evaluated using a single institutional dataset; temporal prediction

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

	using institutional, academic, and demographic data	performance: Recall = 0.92 and AUC = 0.92	relied primarily on structured academic records
--	---	---	---

TABLE 1: A comparison of current models with our proposed ensemble architecture

A1DE, Averaged One-Dependence Estimators; AUC, Area Under the Curve; DT, Decision Tree; GBM, Gradient Boosting Machine; KNN, K-Nearest Neighbors; LR, Logistic Regression; LIME, Local Interpretable Model-Agnostic Explanations; ML, Machine Learning; MLP, Multilayer Perceptron; NB, Naive Bayes; RF, Random Forest; SHAP, SHapley Additive exPlanations; SMOTE, Synthetic Minority Over-sampling Technique; SVM, Support Vector Machine; XGBoost, Extreme Gradient Boosting

Materials And Methods

Dataset

The present research examines the application of predictive modeling in educational settings, with a particular focus on the two related Student Performance datasets from the UCI Machine Learning Repository: the Mathematics (student-mat.csv) dataset, containing 395 student records, and the Portuguese Language (student-por.csv) dataset, containing 649 student records. Both datasets contain the same demographic, social, family, and academic attributes but correspond to different school subjects. These attributes provide a comprehensive basis for developing and evaluating predictive models for student performance. In this research, the two datasets were merged to create a unified dataset comprising 1,044 observations and 33 predictor variables before feature selection. The final academic outcome was converted into a binary target variable. Students with a final grade $G3 \geq 10$ were classified as Pass, and those with $G3 < 10$ were classified as Fail.

Prior to model development, the merged dataset underwent preprocessing stages, including data cleaning, encoding categorical variables, handling missing or inconsistent values (if any), feature transformation, and normalization. The dataset was first divided into training and held-out test subsets with `random_state = 42` by applying a stratified 75:25 split. The test set was isolated before performing feature selection and model training and was not used during feature-selection optimization, Autoencoder training, hyperparameter tuning, or model fitting.

Data Preprocessing

Preceding algorithmic deployment, the raw empirical dataset went through a systematic data preprocessing pipeline to ensure structural integrity, eliminate statistical noise, and standardize heterogeneous variable scales. Real-life academic information is inherently susceptible to missing values, high-cardinality categorical biases, and diverse measurement scales. These anomalies distort gradient descent optimization if left unmitigated, induce premature convergence, and compromise classifier generalizability. The preprocessing framework executed three sequential operations:

Data sanitation and imputation: The provided input matrix was first examined for structural irregularities and missing data. Rows containing a significant percentage of missing data fields were systematically discarded, while minimal missing elements in continuous educational academic metrics were resolved using median imputation to preserve sample size without altering the basic feature distribution. Outliers within performance-tracking metrics - identified via Interquartile Range filtering - were capped at the distribution boundaries to prevent extreme variance anomalies from disrupting downstream weights.

High-cardinality target encoding: There were numerous high-cardinality nominal features in the dataset, such as institutional identifiers representing the unique schools within the district. Utilizing traditional one-hot encoding on these attributes would result in a dimensionality explosion, leading to excessive matrix sparsity and the curse of dimensionality. In an effort to prevent this, a Target Encoding protocol was implemented. This technique replaces each categorical string value with the expected average value of the target variable computed for that specific category, mapping nominal features into continuous numeric spaces without extending feature-set dimensionality.

How to cite this article:

Z-score standardization (StandardScaler): To facilitate the deep representation training layers and minimize distance-based optimization bias, all continuous parameters were mapped to a uniform mathematical scale. Each continuous feature matrix was transformed using the standard normalization formula shown in Equation (1):

$$z = (x - \mu) / \sigma(a) \tag{1}$$

where feature mean is represented by μ , and σ denotes the standard deviation. By this operation, the data distributions are close to a mean of zero with one-standard deviation ($\mu = 0, \sigma = 1$), which accelerates model convergence and stabilizes gradient propagation across the entire framework.

GA, recursive feature elimination, information gain, and chi-square-based feature selection

Leakage-Safe Feature Selection

To prevent information leakage, the held-out test set was isolated before feature selection. The GA was therefore applied exclusively to the training data. The fitness function evaluated candidate feature subsets using only the training data, while the test data remained completely unseen during the optimization process.

To assess the key attributes influencing academic performance, GA, recursive feature elimination, information gain, and chi-square-based feature selection techniques were implemented for feature selection. To maintain a standardized evaluation framework, each technique was restricted to extracting the top 19 features. Results in Table 2 highlight that the GA-selected feature subset achieved the highest cross-validation accuracy at 89.22%, surpassing information gain (88.86% ± 0.61%), chi-square (88.74% ± 1.33%), and recursive feature elimination (87.78% ± 1.11%). Based on these findings, the GA-selected feature subset was used as the foundation for the proposed Autoencoder-MLP framework.

Feature Selection Method	Selected Features	Cross-Validation Accuracy (%)
Recursive Feature Elimination	19	87.78 ± 1.11
Chi-square (χ^2)	19	88.74 ± 1.33
Information Gain	19	88.86 ± 0.61
Genetic Algorithm	19	89.22

TABLE 2: Performance comparison of feature selection techniques

GA Feature Optimization

The GA repeatedly searched for an optimal subset of features by modeling development processes such as selection, crossover, and mutation. A fitness function was implemented to determine candidate feature subsets based on classification performance, enabling the algorithm to identify a subset of informative attributes. This feature selection step reduced the dimensionality of the dataset by removing unnecessary and redundant features and preserving practical information for representation learning [25].

Figure 7 illustrates a GA-based feature selection pipeline, where the raw student dataset comprising 33 attributes is first passed to the data preprocessing stage, including data cleaning, feature scaling, and feature encoding. After this, the GA obtains the final preprocessed feature subset as input and searches for an informative subset (19 significant features) by

How to cite this article:

evaluating possible feature combinations. The optimization process chose 19 significant features, including demographic, behavioral, and academic indicators, namely 'school', 'Pstatus', 'Fedu', 'sex', 'address', 'Mjob', 'Fjob', 'reason', 'absences', 'guardian', 'famrel', 'freetime', 'goout', 'Dalc', 'failures', 'romantic', 'Walc', 'G1', and 'G2'.

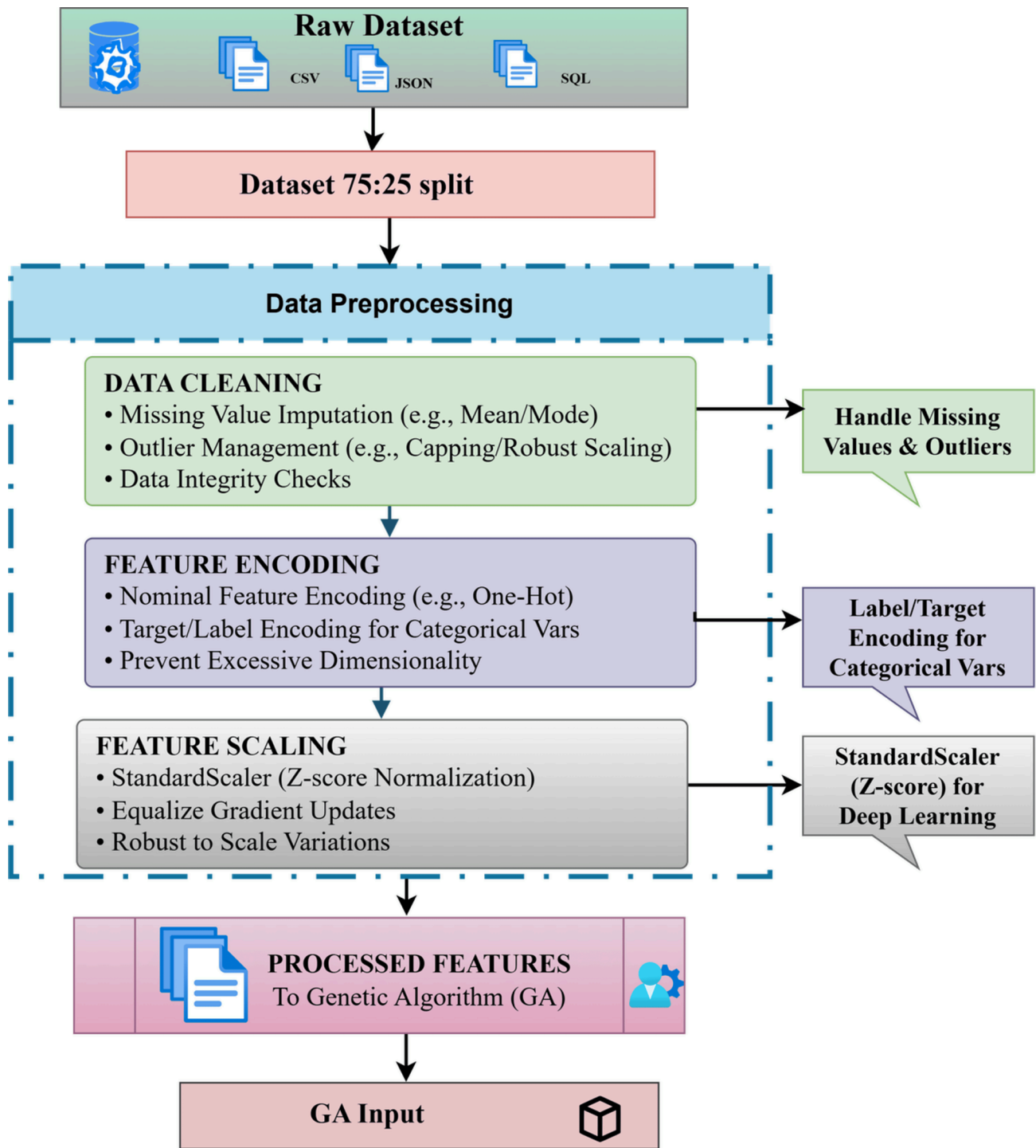


FIGURE 1: Evolutionary workflow of the genetic algorithm-based feature selection

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

The reduction from 33 to 19 features resulted in a more compact predictive model while maintaining high classification performance. This illustrates the capability of GAs to eliminate redundant and less informative variables, thereby improving model interpretability and computational efficacy.

The convergence curve of the GA demonstrated in Figure 2 displays a quick improvement in the objective function during the initial iterations, and after reaching the highest possible number of generations, the GA converged without further improvement. The optimal feature subset achieved a cross-validation accuracy of 89.5%, emphasizing that the search process had stabilized and identified a near-optimal solution. This action solidifies the effectiveness of the GA in effectively inspecting the search space and identifying an optimal subset of features for predicting students' performance. GA is used to reduce dimensionality, remove irrelevant and redundant features, improve model generalization and reduce computational complexity. In short, GA identifies the most influential student-related factors affecting academic performance.

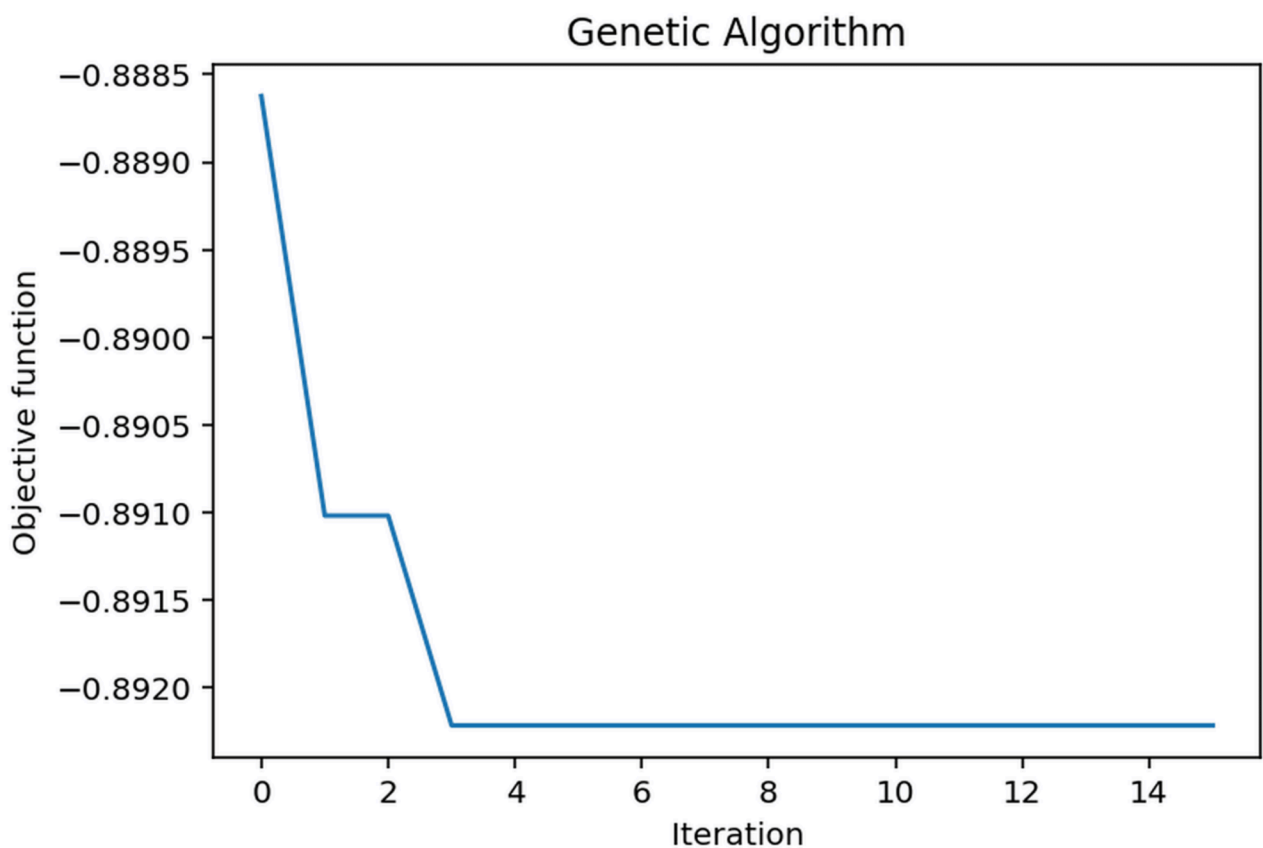


FIGURE 2: Convergence curve of the genetic algorithm

Methodology

The proposed methodological framework develops a strict comparative framework to assess the predictive capability of traditional tree-based ensemble pipelines against a deep implicit representation architecture for learning. The complete workflow is designed to map varied student indicators into distinct academic performance classes while mitigating class asymmetry and data noise. Figure 3 depicts the conceptual process of the suggested deep learning architecture. The workflow outlines the complete sequence of operations, beginning with data collection and preprocessing, and then proceeding to feature engineering and feature selection using the GA. The subset of optimized features is then supplied

to the machine learning and deep learning algorithms for training, validation, and performance evaluation. The workflow ensures a clear understanding of the overall architecture and data flow employed throughout the proposed deep framework.

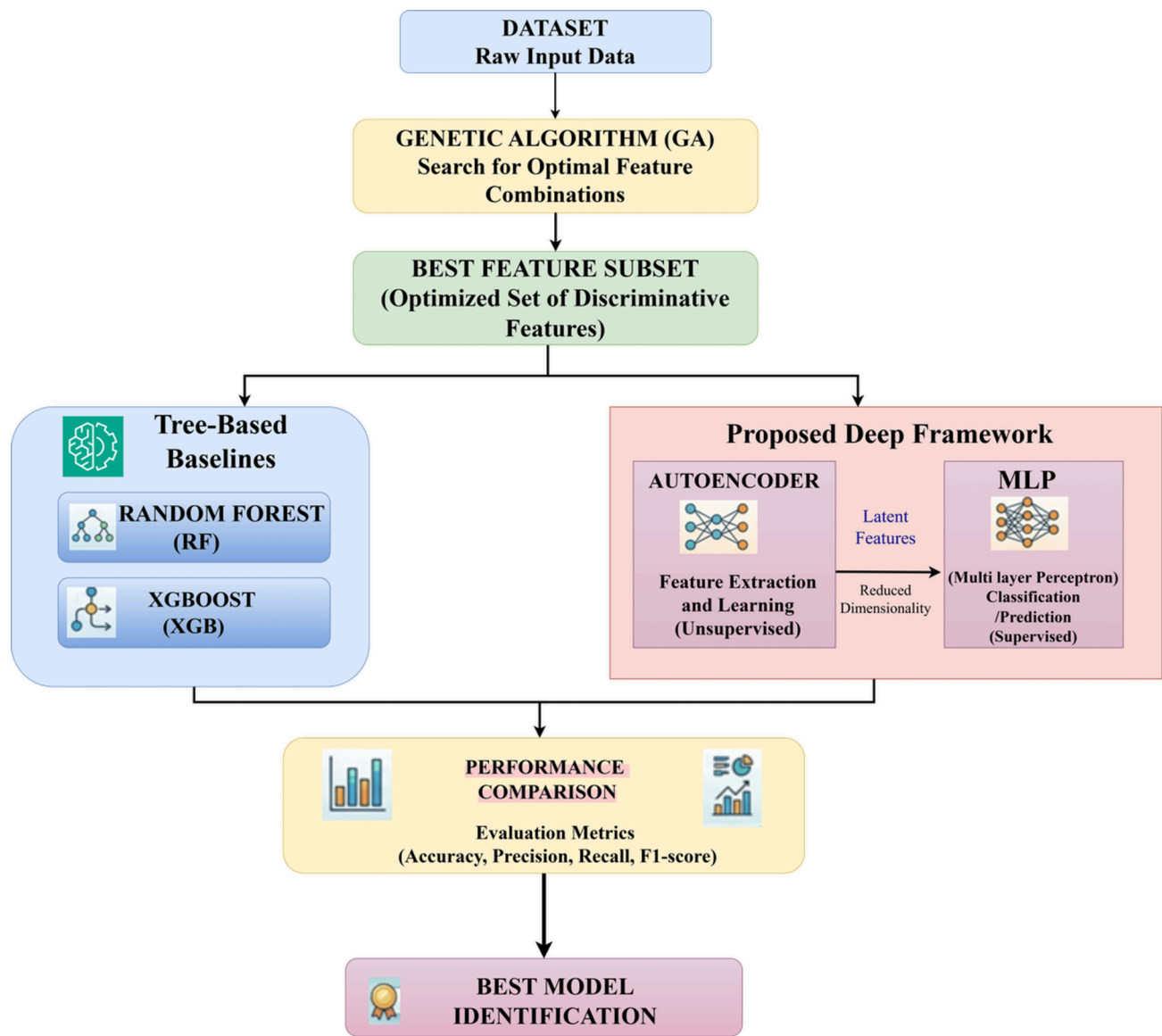


FIGURE 3: Conceptual workflow of the GA-optimized model

GA, Genetic Algorithm

Baseline Paradigm: Tree-Based Baselines

To establish a robust algorithmic benchmark, two state-of-the-art tree-based architectures are employed to operate natively on the preprocessed feature space:

Random forest: This methodology leverages a bootstrap aggregating (bagging) methodology, constructing a large forest of uncorrelated decision trees trained in parallel. By adding feature randomness at each node split, the RF methodology reduces model variance and prevents overfitting to dominant institutional indicators. The binary classification problem was formulated by categorizing students into Pass and Fail classes based on their final academic performance. The dataset contained 1,044 student records, of which 814 students (77.97%) belonged to the Pass category and 230 students

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

(22.03%) belonged to the Fail category. The RF model was evaluated using a stratified train-test split, where 75% of the data was utilized for training and the remaining 25% for testing. This resulted in 783 training and 261 test samples, which were utilized to generate the confusion matrix and classification metrics.

RF classification performance: The RF baseline was trained using the 19 influential features selected through the GA-based feature selection process and attained a total accuracy of 90%, with weighted precision, recall, and F1-score of 90%, 90%, and 90%, respectively, as presented in Table 3. The macro-average precision of 85%, macro-average recall of 87%, and macro-average F1-score of 86% further indicate that the classifier maintained balanced performance across both classes.

Class	Precision (%)	Recall (%)	F1-Score (%)	Support (n)
Fail	76	81	78	58
Pass	94	93	94	203
Accuracy	-	-	90	261
Macro Average	85	87	86	261
Weighted Average	90	90	90	261

TABLE 3: Classification report of the random forest baseline

The confusion matrix of the RF model is illustrated in Figure 4. The model correctly classified 47 out of 58 Fail instances and 188 out of 203 Pass instances, while misclassifying 11 Fail instances as Pass and 15 Pass instances as Fail. The low count of false positives and false negatives highlights the model's strong classification capability, with a high number of correctly predicted instances and relatively few misclassifications. The confusion matrix indicates that the RF classifier provides reliable classification performance and effectively identifies both at-risk and successful students.

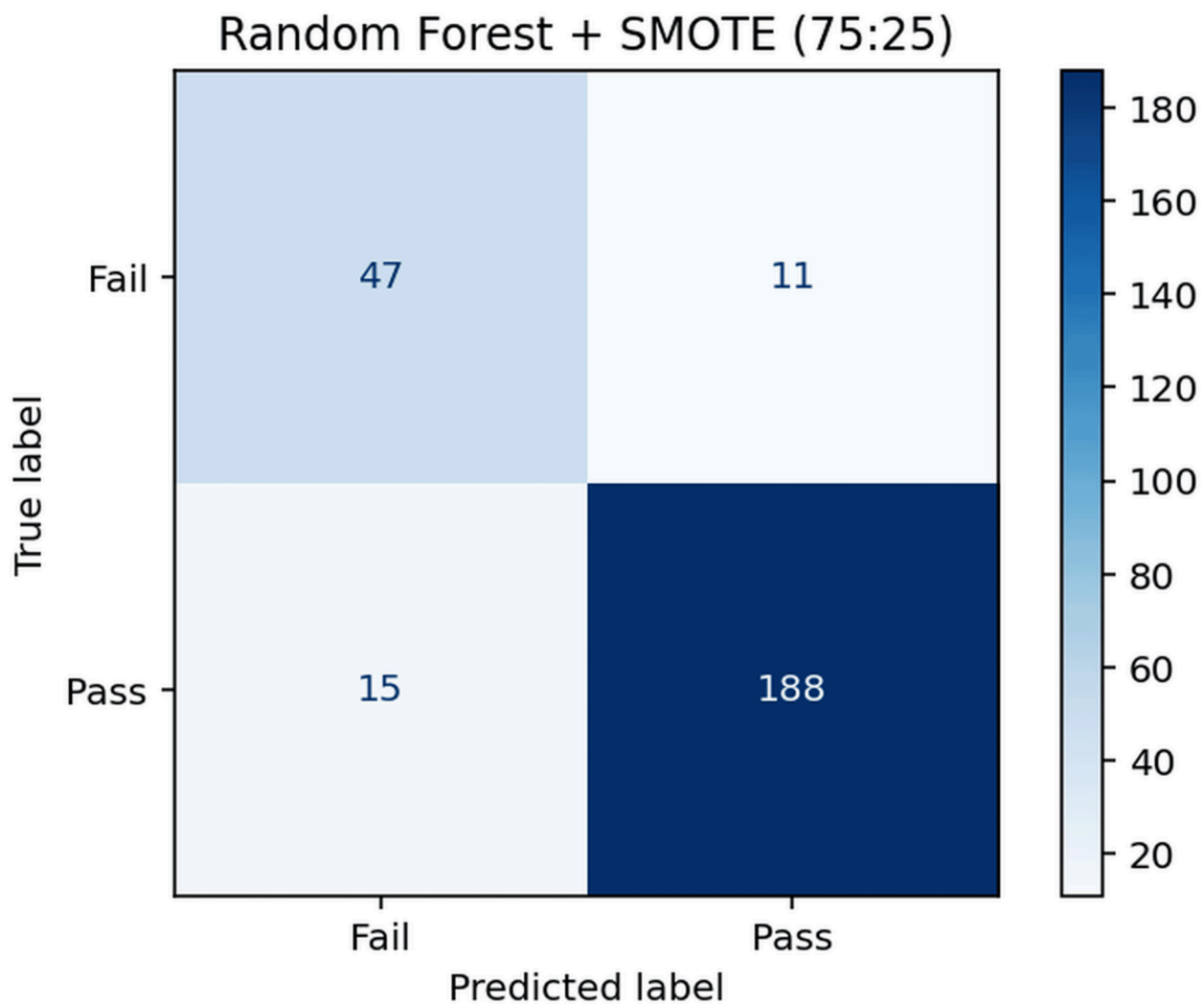


FIGURE 4: Confusion matrix of the random forest classifier

SMOTE, Synthetic Minority Over-sampling Technique

The experimental findings demonstrate that the incorporation of GA-based feature selection with RF baseline provides a robust framework for student academic performance prediction. The selected 19 features were sufficient to achieve high predictive accuracy while reducing model complexity, making the RF baseline suitable for practical educational decision-support systems.

XGBoost baseline: In contrast to parallel bagging, XGBoost executes a sequential gradient boosting paradigm. The model minimizes a regularized objective function by computing second-order Taylor expansions of the loss function. It iteratively builds subsequent DTs to fit the residual errors of preceding iterations, making it highly efficient at capturing intricate, non-linear thresholds within tabular matrices. To determine the most influential attributes of students’ academic success, a GA approach was employed prior to deploying all three models. GA feature selection performs feature evaluation and identifies most relevant attributes, and removing less informative attributes helps reduce noise and enhances the model's generalization capability.

How to cite this article:

The XGBoost model was assessed using a 75:25 train-test split, where the dataset was partitioned into 75% for the training set and the remaining 25% for the test set. The performance metrics of the XGBoost model are presented in Table 4, showing an average accuracy of 91%, precision of 93%, recall of 95%, and an F1-score of 94%. The small number of misclassifications highlights the model's robustness and suitability for detecting students at risk of poor academic performance. Overall, the AUC curve indicates excellent predictive capability and strong class discrimination for this model.

Class	Precision (%)	Recall (%)	F1-Score (%)	Support (%)
Fail	81	76	79	58
Pass	93	95	94	203
Accuracy	-	-	91	261
Macro Average	87	85	86	261
Weighted Average	91	91	91	261

TABLE 4: Classification metrics for the XGBoost model

XGBoost, Extreme Gradient Boosting

The evidence suggests that the GA-selected 19-feature subset retained most of the predictive information while reducing the original feature space from 33 to 19 attributes. This is a logical and convincing conclusion in educational data analytics. The ROC curve, as depicted in Figure 5, shows the strong discriminative capability of the XGBoost model. The model achieved an AUC of 0.9614, demonstrating exceptional performance in distinguishing between pass and fail students. The ROC curve remains near the upper-left corner of the plot, exhibiting a high true positive rate and a low false positive rate across different decision thresholds.

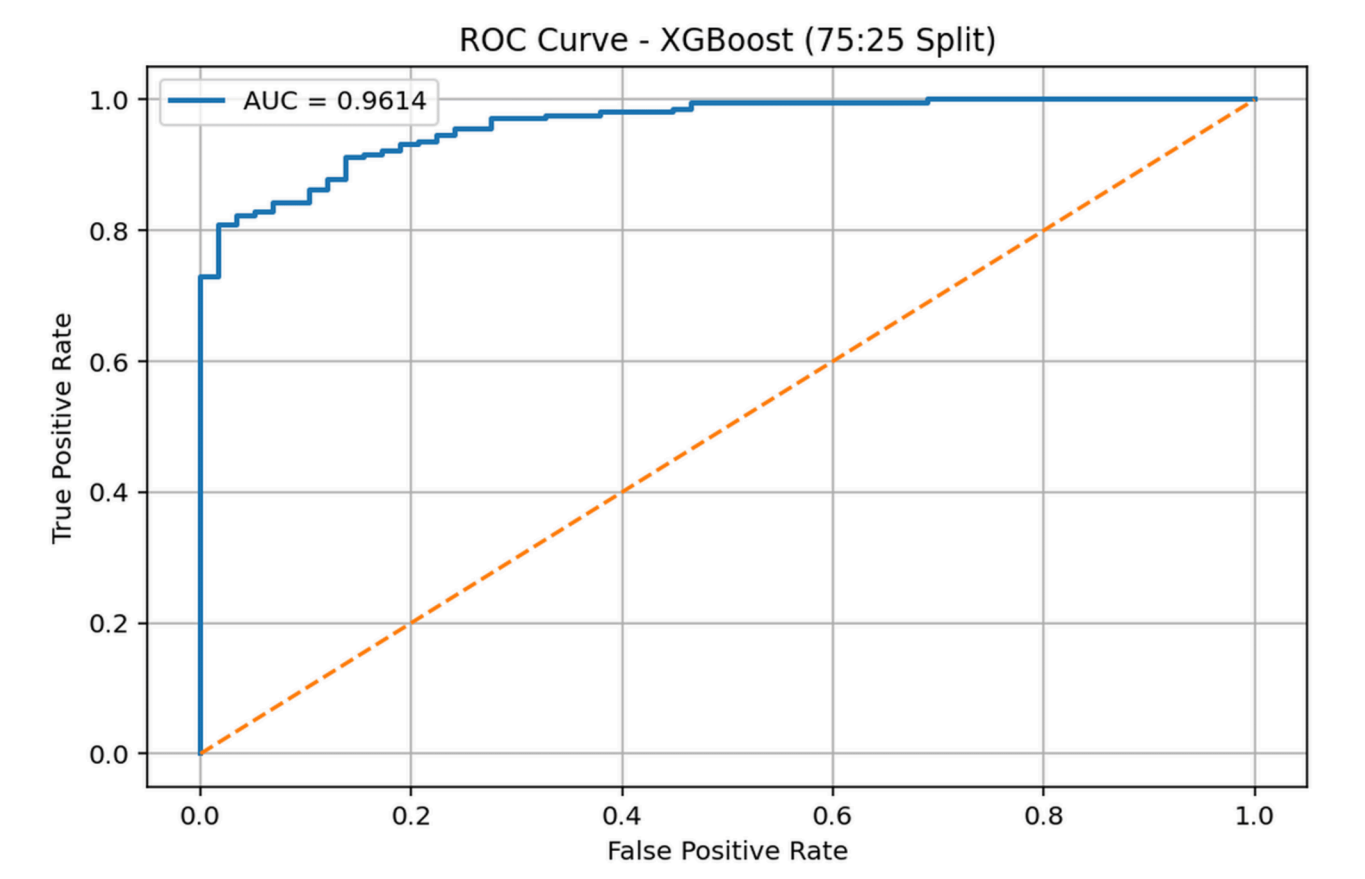


FIGURE 5: ROC curve for the XGBoost model
AUC, Area Under the Curve; ROC, Receiver Operating Characteristic; XGBoost, Extreme Gradient Boosting

The confusion matrix illustrated in Figure 6 highlights the effectiveness of the GA-XGBoost model in predicting student academic outcomes. Out of 261 test instances, the model correctly classified 44 failing students and 193 passing students, while misclassifying only 14 fail students as pass and 10 pass students as fail. These outcomes are consistent with an overall accuracy of 91%.

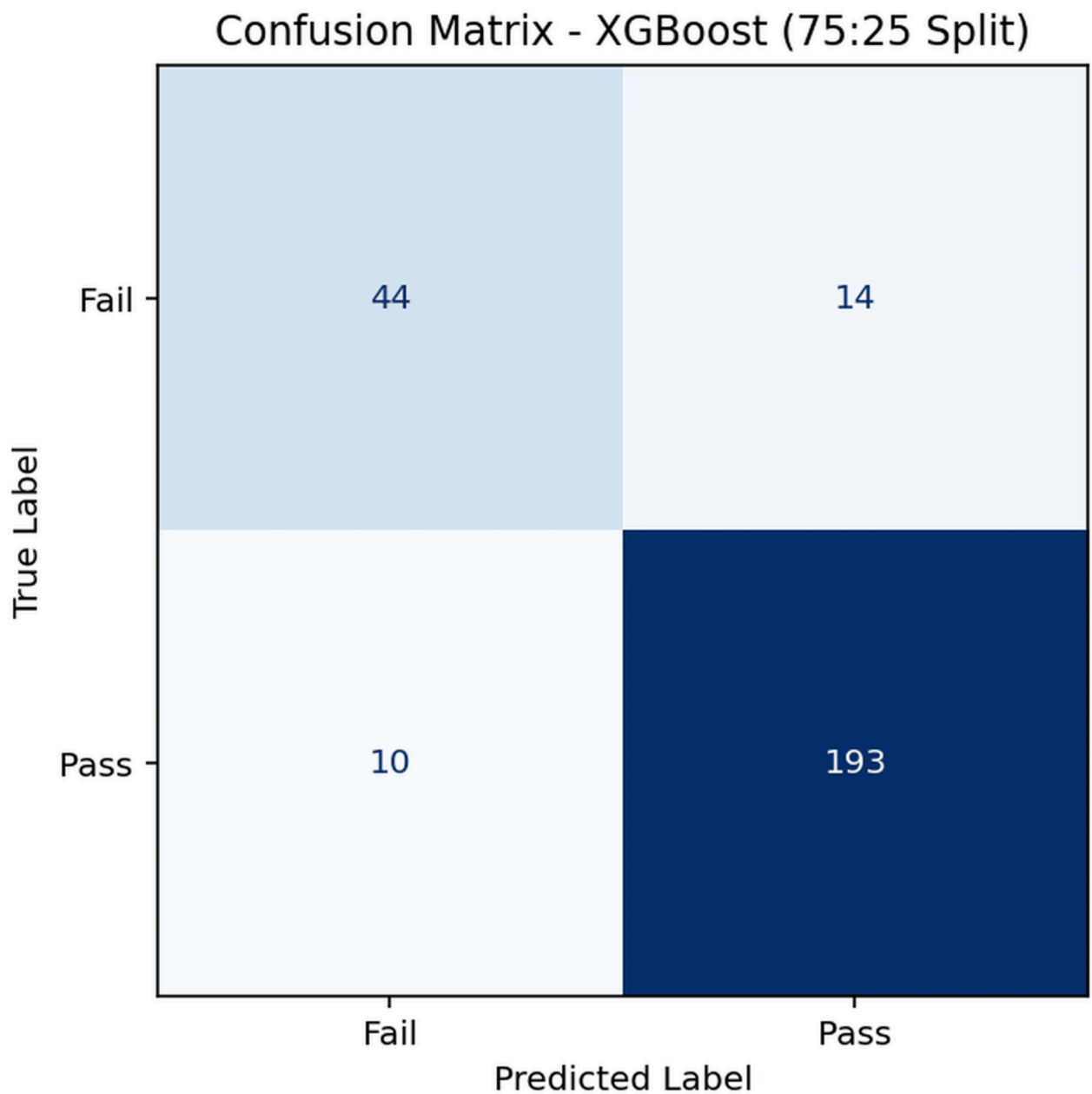


FIGURE 6: Confusion matrix for the XGBoost model

XGBoost, Extreme Gradient Boosting

Proposed Paradigm

Proposed HGA-AMLP framework: As a highly-efficient alternative to tree-based partitioning, this research presents an integrated deep learning framework comprising an unsupervised Autoencoder coupled with a MLP classifier. This framework transitions the predictive task from raw input features into a highly optimized, persistent latent environment.

- Unsupervised Feature Extraction (Autoencoder): The first stage utilizes a symmetric, bottlenecked Autoencoder to execute non-linear dimensionality reduction. The network is enhanced to reduce reconstruction error using Mean Squared Error. This constraint forces the encoder hidden layer to compress the heterogeneous input space into a low-

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

dimensional, high-fidelity latent embedding, effectively filtering out structural noise and uninformative administrative parameters.

- **Deep Latent Classification (MLP):** The dense latent representations generated by the encoder are subsequently consumed by a multilayer feedforward neural network. This classification pipeline uses progressive dense layers mapping to the target vectors. In order to shield the network against variance explosion and prevent an excessive systemic reliance on isolated latent features, stochastic dropout layers and weight-decay regularization penalties are adhered across all hidden layers to restrict feature co-adaptation. The final output layer leverages a sigmoid activation function to estimate the likelihood of each student belonging to the binary Pass/Fail classification.

Evaluation and Paradigm Performance Comparison Protocol

To ensure a mathematically valid comparison and guard against data leakage, all models are trained, tuned, and evaluated. The empirical evaluation establishes a head-to-head paradigm comparison focused on two core dimensions:

1. **Discrete Partitioning vs. Continuous Latent Spaces:** We analyze whether the smooth, continuous latent manifolds synthesized by the Autoencoder-MLP yield superior generalizability compared to the orthogonal, step-wise decision boundaries generated by the tree-based splits of RF and XGBoost.
2. **Robustness to Target Distribution Skews:** Because the student performance dataset exhibits inherent class asymmetry, performance is benchmarked using class-invariant metrics. Beyond overall accuracy, the models are assessed using precision, recall, and macro-F1 score, with particular attention to each model's ability to resist majority-class bias and correctly classify the minority student classes.

Results And Discussion

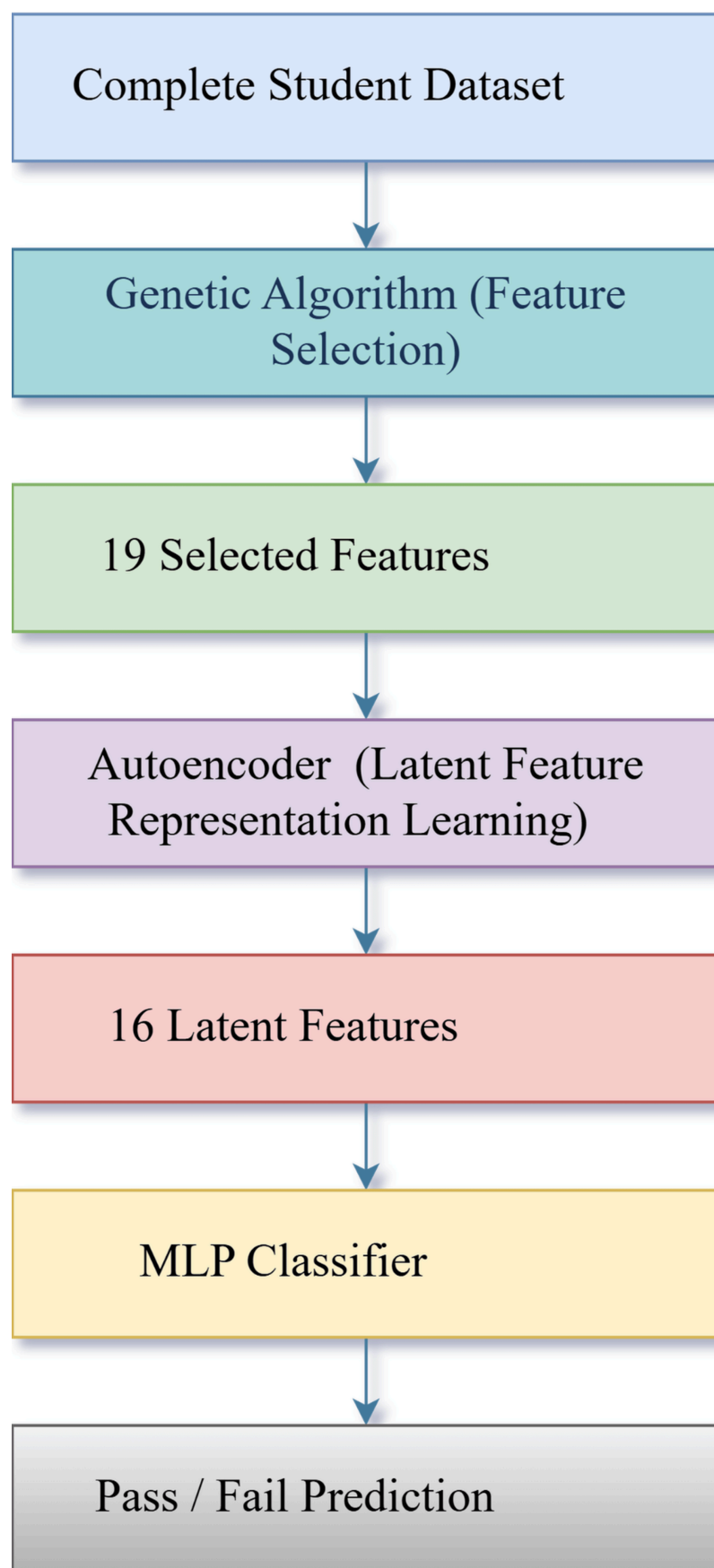
Experimental results and conclusion

The originality of the proposed HGA-AMLP framework exists in the integration of GA-based feature selection, autoencoder-based representation learning, and MLP-based classification within a unified learning analytics pipeline. The autoencoder was used to learn compact latent representations from the selected 19 influential features, which were subsequently provided as input to an autoencoder for unsupervised feature learning. The encoder maps the selected inputs into a compact latent space through multiple hidden layers, thereby capturing essential non-linear patterns among the variables while minimizing redundancy and noise. Simultaneously, the decoder validates feature quality by reconstructing the original input from the latent representation, ensuring that the learned latent features preserved the essential characteristics of the original data to minimize information loss. Finally, these compact latent representations are passed into a MLP classifier.

The proposed pipeline depicted in Figure 7 demonstrated the complete workflow of the GA-Autoencoder-MLP framework. The complete student dataset is first processed by applying a GA for feature selection and selecting the 19 most informative features. Then, these features are encoded into 16-dimensional latent representation, and finally the resulting latent representations are classified by the MLP to predict students' academic performance as Pass or Fail.

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>



How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

FIGURE 7: Proposed hybrid GA–Autoencoder–MLP (HGA-AMLP) pipeline

GA, Genetic Algorithm; MLP, Multilayer Perceptron

The encoded features generated by the trained Autoencoder were used as input to the MLP classifier. The MLP comprised three fully connected hidden layers containing 256, 126, and 64 neurons, respectively, with ReLU activation functions. A single output neuron with sigmoid activation was used to perform binary Pass/Fail classification. The MLP was trained using the Adam optimizer with a learning rate of 0.0005 and binary cross-entropy loss. Training was carried out for a maximum of 200 epochs with a batch size of 128. In order to prevent overfitting by halting training when the validation loss stopped improving, the MLP classifier utilized EarlyStopping during training (monitor = val_loss, patience = 15, restore_best_weights = True). Additionally, ReduceLROnPlateau (factor = 0.2, patience = 5, min_lr = 1×10^{-6}) was utilized implicitly to automatically reduce the learning rate when the validation loss approached a plateau (a point of equilibrium), thereby strengthening the model's generalization capability and boosting the convergence.

The experimental setup utilized 783 training samples and 261 testing samples, with each instance represented by a 16-dimensional latent feature representation generated by the autoencoder. These latent characteristics, or features, were then passed to the MLP classifier for binary student performance prediction, where the MLP uncovered nonlinear relationships among the features and produced the final Pass/Fail classification. The obtained results show that integrating unsupervised feature learning through an autoencoder with supervised classification via an MLP effectively combines evolutionary optimization, unsupervised deep feature learning, and supervised classification to improve the prediction of students' academic success.

The classification results presented in Table 5 show that the proposed HGA-AMLP model achieved an overall accuracy of 93%, along with weighted precision, recall, and F1-score of 93%, indicating robust and balanced classification performance. The high precision value demonstrates that the model effectively minimizes false-positive predictions, while the strong recall value indicates its ability to correctly identify the majority of students belonging to the positive class.

Class	Precision (%)	Recall (%)	F1-Score (%)	Support (%)
Fail	92	76	83	58
Pass	93	98	96	203
Accuracy			93	261
Macro Average	93	87	89	261
Weighted Average	93	93	93	261

TABLE 5: Results of the proposed Autoencoder–MLP (HGA-AMLP) framework

HGA-AMLP, Hybrid Genetic Algorithm-Autoencoder-Multilayer Perceptron; MLP, Multilayer Perceptron

The model achieved a recall of 76% for the Fail class, successfully identifying the majority of at-risk students in the test set, with a precision of 92%. For the Pass class, the framework attained a precision of 93% and a recall of 98%, indicating excellent reliability in predicting successful student outcomes. The weighted F1-score of 93% confirms the overall

How to cite this article:

robustness of the proposed framework across both classes.

The confusion matrix illustrated in Figure 8 shows the effectiveness of the proposed HGA-AMLP framework in predicting students' academic success. Out of 261 test instances, 243 were classified correctly, achieving an overall accuracy of 93%. This is of particular importance in educational analytics, where the early identification of academically at-risk students is a primary objective. These results indicate that the proposed framework is highly suitable for educational decision-making processes and early warning systems.

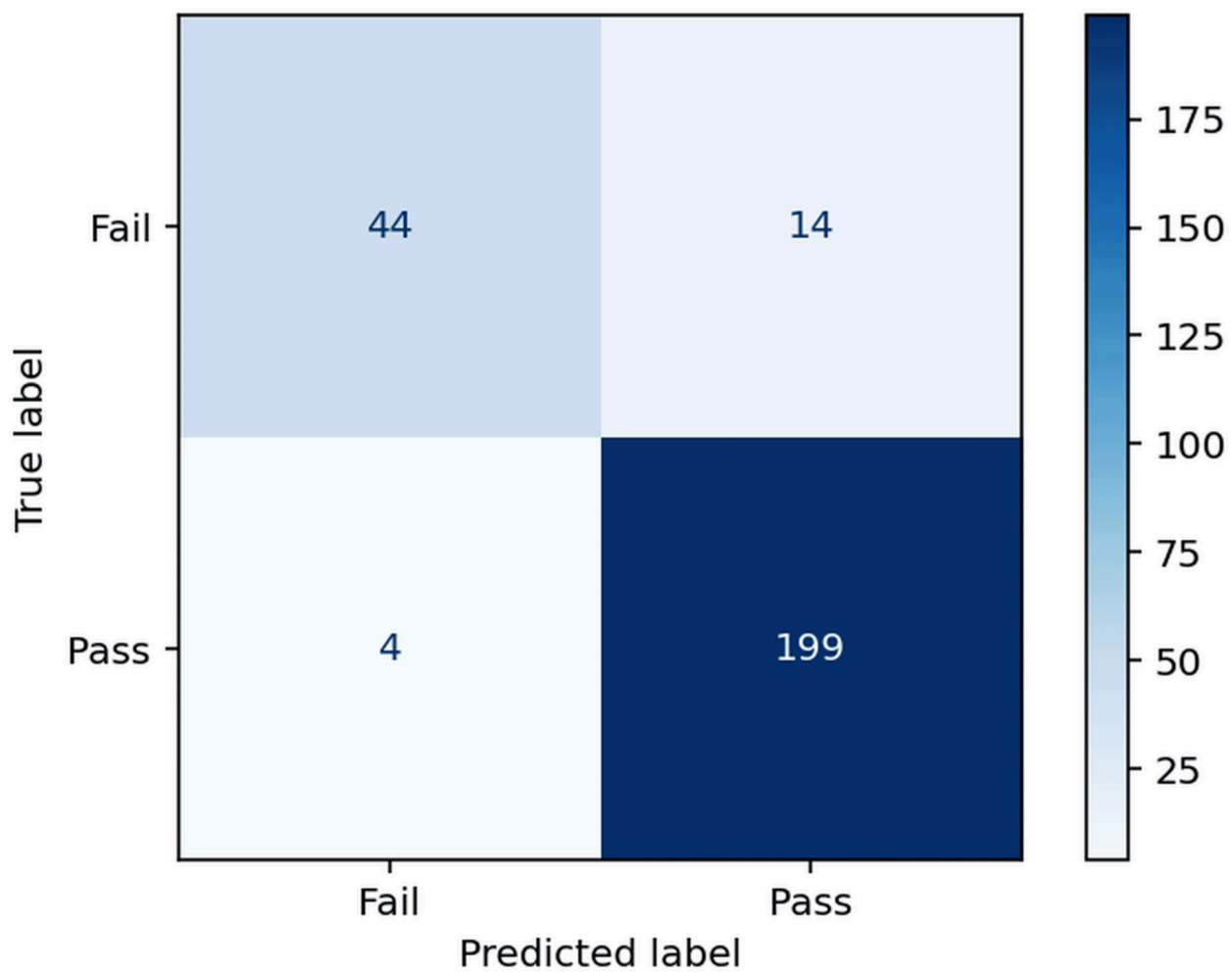


FIGURE 8: Confusion matrix of the proposed Autoencoder-MLP model

MLP, Multilayer Perceptron

Comparative study

Table 6 highlights the comparative performance of the three machine/deep learning models evaluated for predicting students' academic outcomes. Among the evaluated classifiers, our proposed HGA-AMLP framework exceeded the classification performance, with 93% accuracy, and demonstrated its superior ability to learn discriminative feature representations. XGBoost also achieved competitive performance, demonstrating the nearest F1-score (91%), emphasizing a balanced trade-off between precision and recall. Although RF also produced competitive results, its

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

performance was marginally lower than the other two approaches. The outcomes reveal the potential of deep learning techniques for enhancing predictive analytics in educational environments and supporting data-driven academic interventions.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	90	90	90	90
XGBoost	91	91	91	91
Autoencoder-MLP	93	93	93	93

TABLE 6: Comparative performance of the Random Forest, XGBoost, and Autoencoder–MLP classifiers

MLP, Multilayer Perceptron; XGBoost, Extreme Gradient Boosting

The comparative analysis bar graphs shown in Figure 9 reveal that all three models achieved accuracy exceeding 90%, indicating their effectiveness for student performance prediction.

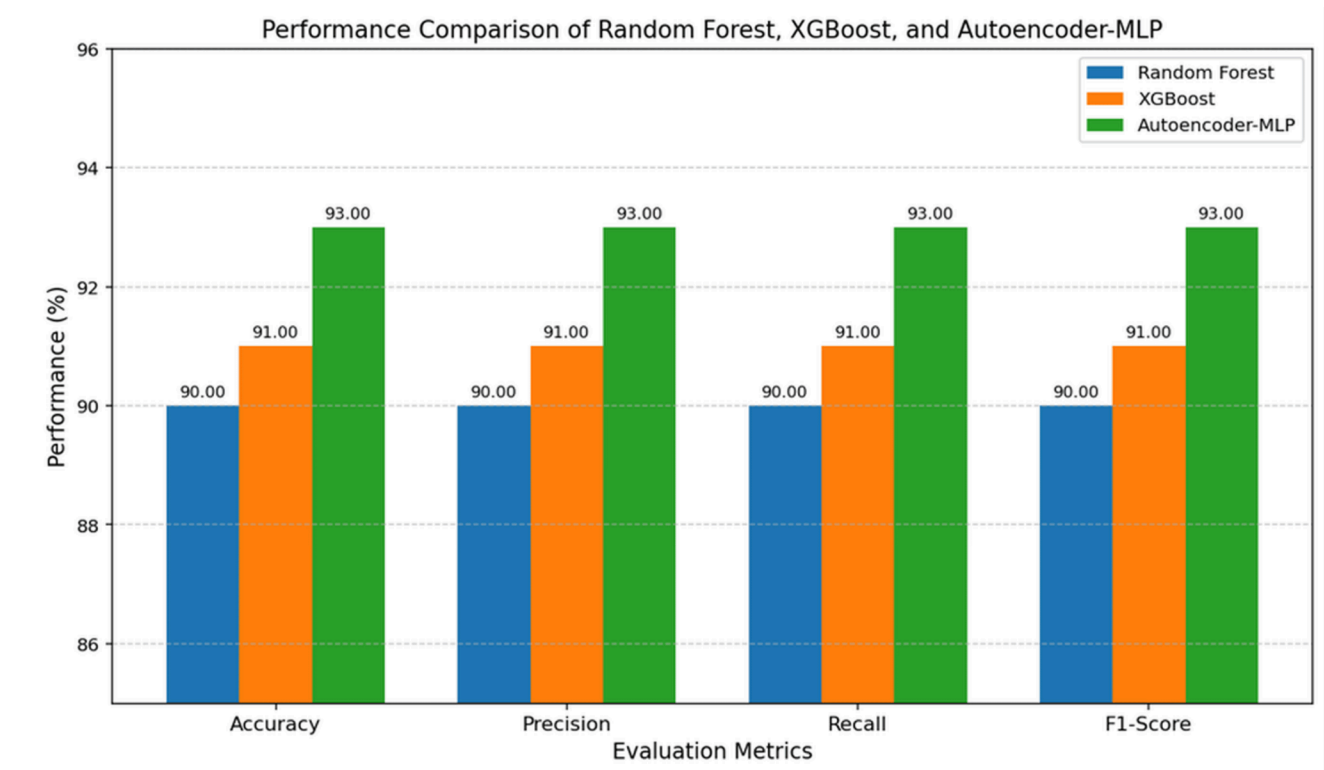


FIGURE 9: Comparative bar graphs of classification performance for Random Forest, XGBoost, and Autoencoder-MLP

MLP, Multilayer Perceptron; XGBoost, Extreme Gradient Boosting

How to cite this article:

Conclusions

This study successfully offers a comprehensive, feature-optimized hybrid deep learning framework specifically designed to maximize predictive performance on academic datasets. The core contributions and findings of this study are effective feature optimization, superiority of the proposed framework, and the value of latent representations. The incorporation of a GA as a front-end selector successfully reduced the raw feature space, selecting a highly discriminative, low-dimensional best feature subset while eliminating non-informative noise. Experimental model performance was assessed on a held-out test set corresponding to 25% of the dataset, while 75% of the dataset was utilized for model development, and the results proved that the proposed HGA-AMLN framework consistently outperformed traditional tree-based baselines (RF and XGBoost). The proposed framework's success highlights that utilizing an unsupervised Autoencoder to map optimized features into a continuous latent space provides a significantly richer signal for dense MLPs than raw or basic scaled inputs. To summarize, this framework establishes a robust, highly accurate blueprint for deploying deep learning models on complex tabular datasets where feature interaction boundaries are highly non-linear. This research bridges the gap between evolutionary computing and tabular deep learning, offering a highly scalable and potent methodology for complex predictive analytics.

Even if the proposed framework achieved excellent validation metrics, there are still several promising avenues for extending this research. We plan to generalize the framework and apply it across independent educational institutions and geographically diverse student populations. We will also focus on validating this hybrid pipeline across diverse public tabular benchmarks (e.g., UCI Machine Learning Repository datasets) to evaluate its out-of-domain stability and generalizability. We aim to apply alternative metaheuristics and benchmark the GA against alternative swarm intelligence and metaheuristic optimization methods, such as Particle Swarm Optimization and the Grey Wolf Optimizer, to compare convergence speeds and feature subset efficiencies. Given the evolving landscape of deep learning for tabular data, future iterations of this framework will explore replacing the MLP classification head with contemporary transformer-based architectures such as TabNet, FT-Transformer, or SAINT to assess whether attention mechanisms can yield further performance breakthroughs. To transition the framework from a high-performing "black box" to a transparent system, we plan to integrate explainability tools, such as SHapley Additive exPlanations and Local Interpretable Model-Agnostic Explanations. This will allow for the post hoc interpretation of both the Autoencoder's latent features and the final MLP decision paths.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Minakshi Sharma, Vibhakar Mansotra

Acquisition, analysis, or interpretation of data: Minakshi Sharma

Drafting of the manuscript: Minakshi Sharma

Critical review of the manuscript for important intellectual content: Vibhakar Mansotra

Supervision: Vibhakar Mansotra

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

Conflicts of interest: In compliance with [the ICMJE uniform disclosure form](#), all authors declare the following:

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

- **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work.
- **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work.
- **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The data (and/or code) supporting this study are openly available in Portuguese Student Performance Dataset from the UCI Machine Learning Repository at <https://archive.ics.uci.edu/ml/datasets/student+performance>, reference 320.

Acknowledgements

The authors would like to thank the UCI Machine Learning Repository and Paulo Cortez for making the Student Performance dataset publicly available. The authors also express their gratitude to the editors and anonymous reviewers for their valuable comments and suggestions, which helped improve the quality of this manuscript. Author Contributions: Minakshi Sharma: Conceptualization, Methodology, Software, Data Curation, Formal Analysis, Investigation, Validation, Visualization, Writing – Original Draft. Vibhakar Mansotra: Conceptualization, Methodology, Supervision, Validation, Writing – Review & Editing, Project Administration.

References

1. Simon PD, Jiang J, Fryer LK, King RB, Frondozo CE: [An assessment of learning management system use in higher education: perspectives from a comprehensive sample of teachers and students](#). Technology, Knowledge and Learning. 2025, 30:741-767. [10.1007/s10758-024-09734-5](#)
2. Wang H, Ge X, Jiang J: [GWO-MLP based achievement prediction model for English learning with different learning motivations](#). Sage Open. 2026, 16:1-21. [10.1177/21582440251405323](#)
3. Kocyigit E, Korkmaz M, Sahingoz OK, Diri B: [Enhanced feature selection using genetic algorithm for machine-learning-based phishing URL detection](#). Applied Sciences. 2024, 14:6081. [10.3390/app14146081](#)
4. Sharma M, Mansotra V: [Implementing a hybrid machine learning technique that merges SVM with KNN to anticipate students' academic achievements](#). International Journal of Information Technology. 2025, [10.1007/s41870-025-02866-6](#)
5. Sharma M, Mansotra V: [Exploring machine learning techniques in school education: a comprehensive analysis of student academic performance](#). Modern Practices and Trends in Expert Applications and Security. Rathore VS, Murala S, Tuba E, Singh MM (ed): Springer, Singapore; 2025. 1379:47-56. [10.1007/978-981-96-5784-1_5](#)
6. Ramaraj M, Dhendapani S, Chembath J, Srividhya S, Thangarasu N, Ilango B: [Educational data mining approach for predicting student performance and behavior using deep learning techniques](#). IAES International Journal of Artificial Intelligence (IJ-AI). 2025, 14:4113-4122. [10.11591/ijai.v14.i5.pp4113-4122](#)
7. Mahesh Kumar NB, Chithrakumar T, Thangarasan T, Dhanasekar J, Logamurthy P: [AI-powered early detection and prevention system for student dropout risk](#). International Journal of Computational and Experimental Science and Engineering. 2025, 11:78-86. [10.22399/ijcesen.839](#)
8. Abu Zohair LM: [Prediction of student's performance by modelling small dataset size](#). International Journal of Educational Technology in Higher Education. 2019, 16:27. [10.1186/s41239-019-0160-3](#)
9. Berhanu F, Abera A: [Students' performance prediction based on their academic record](#). International Journal of Computer Applications. 2015, 131:27-35. [10.5120/ijca2015907348](#)
10. Oyediji AO, Salami AM, Folorunsho O, Abolade OR: [Analysis and prediction of student academic performance using machine learning](#). JITCE (Journal of Information Technology and Computer Engineering). 2020, 4:10-15.

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>

11. Ismail L, Materwala H, Hennebelle A: [Comparative analysis of machine learning models for students' performance prediction](#). Advances in Digital Science. Antipova T (ed): Springer, Cham; 2021. 1352:149-160. [10.1007/978-3-030-71782-7_14](#)
12. Ojajuni O, Ayeni F, Akodu O, et al.: [Predicting student academic performance using machine learning](#). Computational Science and Its Applications - ICCSA 2021. Gervasi O, Murgante B, Misra S, et al. (ed): Springer, Cham; 2021. 12957:481-491. [10.1007/978-3-030-87013-3_36](#)
13. Derveniz C, Kyriatizis V, Stoufis S, Fitsilis P: [Predicting students' performance using machine learning algorithms](#). ICACS '22: Proceedings of the 6th International Conference on Algorithms, Computing and Systems. 2022, 1-7. [10.1145/3564982.3564990](#)
14. Roy K, Farid DM: [An adaptive feature selection algorithm for student performance prediction](#). IEEE Access. 2024, 12:75577-75598. [10.1109/ACCESS.2024.3406252](#)
15. Kalita E, Alfarwan AM, El Aouifi H, Kukkar A, Hussain S, Ali T, Gaftandzhieva S: [Predicting student academic performance using Bi-LSTM: a deep learning framework with SHAP-based interpretability and statistical validation](#). Frontiers in Education. 2025, 10:1581247. [10.3389/feduc.2025.1581247](#)
16. Sharma M, Mansotra V: [A dual attention-driven neural ensemble framework for student outcome prediction](#). 2026 International Conference on Emerging Systems and Intelligent Computing (ESIC), Bhubaneswar, India. 2026, 182-187. [10.1109/ESIC68176.2026.11495867](#)
17. Yağcı M: [Educational data mining: prediction of students' academic performance using machine learning algorithms](#). Smart Learning Environments. 2022, 9:11. [10.1186/s40561-022-00192-z](#)
18. Waheed H, Hassan SU, Aljohani NR, Hardman J, Alelyani S, Nawaz R: [Predicting academic performance of students from VLE big data using deep learning models](#). Computers in Human Behavior. 2020, 104:106189. [10.1016/j.chb.2019.106189](#)
19. Gufroni AI, Purwanto P, Farikhin F: [Academic performance prediction using supervised learning algorithms in university admission](#). JOIV: International Journal on Informatics Visualization. 2025, 9:184-194. [10.62527/joiv.9.1.2974](#)
20. Ahmed E: [Student performance prediction using machine learning algorithms](#). Applied Computational Intelligence and Soft Computing. 2024, 2024:1-15. [10.1155/2024/4067721](#)
21. Saini BK: [Optimizing student academic performance prediction using heterogeneous ensemble learning](#). European Journal of Artificial Intelligence and Machine Learning. 2025, 4:1-6. [10.24018/ejai.2025.4.4.77](#)
22. Almazroei EE: [Ensemble machine learning framework with SHAP and LIME for accurate early prediction of student success in online learning environments](#). Scientific Reports. 2026, 16:19506. [10.1038/s41598-026-49894-1](#)
23. Korkmaz Ö, Aydın MN: [Detection of early school drop out in vocational and technical high schools in Turkey](#). Sage Open. 2025, 15:1-17. [10.1177/21582440251370443](#)
24. Taqatqeh S, Agoyi M: [Early warning and recommendation system for undergraduate students at higher education institutions based on machine learning methods](#). PeerJ Computer Science. 2026, 12:e3792. [10.7717/peerj-cs.3792](#)
25. Aldamani R, Aoudia M, Abuhani DA, Arif O: [Recent advances in genetic algorithm-based feature selection: A comprehensive survey](#). Swarm and Evolutionary Computation. 2026, 102:102318. [10.1016/j.swevo.2026.102318](#)

How to cite this article:

Sharma M, Mansotra V (September 09, 2026) A Genetic Algorithm-Based Feature Selection Autoencoder-Multilayer Perceptron Framework for Student Academic Performance Prediction . Cureus J Comput Sci 3 : es44389-026-00269-7. DOI <https://doi.org/10.7759/s44389-026-00269-7>