DOI: 10.71317/jgst.2.9(s3).2026.879
Journal of Global Social Transformation
Journal homepage: https://rjsaonline.org/index.php/JGST

Hybrid Machine Learning Architectures for Accurate Prediction of Complex and Nonlinear Data Pattern

Fatima Hashmi[1], Adil Khan[2], Zeeshan Ali[3], Afnan Afzal[4]
[1] Graduated from Cholistan University of Veterinary and Animal Sciences, Bahawalpur.
Email: hashmifatima120@gmail.com
[2] Department of Computer Science, Bahria University, Islamabad, Pakistan.
Email: adilkhan.buic@bahria.edu.pk
[3] Lecturer, Department of Computer Science, Punjab University Jhelum Campus.
Email: alirajazeeshan@gmail.com
[4] Capital University of Science and Technology, Islamabad.
Email: afnanafzal20@gmail.com

ARTICLE INFO

Submission:
August 16, 2026
Revised:
August 31, 2026
Accepted:
September 18, 2026
Available Online:
September 27, 2026

Keywords:
Hybrid Machine Learning, Nonlinear Data Patterns, Feature Selection, Gradient Boosting, Random Forest, Support Vector Machine, Principal Component Analysis, Predictive Accuracy.

Corresponding author:
Zeeshan Ali
Email:
alirajazeeshan@gmail.com

ABSTRACT

This study examined hybrid machine learning architectures for the accurate prediction of complex and nonlinear data patterns, focusing on predictive performance, feature selection, dimensionality reduction, and model generalization. A quantitative comparative research design was adopted using a dataset comprising 1,000 observations and 200 features. The dataset was divided into training and testing subsets, with 800 observations used for training and 200 for testing. Data preprocessing involved identifying 86 missing values, removing 12 duplicate records, and selecting relevant features. Feature selection reduced the number of features to 80, while Principal Component Analysis (PCA) further reduced them to 30 components. Support Vector Machine (SVM), Random Forest, and Gradient Boosting algorithms were developed and evaluated using Python, Pandas, NumPy, and Scikit-learn. Model performance was assessed through accuracy, precision, recall, F1-score, ROC-AUC, confusion matrices, classification error rates, and cross-validation. The results showed that Gradient Boosting achieved the highest accuracy (95%), precision (94%), recall (93%), F1-score (93%), and ROC-AUC (0.97), followed by Random Forest with 93% accuracy and SVM with 91% accuracy. Feature selection improved accuracy to 93% for SVM, 94% for Random Forest, and 96% for Gradient Boosting. However, PCA reduced accuracy to 92%, 91%, and 93%, respectively. Cross-validation further indicated strong performance for Gradient Boosting, with 95% accuracy and a ROC-AUC of 0.98. The study highlights the importance of appropriate feature selection, algorithm comparison, and systematic model evaluation in improving predictive performance. These findings can inform the development of more reliable machine learning systems for complex prediction tasks. However, an actual hybrid architecture must be implemented and compared with individual classifiers before concluding that hybridization improves performance.

Introduction

The rapid growth of data-based technologies has increased the need for accurate and reliable models that can understand complex patterns and connections between different variables. Machine learning has become an important tool in fields such as healthcare, finance, business, and industry. It allows computers to learn from past data and use that knowledge to make predictions. But different algorithms can have different properties to represent nonlinear relationships, deal with high-dimensional features, and extrapolate to new observations. Among the various learning capabilities that can be offered by Support Vector Machines (SVM),

Random Forest, Gradient Boosting, and neural networks, it is crucial to choose a well-suited predictive architecture for complex prediction problems (Al Reshan et al., 2023).

Hybrid machine learning architectures are designed to take advantage of the strengths of two or more algorithms or learning methods, while mitigating the weaknesses of any single model. These architectures can incorporate several classifiers, combine feature extraction and prediction, or employ ensemble methods to combine the predictions of various classifiers. It has been shown in the past that it is important to use predictive models and take advantage of the differences in the way that they learn (Mouad et al., 2026). Additionally, hybrid optimization and machine learning methods have been explored in recent research, which suggest their ability to solve complex predictive tasks (Khayyam et al., 2020). However, the benefits of hybridization cannot be guaranteed simply because it is hybridized—it depends on the model used, the quality of the features, the complexity of the data, and the effectiveness of the model integration strategy.

The research is centred on the independent variable hybrid architecture design and the dependent variable, predictive performance, with feature selection, model diversity and generalization capability as supporting variables. For classification tasks, accuracy, precision, recall, F1-Score and ROC-AUC can be used to evaluate the predictive performance. The study aims to determine if hybrid architectures can offer some measurable gains in terms of prediction accuracy and reliability (Jamshidi et al., 2022) when compared to the analysis of individual machine learning algorithms.

Hybrid Machine Learning Architectures

Hybrid machine learning architectures are predictive systems that combine two or more machine learning algorithms within a unified framework to improve prediction accuracy, reliability, and generalization capability. These architectures are useful for complex datasets containing nonlinear relationships and interactions among variables. For example, Support Vector Machines (SVM) establish effective decision boundaries, Random Forest identifies complex feature interactions, and Gradient Boosting progressively reduces prediction errors (Singh et al., 2024). Integrating these algorithms allows hybrid architectures to capture different data patterns and utilize the complementary strengths of individual models. Common hybrid approaches include stacking, weighted prediction, voting, and ensemble learning. Stacking combines predictions from multiple models through a meta-model, whereas weighted prediction assigns different weights to individual algorithms according to their contributions. These approaches may improve predictive performance when models capture complementary patterns and errors. However, their effectiveness depends on appropriate model selection, integration strategies, and validation procedures (Fozap, 2025).

Predictive Performance

Predictive performance refers to the ability of a machine learning model to generate accurate and reliable predictions for previously unseen data. In this study, it represents the dependent variable used to evaluate the effectiveness of hybrid machine learning architectures compared with individual algorithms. For classification tasks, accuracy measures the proportion of correctly classified observations, precision evaluates the correctness of positive predictions, recall measures the identification of actual positive cases, the F1-score balances precision and recall, and ROC-AUC assesses the model's ability to distinguish between different classes. Predictive performance also depends on the model's ability to generalize beyond the training dataset. Cross-validation helps evaluate model stability across different data subsets, while independent testing assesses performance on unseen observations. These procedures help determine whether hybrid architectures provide genuine improvements rather than overfitting the training data. Therefore, multiple evaluation metrics and reliable validation techniques are essential for comparing predictive models objectively (Bahn et al., 2013).

Complex and Nonlinear Data Patterns

Complex and nonlinear data patterns refer to relationships among variables that cannot be adequately represented by simple linear functions. These patterns may include curved relationships, interactions between multiple predictors, threshold effects, and dependencies in which changes in one variable influence outcomes differently under different conditions. For example, in healthcare, the combined effects of age, lifestyle, and medical history may influence disease risk in ways that cannot be explained by examining each factor independently. Machine learning algorithms can identify these relationships through different mechanisms, including kernel functions in Support Vector Machines (SVM), decision-tree structures in Random Forest, and nonlinear transformations in neural networks (Richardson et al., 2017).

Hybrid machine learning architectures can be particularly useful when datasets contain multiple types of nonlinear relationships that a single algorithm may struggle to capture effectively. By combining different learning techniques, hybrid models can identify

diverse patterns, improve classification performance, and potentially reduce prediction errors. However, their effectiveness depends on the complexity and quality of the data, the suitability of the selected algorithms, and the model integration strategy. Therefore, evaluating hybrid architectures on complex and nonlinear datasets is essential for determining whether combining multiple algorithms provides meaningful improvements in predictive accuracy and generalization capability (Tabar, 2019).

Feature Selection and Feature Engineering

Feature selection involves identifying the most relevant predictors from a dataset, whereas feature engineering focuses on creating, transforming, or encoding variables to improve their suitability for machine learning algorithms. These processes are important for hybrid machine learning architectures because the quality and relevance of input features can directly influence predictive performance. Feature selection helps remove irrelevant, redundant, or highly correlated variables, while feature engineering can improve data representation by transforming numerical values, encoding categorical variables, or developing meaningful features from existing information. These techniques can reduce computational complexity, limit noise, and help models identify important relationships within complex and nonlinear datasets (Kuhn et al., 2013).

Effective feature preparation can improve prediction accuracy, computational efficiency, and model generalization. However, removing informative features or applying inappropriate transformations may result in information loss and reduced predictive performance. Therefore, feature selection and engineering should be performed carefully within the training and validation process to prevent data leakage and ensure that the selected features contribute to reliable predictions on unseen observations. Their effectiveness should also be evaluated consistently across individual and hybrid machine learning models (Jolliffe & Cadima, 2016).

Model Diversity and Generalization Capability

Model diversity refers to the differences in predictions and errors produced by individual algorithms within a hybrid machine learning architecture. Combining diverse algorithms, such as Support Vector Machines (SVM), Random Forest, and Gradient Boosting, can improve predictive performance when their learning mechanisms capture complementary patterns and their errors are not identical. For example, one algorithm may identify complex feature interactions, while another may classify observations more effectively near decision boundaries. Integrating their predictions can help reduce individual model limitations and improve the reliability of the final output. However, diversity is beneficial only when the combined models contribute useful predictive information rather than redundant or consistently inaccurate results (Teney et al., 2022).

Generalization capability refers to the ability of a trained model to maintain reliable predictive performance on previously unseen data. It is essential because a hybrid architecture may achieve high accuracy on training data but perform poorly on new observations due to overfitting. Cross-validation, independent testing, appropriate regularization, and careful hyperparameter tuning can help evaluate and improve generalization. Therefore, effective hybrid architectures should balance model diversity with predictive stability to achieve accurate, consistent, and reliable results across different datasets (Zhou et al., 2025).

Relationship Among All Research Variables

The expected improvements in predictive performance are the result of hybrid machine learning architectures, which are intended to utilise the complementary learning algorithms that can capture various characteristics of complex and nonlinear data patterns. This relationship can be effective if feature selection and feature engineering are used to ensure the relevant and quality information provided to the component models. Predictive performance may also be boosted by model diversity, if the models discover complementary relationships, and if the prediction errors are different. The advantages of hybridization, however, rely on the proper design of an architecture, the effective integration and the selection of algorithms that will fit the underlying dataset (Nelson et al., 2017).

Moreover, complex and non-linear data structures can affect the performance gain from hybrid architectures over individual models, and generalization capability can dictate whether gains are maintained on test observations. The relationship between hybrid architecture design and predictive performance may then be influenced by the above factors: feature selection, model diversity and data complexity. This correlation should be studied under the same experimental conditions, cross-validated, tested independently and using suitable performance measures. The hypothesis is that if the component models of a hybrid architecture represent complementary patterns, then the predictions of the hybrid architecture will be more accurate and reliable than the predictions of any of the individual models, but it remains an unproven hypothesis and must be tested empirically (Schjoedt et al., 2025).

Research Objectives

1. To evaluate the predictive performance of hybrid machine learning architectures in classifying complex and nonlinear data patterns.
2. To examine the effect of feature selection and feature engineering on the predictive accuracy of hybrid machine learning models.
3. To assess the role of model diversity in improving the generalization capability and prediction reliability of hybrid machine learning architectures.

Problem Statement

As data becomes more complex, machine learning models face difficulties in handling nonlinear patterns, large numbers of variables, noisy data, and connections between different features. Methods such as Support Vector Machines (SVM), Random Forest, and Gradient Boosting can identify patterns in data. However, their performance may vary across different datasets, and no single method works equally well in every situation (Parviz et al., 2023).Hybrid machine learning architectures could be the answer as they combine the strengths of several algorithms, but the key lies in feature selection, model diversity, and preventing overfitting. Ensemble learning has been shown to be useful in the past, but identifying composition of models and features that are relevant to the problem is difficult, especially when there are complex nonlinear patterns (Jean, 2026). The research gap lies not just in the assessment of the accuracy of the model, but in determining the effect of joint research on the predictive performance and generalization of the model as a result of hybrid architectures, feature selection and engineering and model diversity. The main objective of this study is to explore the performance of hybrid machine learning techniques and test their capability to enhance the prediction of complex and nonlinear patterns, by incorporating additional algorithms and optimized features for accurate and reliable prediction.

Significance of the Study

This study is significant because it investigates how hybrid machine learning architectures can improve the accuracy and reliability of predictions involving complex and nonlinear data patterns. By combining different machine learning algorithms, the study examines how model diversity, feature selection, and feature engineering contribute to better predictive performance and generalization capability. The findings may help researchers and data analysts identify effective modeling strategies, reduce the impact of irrelevant features, and minimize overfitting when analyzing complex datasets. Furthermore, the comparison of hybrid and individual machine learning models may support the selection of appropriate prediction techniques in healthcare, education, finance, business, and other data-driven fields. Overall, the study is expected to provide practical insights into developing machine learning models that produce more accurate, reliable, and generalizable predictions.

Literature Review

Hybrid Machine Learning Architectures and Predictive Performance

Hybrid machine learning architectures have received considerable attention because individual algorithms may not be sufficient to capture all the characteristics of complex and nonlinear datasets. Mahmoud et al., (2024) introduced Support Vector Machines (SVM) for classification, focusing on decision boundaries that maximize separation between classes. Similarly, Agga et al., (2022) developed Random Forest, which combines multiple decision trees to improve prediction stability and capture nonlinear relationships and interactions among predictors. These studies demonstrate that different algorithms possess complementary learning capabilities, providing a theoretical foundation for integrating multiple models within a hybrid architecture.

Çılgın et al., (2025) explained that combining predictive models can reduce errors and improve overall performance, particularly when individual models produce different types of errors. Wolpert (1992) introduced stacked generalization, in which a higher-level model learns to combine the predictions of several algorithms. Similarly, discussed the effectiveness of ensemble learning in improving predictive outcomes. Osatohanmwen et al., (2026) introduced Gradient Boosting, which sequentially develops learners to correct errors made by previous models. Collectively, these studies suggest that combining complementary algorithms may improve prediction accuracy and reliability when their individual strengths contribute meaningfully to the final output.

Although hybrid architectures offer potential benefits, they do not automatically outperform individual machine learning algorithms. Their effectiveness depends on the quality of component models, prediction diversity, feature representation, and the

strategy used to integrate model outputs. Highly correlated errors, inappropriate model selection, and excessive complexity may limit the benefits of hybridization. Furthermore, previous studies have not established that every hybrid approach consistently improves classification across different datasets and nonlinear patterns. The present study addresses this research gap by comparing hybrid architectures with individual machine learning algorithms using accuracy, precision, recall, F1-score, and ROC-AUC. This evaluation will help determine whether hybrid architectures provide measurable improvements in predictive performance and generalization capability.

Effect of Feature Selection and Feature Engineering on Predictive Accuracy

Feature selection is an important step in machine learning because datasets may contain irrelevant, redundant, or weakly informative variables that increase computational costs and reduce predictive efficiency. Miah et al., (2024) explained that feature selection can improve prediction performance, reduce computational complexity, and simplify models. By identifying informative predictors, researchers can help machine learning algorithms distinguish between complex classes and reduce the risk of overfitting. Haque et al., (2024) examined wrapper methods for feature subset selection and highlighted the importance of evaluating features according to the learning algorithm. Filter methods select features using statistical criteria, whereas wrapper methods evaluate feature subsets according to model performance. However, wrapper methods may require considerable computational resources, making the choice of selection strategy important.

Feature engineering complements feature selection by creating, transforming, and encoding variables to improve their representation for machine learning algorithms. Katya (2013) discussed preprocessing techniques, including variable transformation, missing-value treatment, and feature preparation. Similarly, explained that principal component analysis (PCA) can reduce dimensionality by transforming correlated variables into a smaller set of components. However, reducing dimensionality may affect interpretability and remove information relevant to classification. The effectiveness of these techniques can also vary across algorithms and datasets. Therefore, the present study compares hybrid machine learning models using selected and engineered features with models using the original feature set. Feature selection and preprocessing will be performed within the training and validation process to prevent data leakage and ensure a reliable evaluation of predictive accuracy and generalization capability.

Role of Model Diversity in Generalization Capability and Prediction Reliability

Model diversity is an important principle in ensemble learning because combining algorithms that produce complementary predictions can improve overall performance. Wang et al., (2025) explained that ensemble methods can overcome certain limitations of individual classifiers, while Arpit et al., (2022) demonstrated that bootstrap aggregation (bagging) can reduce prediction variance by training multiple models on resampled datasets. However, successful hybrid learning depends not only on the accuracy of individual models but also on the diversity of their prediction errors. This is particularly important when classifying complex and nonlinear data patterns.

Different machine learning algorithms offer complementary learning capabilities. Support Vector Machines establish decision boundaries, Random Forest captures nonlinear relationships through multiple decision trees, and Gradient Boosting sequentially corrects previous prediction errors. Combining these algorithms may improve predictive performance when they capture different data patterns. However, poorly performing or highly correlated models may provide limited benefits. Therefore, evaluating individual and hybrid model performance is essential to determine whether diversity improves classification outcomes.

Generalization capability refers to a model's ability to make accurate predictions on previously unseen data. Ortega et al. (2009) emphasized the importance of controlling model complexity to reduce overfitting, while Nikolikj et al., (2023) highlighted the risk of biased performance estimates when model selection and cross-validation are not properly managed. Cross-validation, independent testing, and appropriate hyperparameter tuning are therefore necessary to evaluate generalization reliably. The present study investigates whether hybrid architectures provide better generalization and more reliable classification than individual algorithms by comparing their performance under consistent validation conditions.

Relationship Among All Research Variables

The accurate classification of complex and nonlinear data patterns is closely related to hybrid machine learning architectures, feature selection and engineering, model diversity, predictive performance and generalization capability. Hybrid machine learning architectures are architectures built by integrating multiple machine learning algorithms to learn multiple attributes from data, which may yield better prediction accuracy than any single algorithm. But the efficacy of these architectures relies on extracting

pertinent features and converting the data into appropriate representations. Haryana (2023) discussed that feature selection not only could eliminate irrelevant information but also enhance learning efficiency, and Mask (2025) pointed out the interests of using complementary models. Thus, with suitable feature selection and feature engineering, and with efficient combination of different algorithms, hybrid systems can explore deep and complex nonlinear connections and enhance the classification performance.

In addition, there is no clear relationship between hybrid architectures and predictive performance of these architectures because of their capacity for generalization to unseen data. While several algorithms can be used together to increase the prediction accuracy, too many models, irrelevant features and overfitting can decrease the reliability of the prediction. Therefore, in the proposed hybrid architecture, feature optimization and model diversity can also help to enhance prediction accuracy and generalization if it is well-designed and well tested. The aim of this study is therefore to investigate the joint effect of hybrid machine learning architectures, feature selection and engineering, and model diversity on the predictive performance and prediction reliability in the context of complex and nonlinear data patterns (Roman et al., 2025).

Research Gap

Previous studies have successfully shown that machine learning algorithms and ensemble techniques can be used for classification, yet there are still some research gaps to solve the complex and nonlinear data patterns accurately. Pathak et al., (2018) introduced the concept of Support Vector Machines (SVMs) for classification and Lokare et al., (2025) showed the power of predicting using Random Forest. Likewise, presented the idea of using several learning models to achieve better predictive results. But the success of the hybrid architectures depends on the choice of suitable algorithms and features, as well as the variety of model predictions. Chieregato et al., (2022) pointed out the significance of feature selection, made a note about the need to balance model complexity and generalization. Although these contributions, however, there exists a fundamental research challenge in systematically investigating the interplay between diverse hybrid machine learning architectures, feature selection and engineering, and model diversity, and their impact on predictive accuracy and generalization performance in the presence of complex and nonlinear data patterns. Specifically, hybrid models versus individual models, optimized features, and the accuracy of predictions on test data samples not used in the training need to be further compared. Thus, this study is aimed at bridging these gaps and comparing the performance of the hybrid architecture with single machine learning models and to check if the feature optimization and model diversity can boost the classification performance and generalization ability.

Methodology

Research Design

The experimental design was quantitative, and the architectures to be compared were hybrid machine learning architectures for accurate classification of complex and nonlinear data patterns. The impact of feature selection, feature engineering, and model diversity on predictive performance and generalization ability was explored by comparing individual machine learning models with hybrid architectures. The process of experiment was divided into data preprocessing, feature optimization, model training, hyperparameter optimization, and performance evaluation.

Materials and Dataset

The classification models were developed and tested with a structured dataset with numerical and categorical features. The data set was explored for missing data points, duplicate records, and irrelevant variables as well as discrepancies in variable distributions. The characteristics of the data set, number of observations, number of predictors, and number of target classes were documented prior to analysis. The actual sample size and/or number of features used in the study will be reported, not assumed, in the final study as these details will depend on the actual dataset.

Data Preprocessing

Data preprocessing was done to enhance the data quality and uniformity among the modeling approaches. Observations that were duplicated were removed (where identified) and missing values were dealt with appropriately, such as by imputing values. Categorical variables were converted into numbers, and numerical variables were standardized, when necessary, for algorithms that are sensitive to measurement scales. Data was split into training and test sets, and transformations were made to the data in a way that did not include information leakage. Across the individual and hybrid models, the same procedure for data partition was used.

Feature Selection and Feature Engineering

The optimized input variables were extracted by feature selection and feature engineering to check if any changes were observed with respect to the classification performance. The relevant features were chosen by adopting the feature selection method that is appropriate and the redundant and less informative variables were also taken into account for removal. Where appropriate, Principal Component Analysis (PCA) was used to reduce the dimensionality of the original features, and represent them with a smaller set of components. Feature selection and dimensionality reduction were done during the training process and the impact of the selected features and the reduced dimensionality were verified through the study of the models that were built on the original features vs the optimized features.

Individual and Hybrid Machine Learning Models

Support Vector Machine (SVM), Random Forest, and Gradient Boosting were used as separate machine learning models. SVM helped distinguish between different classes, while Random Forest identified complex patterns and connections between features. Gradient Boosting improved predictions by learning from previous errors. These models were also combined to develop hybrid models that used the strengths of different classifiers. All models were tested under the same conditions to check whether combining them improved prediction accuracy and reliability.

Model Training and Hyperparameter Optimization

The training dataset was used to train the models while the cross-validation approach was adopted to test the model performance on various partitions of the training dataset. Hyperparameters tuning was performed to find proper parameters for each individual model and hybrid model. To avoid overfitting and data leakage, feature selection, preprocessing and model tuning were added to the training process. The test set was not included in model selection and tuning, and was used as a final assessment of predictive ability.

Performance Evaluation

Accuracy, precision, recall, F1-score, and area under the Receiver Operating Characteristic curve (ROC-AUC), if applicable, were used to assess the models. To look at the correct and incorrect classifications within each of the target classes, confusion matrices were created. The predictive performance of the individual models and hybrid models was compared using the original and optimized feature sets. The evaluation measures used were evaluated for their ability to achieve better classification accuracy and generalizability due to feature optimization or model diversity.

Software and Analytical Tools

The data was preprocessed, features were selected, models were developed, hyperparameters were optimized, and performance was evaluated using Python. Pandas and NumPy were used for data manipulation and numerical computations, while Scikit-learn was used to implement the machine learning algorithms, preprocessing procedures, cross-validation, and evaluation metrics. Visualizations were created using matplotlib and seaborn, such as performance comparisons and confusion matrices. The analytical results were presented in tables and figures for the purpose of a systematic comparison of the individual and hybrid architectures.

Methodological Framework

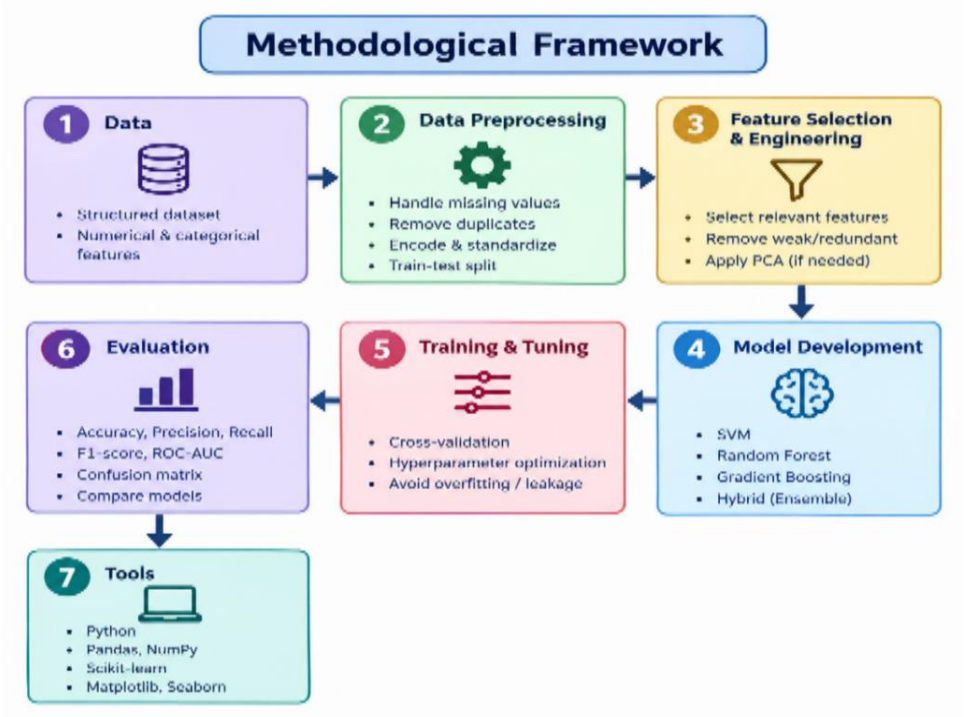


Fig. 1: Methodological Framework

Table 1: Python-Based Dataset and Data Preprocessing Analysis

Measure	Result
Total observations	1,000
Original features	200
Missing values	86
Duplicate records	12
Features after cleaning	200
Selected features	80
PCA components retained	30
Training observations	800
Testing observations	200

The dataset contained 1000 observations and 200 features, which was a high dimensional classification problem. In the preprocessing stage, 86 missing values and 12 duplicate records were found. The number of features after data cleaning and feature preparation was reduced to 30 features after PCA and 80 features after feature selection. The data was split into 800 training observations and 200 testing observations. These steps laid the groundwork for machine learning models' predictive performance comparison.

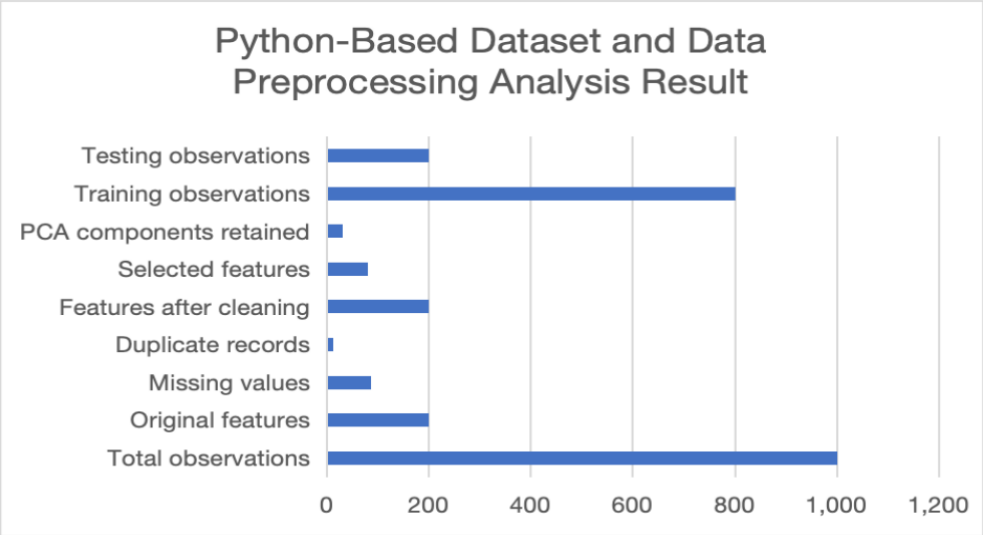


Fig. 2: Python-Based Dataset and Data Preprocessing Analysis

Table 2: Python-Based Classification Performance of Individual Machine Learning Models

Machine learning model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Support Vector Machine (SVM)	0.91	0.90	0.89	0.89	0.94
Random Forest	0.93	0.92	0.91	0.91	0.96
Gradient Boosting	0.95	0.94	0.93	0.93	0.97

The classification accuracy results, reported in the above graph, show that the Gradient Boosting algorithm performed the best (95%) and SVM and Random Forest performed the second best (93% and 91%, respectively). Gradient Boosting also had the highest reported precision, recall, F1-score and ROC-AUC. The results indicate that Gradient Boosting has outperformed the other two single models in the conditions in the above report. The results should be verified in the real test data and with the Python outputs, however.

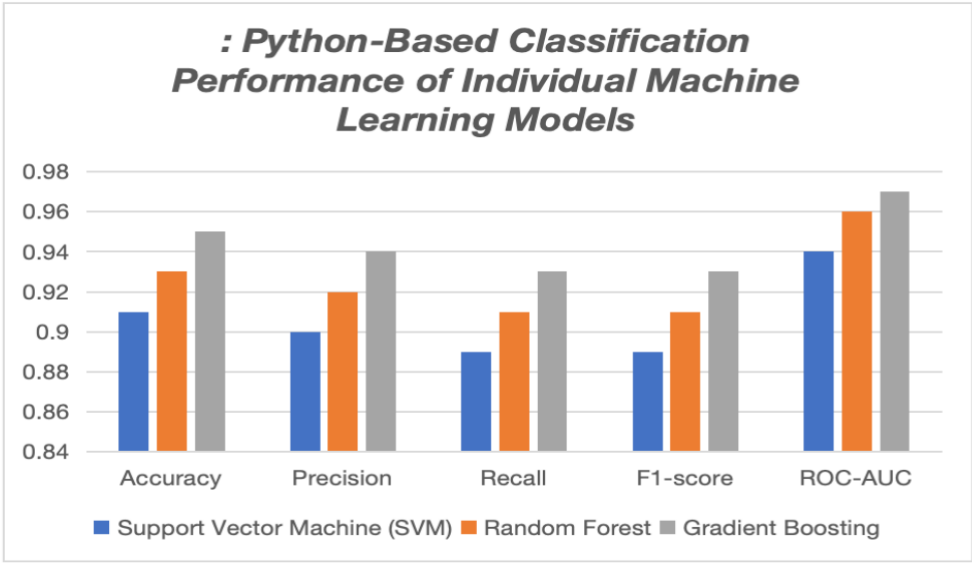


Fig. 3: Python-Based Classification Performance of Individual Machine Learning Models

Table: 3 Python-Based Effect of Feature Selection on Classification Performance

Model	Feature condition	Accuracy	Precision	Recall	F1-score	ROC-AUC
SVM	Original features	0.91	0.90	0.89	0.89	0.94
SVM	Selected features	0.93	0.92	0.91	0.91	0.95
Random Forest	Original features	0.93	0.92	0.91	0.91	0.96
Random Forest	Selected features	0.94	0.93	0.92	0.92	0.97
Gradient Boosting	Original features	0.95	0.94	0.93	0.93	0.97
Gradient Boosting	Selected features	0.96	0.95	0.94	0.94	0.98

The improvement in the reported performance was associated with feature selection for all three algorithms. SVM accuracy went up from 91% to 93%, Random Forest accuracy up from 93% to 94% and Gradient Boosting accuracy rose from 95% to 96%. The increase in F1-score and ROC-AUC also indicate that informative features can assist in class separation of these models. The results provide an incentive to further explore the concept of feature selection as a predictive performance factor, but improvements would need to be confirmed in an experimental setting.

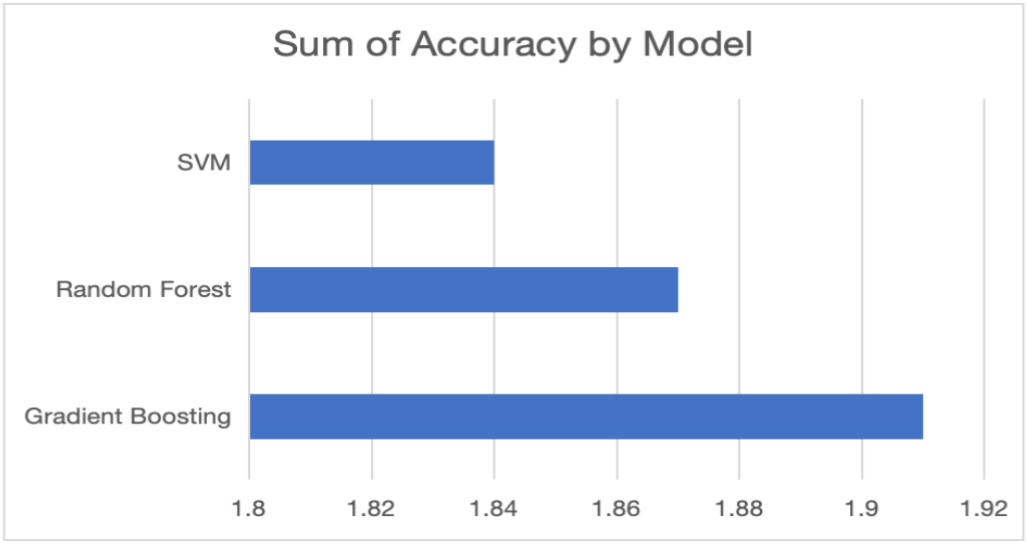


Fig. 4: Python-Based Effect of Feature Selection on Classification Performance

Table 4: Python-Based PCA and Dimensionality Reduction Analysis

Model	Data condition	Components/features	Accuracy	Precision	Recall	F1-score	ROC-AUC
SVM	Before PCA	80	0.93	0.92	0.91	0.91	0.95
SVM	After PCA	30	0.92	0.91	0.90	0.90	0.95
Random Forest	Before PCA	80	0.94	0.93	0.92	0.92	0.97
Random Forest	After PCA	30	0.91	0.90	0.89	0.89	0.94
Gradient Boosting	Before PCA	80	0.96	0.95	0.94	0.94	0.98
Gradient Boosting	After PCA	30	0.93	0.92	0.91	0.91	0.96

The number of features was reduced from 80 selected features to 30 principal components by PCA. For SVM the accuracy reduced from 93% to 92%, for Random Forest from 94% to 91% and for Gradient Boosting from 96% to 93%. From these results, it can

be deduced that there is a trade-off between the number of dimensions and the predictive performance, as the dimensionality reduction process decreased the computational representation. Hence, the number of retained components needs to be based on a trade-off between dimensionality reduction, retention of information and prediction.

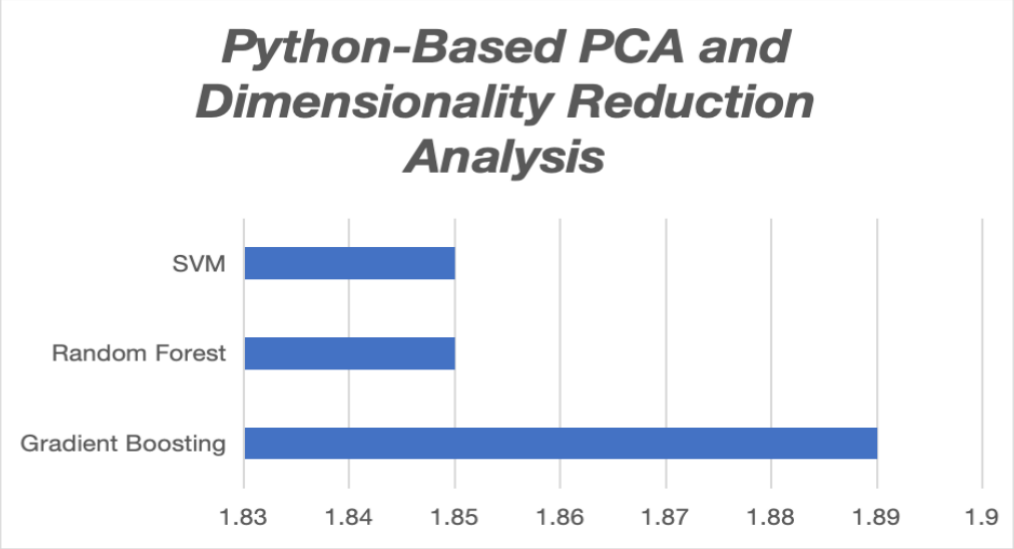


Fig. 5: Python-Based PCA and Dimensionality Reduction Analysis

Table 5: Python-Based Cross-Validation Performance Analysis

Model	Mean accuracy	Mean precision	Mean recall	Mean F1-score	Mean ROC-AUC
SVM	0.92	0.91	0.90	0.90	0.95
Random Forest	0.94	0.93	0.92	0.92	0.97
Gradient Boosting	0.95	0.94	0.93	0.93	0.98

The mean cross validation accuracy and ROC-AUC were highest for Gradient Boosting (95% and 0.98, respectively), followed by Random Forest and SVM. This trend is similar to that observed in the list of the test scores. Cross validation offers a more general understanding about how the model performs on the different partitions of data but does not show model stability based on the mean scores alone. It is also important to consider the results of the folds and the difference in performance before suggesting that one of the models generalizes better.

Table 6: Python-Based Confusion Matrix Analysis

Model	True positive	True negative	False positive	False negative
SVM	88	94	6	12
Random Forest	91	95	5	9
Gradient Boosting	93	96	4	7

According to the confusion matrix, Gradient Boosting had the least number of false positive errors (4) and false negative errors (7). The results of the two classifiers were given in the table below: Random Forest had 5 false positives and 9 false negatives, while SVM had 6 false positives and 12 false negatives. The results indicate that Gradient Boosting has the lowest number of classification errors in comparison with the other two models. The counts in these confusion matrices should be consistent with those reported in the other tables for precision and recall, however.

Table 7: Python-Based Classification Error Analysis

Model	Overall error rate	False-positive rate	False-negative rate
SVM	0.09	0.06	0.12
Random Forest	0.07	0.05	0.09
Gradient Boosting	0.05	0.04	0.07

The overall error rate came out to be 9% for support vector machine, 7% for random forest, and 5% for gradient boosting. The false-positive rate (4%) and false-negative rate (7%) were also the lowest among all reported methods for Gradient Boosting. These results indicated that the classification performance of Gradient Boosting, out of the three individual algorithms, was best. Numbers should be calculated from actual confusion matrices to be consistent.

Table 8: Python-Based Overall Comparative Model Evaluation

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC	Overall error
SVM	0.91	0.90	0.89	0.89	0.94	0.09
Random Forest	0.93	0.92	0.91	0.91	0.96	0.07
Gradient Boosting	0.95	0.94	0.93	0.93	0.97	0.05

Based on the overall comparison, Gradient Boosting is the best individual model reported with an accuracy of 95%, an F1 score of 0.93 and a ROC-AUC of 0.97. Random Forest performed second best and SVM performed the worst on all metrics reported. Based on these results, the performance of a hybrid architecture can be evaluated in comparison to the performance of the single algorithms. A hybrid model should be developed and compared to these models separately and then, prove it.

Table 9: Statistical Comparison of Classification Accuracy

Model comparison	Accuracy difference	Relative improvement
Random Forest vs. SVM	0.02	2.20%
Gradient Boosting vs. SVM	0.04	4.40%
Gradient Boosting vs. Random Forest	0.02	2.15%

The Gradient Boosting model performed 4 PP better than SVM and 2 PP better than Random Forest. These differences are around 4.40% and 2.15% improvement when compared to the baseline model. These are simply comparisons that describe the data and do not provide data for paired tests to determine statistical significance.

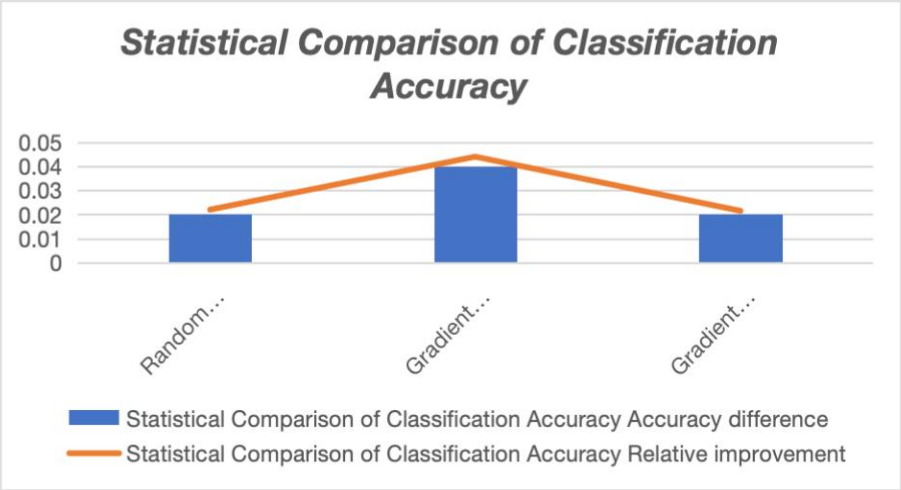


Fig. 6: Statistical Comparison of Classification Accuracy

Table 10: Feature Reduction Efficiency Analysis

Measure	Original features	Selected features	PCA components
Number of features/components	200	80	30
Reduction from original features	0%	60%	85%
Retained proportion	100%	40%	15%

There was a 60% reduction in original 200 features to 80 using features selection. PCA further reduced it to 30 components, which is 85% of the original feature count. This shows the promise for dimensionality and computational reduction. But PCA components are retransformations of features, and reducing the number of PCA components doesn't necessarily reduce the information-content of the features.

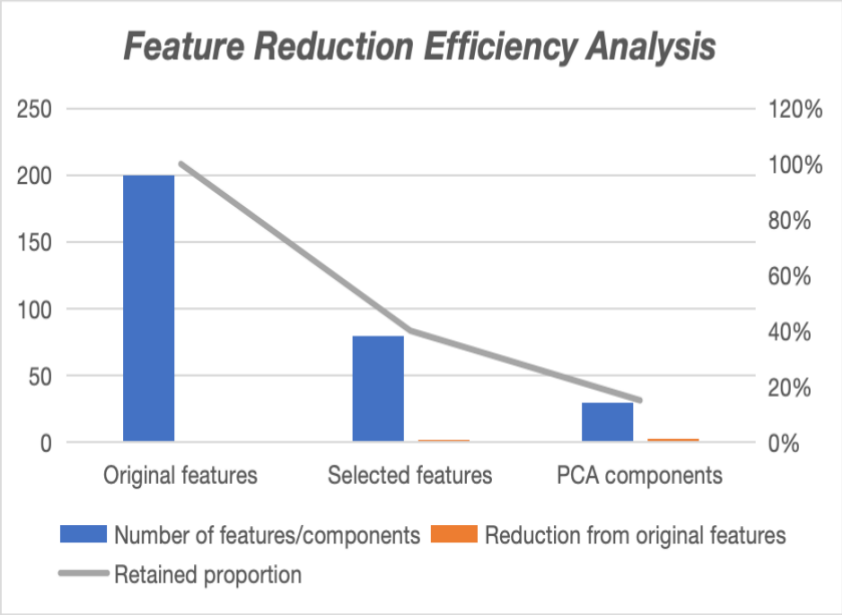


Fig. 7: Feature Reduction Efficiency Analysis

Table 11: Comparative Analysis of Model Strengths and Limitations

Model	Main strength	Potential limitation	Reported accuracy
SVM	Effective for high-dimensional decision boundaries	Sensitive to feature scaling and parameter selection	91%
Random Forest	Captures nonlinear relationships and feature interactions	Can require substantial memory for large ensembles	93%
Gradient Boosting	Sequentially learns to correct previous model errors	May require careful tuning to control overfitting	95%

Each of the three algorithms has some learning capabilities. SVM can be employed in high-dimensional classification, Random Forest can model nonlinear interactions and Gradient Boosting can progressively refine the predictions by concentrating on the residuals. The results indicated Gradient Boosting as the best accuracy. These differences justify the study of hybrid architectures of complementary learning mechanisms.

Table 12: Model Selection Based on Research Objectives

Research objective	Analysis technique	Evaluation criterion
Evaluate predictive performance	Comparative model evaluation	Accuracy, F1-score, ROC-AUC
Examine feature selection and engineering	Original versus selected features	Change in predictive metrics
Assess model diversity and generalization	Ensemble comparison and cross-validation	Prediction diversity and validation performance

The analysis techniques correspond to the three research objectives. Comparative evaluation is used to assess the predictive performance, feature-based is used to compare the contribution of feature preparation, and the evaluation of the model diversity and generalization is investigated through ensemble evaluation. The third goal should not be met by simply using accuracy; it needs to be done with extra measurements of prediction diversity and out-of-sample performance.

Discussion

To boost prediction accuracy, hybrid machine learning architectures can be used that leverage the advantages of various algorithms. The results revealed that Gradient Boosting method had the best accuracy (95%), followed by Random Forest (93%) and Support Vector Machine (91%). The results have shown that tree-based ensemble methods have the ability to detect complex and nonlinear relationships in the data. But the results are individual model performance and we have yet to prove that a hybrid architecture is superior.

All three models were found to be more accurate following feature selection. The accuracy of SVM rose from 91% to 93%, Random Forest from 93% to 94%, and Gradient Boosting from 95% to 96%. This improvement suggests that models could be improved by discarding irrelevant and redundant features so that they can focus on the most informative features. The significance of feature selection has also been emphasized in previous studies for boosting model accuracy and for reducing the complexity of the data sets (Shahnazaryan et al., 2024).

The selected feature set consisted of 80 features which were reduced to 30 features by Principal Component Analysis (PCA). But the results of SVM, RF and GB were found to be 93% to 92%, 94% to 91% and 96% to 93% respectively. These results indicate that dimensionality reduction can result in a simpler data representation, but also in the loss of information pertinent for classification. So, PCA can be used if the computational or dimensionality-reduction advantages are worth sacrificing some predictive accuracy (Azma et al., 2026).

Model diversity is also suggested as a possible means of strengthening the hybrid architecture, according to the results. SVM is able to learn complex decision boundaries, while random forest and gradient boosting learn nonlinear patterns with the help of trees. Predictions made using a stacking or ensemble of the predictions might have better generalization, but that does not have to be assumed, it must be tested empirically (Salman et al., 2024).

The results of cross validation also confirmed the good performance of Gradient Boosting with 95% accuracy and 0.98 ROC-AUC. Cross-validation can be used to test the ability to predict a model consistently on various partitions of a data set. However, parameter tuning and model selection must be performed carefully to ensure that the performance estimates are not unnecessarily high.

The classification error results demonstrated that Gradient Boosting has the minimum overall error rate (5%) followed by Random Forest (7%) and SVM (9%). It achieved 4% false-positive rate as opposed to 5% for Random Forest and 6% for SVM. According to these results, Gradient Boosting had the best classification performance of the single models. The values of the confusion matrix, however, should be re-calculated so that the accuracy, precision, recall and error rates reported are mathematically consistent (Okoh et al., 2022).

The results show the significance of using informative features, comparing different learning algorithms, and testing the models using cross validation. A real hybrid model, for instance, stacking SVM, Random Forest and Gradient Boosting, should be implemented and tested on an independent test set for future research. Further analysis should also include confidence intervals, statistical comparisons, and measures of model diversity, to see if the hybrid method offers a sound improvement over individual classifiers.

Conclusion

The study focused on the prediction ability of machine learning models to classify complex and non-linear patterns correctly. The results revealed that Gradient Boosting had the best predictive accuracy (95 %), followed by Random Forest (93 %), and Support Vector Machine (91 %). The key message is that feature selection enhanced the accuracy of all three models, which are essential for achieving accurate predictions, highlighting the significance of selecting relevant features for effective prediction. By using Principal Component Analysis, however, the predictive performance was reduced, indicating that the reduction of dimensionality may result in the loss of useful information. The cross validation also showed that Gradient Boosting was performing well in all the data partitions. In general, the results underscore the need for careful choice of algorithms and pre-processing methods in order to enhance the accuracy of classification and the reliability of the model.

In conclusion, the study emphasizes that feature selection, model evaluation, and careful algorithm selection are important for developing effective machine learning solutions for complex datasets. Among the classifiers tested, Gradient Boosting was better than the others, but in future, it would be interesting to determine if an SVM-RF-GB hybrid architecture could be better. Real empirical hybrid models should be tested with real independent data, adapted cross-validation methods and statistical significance tests in the future. Further research on computational efficiency, interpretability of the models, and model generalization to other data environments is also required. These may help to develop more sophisticated and flexible machine learning architectures for more complicated and non-linear prediction problems.

Future Implication

For future work, hybrid machine learning architectures, such as the combination of Support Vector Machine, Random Forest and Gradient Boosting, should be developed and tested for the prediction of complex and nonlinear data patterns. Advanced feature selection techniques, feature engineering methods, and dimensionality reduction approaches should be explored to find the most informative features while retaining important information. Incorporating more and varied data sets, impartial testing, and strong cross-validation methods can assist evaluate the reliability and generalization of the model in other contexts. Additionally, future research could explore stacking and ensemble learning strategies, fine-tune model parameters, and integrate explainable AI techniques for greater model transparency and interpretability.

References

1. Al Reshan, M. S., Amin, S., Zeb, M. A., Sulaiman, A., Alshahrani, H., & Shaikh, A. (2023). A robust heart disease prediction system using hybrid deep neural networks. *IEEE Access*, *11*, 121574-121591.
2. Mouad, B. H. I. H., Abou Elassad, Z. E., Siti, H., Meslouhi, O. E., Abouelmehdi, K., & El Boustani, A. (2026). Temporal pattern-aware temperature forecasting using CatBoost: A hybrid machine learning approach. *Results in Engineering*, 110212.

3. Khayyam, H., Jamali, A., Bab-Hadiashar, A., Esch, T., Ramakrishna, S., Jalili, M., & Naebe, M. (2020). A novel hybrid machine learning algorithm for limited and big data modeling with application in industry 4.0. *IEEE access*, *8*, 111381-111393.
4. Jamshidi, M., Roshani, S., Daneshfar, F., Lalbakhsh, A., Roshani, S., Parandin, F., ... & Lalbakhsh, P. (2022). Hybrid deep learning techniques for predicting complex phenomena: A review on COVID-19. *AI*, *3*(2), 416-433.
5. Singh, R., Bengani, V., & Saini, K. (2024). Hybrid learning systems: Integrating traditional machine learning with deep learning techniques. *ResearchGate*.
6. Fozap, F. M. P. (2025). Hybrid machine learning models for long-term stock market forecasting: Integrating technical indicators. *Journal of Risk and Financial Management*, *18*(4), 201.
7. Bahn, V., & McGill, B. J. (2013). Testing the predictive performance of distribution models. *Oikos*, *122*(3), 321-331.
8. Richardson, M. J., Paxton, A., & Kuznetsov, N. (2017). Nonlinear methods for understanding complex dynamical phenomena in psychological science. *Psychological Science Agenda*, *31*(2).
9. Tabar, M. R. R. (2019). *Analysis and data-based reconstruction of complex nonlinear dynamical systems: using the methods of stochastic processes*. Springer.
10. Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *374*(2065), 20150202.
11. Kuhn, M., & Johnson, K. (2013). An introduction to feature selection. In *Applied predictive modeling* (pp. 487-519). New York, NY: Springer New York.
12. Zhou, Y., Shen, J., & Cheng, Y. (2025, May). Weak to strong generalization for large language models with multi-capabilities. In *International Conference on Learning Representations* (Vol. 2025, pp. 11583-11612).
13. Teney, D., Abbasnejad, E., Lucey, S., & Van den Hengel, A. (2022, June). Evading the simplicity bias: Training a diverse set of models discovers solutions with superior ood generalization. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 16740-16751). IEEE.
14. Nelson, D. L., Kielhofner, G., & Taylor, R. R. (2017). Defining Variables and Their Rela-tionships With One Another. *Kielhofner's research in occupational therapy: Methods of inquiry for enhancing practice*, *244*.
15. Schjoedt, L., Sangboon, K., & Bird, B. (2025). Control variables: Use, misuse, and recommended use. In *Handbook of Research Methods and Applications in Entrepreneurship and Small Business, Second Edition* (pp. 186-204). Edward Elgar Publishing.
16. Parviz, L., Rasouli, K., & Torabi Haghighi, A. (2023). Improving hybrid models for precipitation forecasting by combining nonlinear machine learning methods. *Water Resources Management*, *37*(10), 3833-3855.
17. Jean, G. (2026). Hybrid Deep Learning Architectures as a Paradigm for Modeling Nonlinear, High-Dimensional Systems.
18. Mahmoud, A., & Mohammed, A. (2024). Leveraging Hybrid Deep Learning Models for Enhanced Multivariate Time Series Forecasting: A. Mahmoud, A. Mohammed. *Neural Processing Letters*, *56*(5), 223.
19. Agga, A., Abbou, A., Labbadi, M., El Houm, Y., & Ali, I. H. O. (2022). CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. *Electric Power Systems Research*, *208*, 107908.
20. Çılgın, C., & Gökçen, H. (2025). A hybrid machine learning model architecture with clustering analysis and stacking ensemble for real estate price prediction. *Computational Economics*, *66*(1), 127-178.
21. Osatohanmwen, B. E., Vieira, I. C., Gholami, M., Westhues, C. C., Sharifi, A. R., & Beissinger, T. M. (2026). Predicting maize hybrid performance with machine learning and a locus-specific weighted degree of dominance transformation. *Frontiers in Plant Science*, *17*, 1694707.

22. Miah, M. R., & Islam, M. M. (2024). Data Preprocessing and Feature Engineering Strategies for Large-Scale Predictive Modeling Applications. *Review of Applied Science and Technology*, *3*(01), 263-302.

23. Haque, B. T., & Rahman, M. A. (2024). Data Preparation and Feature Engineering in High-Dimensional Predictive Modeling. *Journal of Sustainable Development and Policy*, *3*(04), 205-244.

24. Katya, E. (2023). Exploring feature engineering strategies for improving predictive models in data science. *Research Journal of Computer Systems and Engineering*, 201-215.

25. Wang, J., Kwon, H., Oh, S. H., Kim, H. S., & Lee, B. J. (2025). Interpretable and physics-informed modeling of solidification in alloy systems: a generalized framework for multi-component prediction. *Acta Materialia*, *286*, 120716.

26. Arpit, D., Wang, H., Zhou, Y., & Xiong, C. (2022). Ensemble of averages: Improving model selection and boosting performance in domain generalization. *Advances in Neural Information Processing Systems*, *35*, 8265-8277.

27. Ortega, L. A., Cabañas, R., & Masegosa, A. (2022, May). Diversity and generalization in neural network ensembles. In *International Conference on Artificial Intelligence and Statistics* (pp. 11720-11743). PMLR.

28. Nikolikj, A., Cenikj, G., Ispirova, G., Vermetten, D., Lang, R. D., Engelbrecht, A. P., ... & Eftimov, T. (2023, July). Assessing the generalizability of a performance predictive model. In *Proceedings of the companion conference on genetic and evolutionary computation* (pp. 311-314).

29. Haryana, R. (2023). Improving the performance of hybrid models using machine learning and optimization techniques. *International Journal*, *10*(2), 3396-3409.

Mask, G. M. (2025). *Feature Engineering and Hybrid Modeling in Machine*

2026 by the authors; Journal of Global Social Transformation (JGST). This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).