Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15509

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Comparison of XGBoost and Naive Bayes Models in Type 2 Diabetes Prediction with RFE Feature Selection

Hanisa Putri Barus[1], Robet[2] *, Feriani Astuti Tarigan[3]
[1,2] Department of Informatics, STMIK Time, Medan, Indonesia,
[3] Department of Information Systems, STMIK Time, Medan, Indonesia
[1]hanisaputri7662@gmail.com, [2] *robertdetime@gmail.com, [3]ferianiastutitime@gmail.com

Submitted :Oct 29, 2025 | **Accepted** : Nov 20, 2025 | **Published** : Jan 03, 2026

Abstract: Type 2 diabetes mellitus is a chronic disease with an increasing prevalence rate that can cause serious complications if not detected early. The application of machine learning algorithms can aid prediction, but selecting the right model and features greatly determines the accuracy of the results. This study aims to compare the performance of the Extreme Gradient Boosting (XGBoost) and Naive Bayes algorithms in predicting type 2 diabetes with and without Recursive Feature Elimination (RFE) feature selection. The data used were from the UCI Machine Learning Repository, comprising 768 samples and eight clinical features. The research process included data preprocessing, dividing the data into 614 training data and 154 testing data, applying RFE to select the most influential features, model training, and evaluation using accuracy, precision, recall, F1-score, and AUC. The results show that Naive Bayes without RFE achieves 70.77% accuracy, 0.57377 precision, 0.648148 recall, F1-score 0.608696, and 0.772778 AUC, while Naive Bayes with RFE increases the accuracy to 74.02% and the AUC to 0.793333. Meanwhile, XGBoost with RFE provided the best results with an accuracy of 74.67%, precision of 0.653061, recall of 0.592593, F1-score of 0.621359, and the highest AUC of 0.804259. Besides, applying RFE also improves the computational efficiency. These findings indicate that applying RFE significantly improves classification and computation time performance. The practical implication is that this model could aid early detection of diabetes in clinical settings. Further research can be conducted by optimizing parameters and using more diverse datasets.

Keywords: Feature Selection; Naive Bayes; RFE; Type 2 Diabetes; XGBoost

INTRODUCTION

Diabetes mellitus is a non-communicable disease whose prevalence continues to increase worldwide. Disorders of glucose metabolism cause this condition due to the pancreas's inability to produce insulin or the body's resistance to insulin. The increase in the number of diabetes patients, especially type 2, poses a serious challenge in the health sector because it can trigger long-term complications such as kidney failure, heart disease, and stroke (Naki et al., 2025). Therefore, a computational approach is needed to enable early detection, allowing medical treatment to be carried out quickly and accurately.

In recent years, machine learning has been widely used in the development of disease prediction systems, including those for diabetes mellitus. Various algorithms, such as Naïve Bayes, Decision Trees, Random Forests, and XGBoost, have strong potential for classifying medical data (Jawza et al., 2025). However, differences in model performance remain a concern, especially when applied to datasets with complex characteristics and many irrelevant features.

Several studies show that feature selection is an essential factor in improving classification model accuracy. Methods such as Recursive Feature Elimination (RFE), Backward Elimination, and Principal Component Analysis (PCA) have been used to filter out the attributes that most influence prediction results, thereby reducing noise and improving computational efficiency (Idris et al., 2024). Therefore, combining classification algorithms and feature selection techniques has the potential to produce more optimal models.

One of the algorithms widely used in chronic disease prediction is Extreme Gradient Boosting (XGBoost). This method is known for its ability to handle large datasets, overcome overfitting, and achieve high performance

---

*name of corresponding author

compared to other algorithms (Susanto et al., 2025). In addition, other studies also suggest that XGBoost excels because it combines an adaptive boosting process that repeatedly corrects prediction errors, resulting in more accurate final results (Erkamim et al., 2023).

However, some studies show that Naïve Bayes remains relevant for use due to its simplicity, high interpretability, and speed in classifying datasets with a large number of features. This model can achieve competitive results in cases of diabetes diagnosis with low to medium complexity (Khurshid et al., 2025). Therefore, it is crucial to conduct empirical comparative research to observe the performance differences between XGBoost and Naïve Bayes.

In addition to algorithm comparison, previous studies have also highlighted the importance of optimization and attribute selection processes in improving prediction performance. Mutual information-based and random forest feature importance approaches have proven effective in determining the attributes that most influence classification results (Goyal et al., 2025). This technique can be adapted to improve model performance in medical cases, particularly in type 2 diabetes prediction.

Based on this review, this study aims to compare the performance of the XGBoost and Naïve Bayes algorithms in predicting type 2 diabetes using the RFE feature selection technique. This study is expected to contribute to the development of artificial intelligence-based medical decision support systems that can assist healthcare professionals in early diagnosis and clinical decision-making (Aqmar et al., 2025)**.** Thus, the results of this study can serve as a basis for further research aimed at improving model accuracy and generalization to broader medical datasets.

LITERATURE REVIEW

Many studies have been conducted on the application of machine learning algorithms for diabetes prediction, with a variety of focuses ranging from algorithm selection and feature selection to model optimization. One study shows that using Naïve Bayes with the Greedy Forward Selection approach can improve classification accuracy in medical datasets by eliminating irrelevant features early in the learning process (Fitriyani, 2021). This approach has proven simple yet effective for optimizing the performance of probabilistic models.

Meanwhile, another study highlights the importance of comprehensively evaluating the performance of machine learning algorithms in diabetes classification. The combination of algorithms such as Random Forest, Support Vector Machine (SVM), and Logistic Regression shows different performances depending on the feature selection method used (Anasanti et al., 2022). The results of this study indicate that applying an appropriate feature selection method leads to a notable improvement in model accuracy compared to models without attribute selection.

Previous research indicates that combining Naïve Bayes with Backward Elimination can simplify the classification model (Wiratama & Pradnya, 2022). These results confirm that algorithm efficiency depends not only on the model, but also on the quality of the features used.

In the context of ensemble models, XGBoost is one of the most popular algorithms for chronic disease classification. A study combined Principal Component Analysis (PCA) with XGBoost to detect diabetes early, achieving high accuracy (Wardhani & Novayani, 2024). This combination works well because PCA can reduce the dimensionality of the data without losing essential information, while XGBoost maximizes performance through adaptive boosting.

Other studies have also examined the effectiveness of Gradient Boosting in classifying type 2 diabetes (Syahputra & Saputro, 2024). The results of the literature review confirm that boosting methods, including XGBoost and AdaBoost, have advantages for overcoming data imbalance and can improve model performance iteratively. However, most of these studies have not directly compared probabilistic algorithms, such as Naïve Bayes, on the same dataset.

In addition, studies highlight the influence of activation functions on the performance of neural network models on diabetes data (Wijaya Kusuma et al., 2025). The results show that selecting an appropriate activation function can significantly affect the convergence rate and accuracy. This illustrates that each model has unique characteristics that need to be adjusted to the type of data and analysis objectives.

In terms of model optimization, the soft voting ensemble approach combined with the SMOTE-ENN oversampling technique and Bayesian optimization yields more stable, accurate predictions in chronic disease classification, including diabetes (Maulana & Ernawati, 2025). However, this research focuses more on combining multiple models than on comparing the performance of individual algorithms.

Another study also used mutual information and random forest feature importance methods to identify the most significant attributes in the machine learning-based classification process (Alqahtani et al., 2024). This approach has high potential for application in the medical field, as it can help identify the clinical variables that most influence diagnostic results.

Based on the above studies, it can be concluded that there is still a research gap in terms of direct comparison between the XGBoost and Naïve Bayes algorithms with the application of the RFE feature selection technique in the case of type 2 diabetes prediction (Nemer et al., 2025). Most studies only examine one algorithm or use different feature selection methods. Therefore, this study focuses on filling this gap by analyzing the performance of both algorithms on the same dataset and evaluating the impact of RFE implementation on improving model accuracy and efficiency.

METHOD

Research Stages

This study uses a quantitative, comparative experimental approach to evaluate the performance of two Machine Learning algorithms, namely Extreme Gradient Boosting (XGBoost) and Naive Bayes, in predicting the risk of type 2 Diabetes Mellitus. In addition, the Recursive Feature Elimination (RFE) feature selection technique was used to evaluate the effect of feature reduction on the performance of both models.

This study was conducted through a series of systematic stages designed to build a Type 2 Diabetes Mellitus prediction model with optimal accuracy. These stages include:

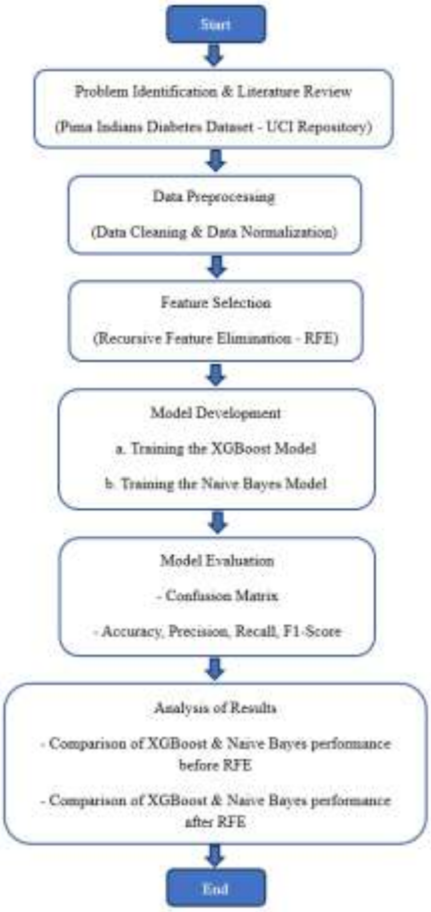

Figure 1. Research Stages

Problem Identification and Literature Review

The initial stage of the research began with problem identification and a literature review, namely by collecting information on diabetes prediction and relevant previous studies (Anasanti et al., 2022). The dataset used was Pima Indians Diabetes, available at the UCI Machine Learning Repository and on the Kaggle platform. This dataset contains medical data on female Pima Indian patients aged 21 years and older. In total, there are 768 samples, consisting of 8 independent variables and 1 target variable (Outcome).

Table 1. Independent Attributes Include

Attribute Name	Data Type	Brief Description
Pregnancies	Integer	Number of pregnancies experienced
Glucose	Integer	Plasma glucose level (mg/dL)
BloodPressure	Integer	Diastolic blood pressure (mm Hg)
SkinThickness	Integer	Triceps skinfold thickness (mm)

Insulin	Integer	Serum insulin level (mu U/ml)
BMI	Float	Body Mass Index (kg/m²)
DiabetesPedigreeFunction	Float	Genetic history of diabetes
Age	Integer	Respondent age (years)
Outcome	Integer	Diagnosis label (0 = Negative, 1 = Positive)

Pre-Processing Data

Data preprocessing is performed to ensure the dataset's quality before it is used in the model training process. This stage includes several steps, namely handling missing values, data normalization, and data division. Several attributes, such as Glucose, Blood Pressure, Skin Thickness, Insulin, and BMI, have zero values, which are considered missing values. These values are imputed using the mean imputation method for numerical attributes, ensuring the data distribution remains consistent. Next, all numerical features are normalized using the Min-Max Scaling method to the range 0 to 1. This normalization process aims to speed up model training and prevent features with larger scales from dominating. After that, the dataset, consisting of 768 samples, was split into 80% for training (614 samples) and 20% for testing (154 samples). This division was carried out using stratified sampling to maintain a balanced distribution of target classes between the training and test data (Parvez & Mufti, 2025).

Feature Selection

Feature selection in this study was performed using the Recursive Feature Elimination (RFE) method to identify the features most strongly influencing diabetes prediction. The RFE process begins by training the initial model using all available features, then calculating the importance ranking of each feature based on the model weight. After that, the features with the lowest contribution are removed, and the training process is repeated. This stage is repeated iteratively until the optimal number of features that achieve the best accuracy is obtained (Siringoringo et al., 2024). The formula for assessing feature importance in RFE can be written as follows:

$$S_J = \frac{|w_j|}{\sum_{k=1}^{n}|w_k|} \quad (1)$$

With:
S_j = feature importance score j
w_j= feature weight j obtained from the model
n = total number of features

Classification Model Training

Classification models in this study were developed by comparing two algorithms: Extreme Gradient Boosting (XGBoost) and Naive Bayes. XGBoost is a gradient-boosting ensemble algorithm that combines multiple weak decision trees into a single strong model. The advantages of XGBoost include its ability to handle non-linear relationships, regularization techniques to reduce the risk of overfitting, and support for parallel processing, which speeds up training (Erkamim et al., 2023). Mathematically, predictions in XGBoost can be written as:

$$y\hat{}_i = \sum_{k=1}^{K} f_k\,(x_i), f_k \in F \quad (2)$$

Where y^i is the prediction for the i^(th) data point, f_k is the decision tree at iteration K, and K is the total number of trees.

Meanwhile, Naive Bayes is a probabilistic classification algorithm based on Bayes' theorem, assuming feature independence (Fitriyani, 2021). The general formula is written as:

$$P(C_k \mid x) = \frac{P(x|C_k).p(C_k)}{P(x)} \quad (3)$$

where $P(C_k \mid x)$is the probability of class C_k given feature x, $P(C_k \mid x)$is the probability of x appearing in class C_k, and $p(C_k)$ is the initial probability of class C_k.

Both models will be trained and tested in two scenarios: using all original features and using features selected with RFE, to compare their performance in predicting type 2 diabetes.

Model Performance Evaluation

The performance of the models in this study was evaluated to assess each algorithm's ability to accurately and consistently predict Type 2 diabetes. Several evaluation metrics were used, namely accuracy, precision, recall, F1-score, and Area Under Curve (AUC) (Jawza et al., 2025). Accuracy was used to measure the percentage of correct predictions from all test data. Precision measured the proportion of optimistic predictions that were actually positive, while recall measured the model's ability to detect all positive cases. F1-score is a harmonized value between precision and recall, which is very useful in datasets with unbalanced class distributions. Meanwhile, AUC measures the model's ability to distinguish between positive and negative classes at different thresholds, with higher values indicating better discrimination. The evaluation was conducted on both models in two scenarios, namely using all original features and using RFE-selected features. The evaluation results will be presented in tables, confusion matrices, and ROC curves to provide a comprehensive overview of each model's performance. The calculation formulas for each metric are as follows:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{4}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FN}} \tag{5}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{6}$$

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision+Recall}} \tag{7}$$

$$\text{AUC} = \frac{1+\text{TPR}-\text{FPR}}{2} \tag{8}$$

a. True Positive (TP): Positive data that is correctly predicted as positive.
b. True Negative (TN): Negative data that is correctly predicted as negative.
c. False Positive (FP): Negative data that is incorrectly predicted as positive.
d. False Negative (FN): Positive data that is incorrectly predicted as negative.

RESULT

Overview of Research Results

This section presents a summary of the results from experiments using two classification algorithms: Naive Bayes and XGBoost. The evaluation process was carried out on both the initial model and the model after applying the feature selection method, Recursive Feature Elimination (RFE).

Classification Metric Evaluation

Classification metrics were evaluated to compare the performance of the Naive Bayes and XGBoost models, both before and after feature selection using Recursive Feature Elimination (RFE). The evaluation results are shown in Table 2 and Figure 2.

Table 2. Classification Metric Evaluation Results

	Accuracy	**Precision**	**Recall**	**F1-Score**	**AUC**
Naive Bayes (Tanpa RFE)	0.707792	0.573770	**0.648148**	0.608696	0.772778
XGBoost (Tanpa RFE)	0.727273	0.611111	0.611111	0.611111	**0.804259**
Naive Bayes (RFE)	0.740260	0.634615	0.611111	0.622642	0.793333
XGBoost (RFE)	**0.746753**	**0.653061**	0.592593	**0.621359**	0.804259

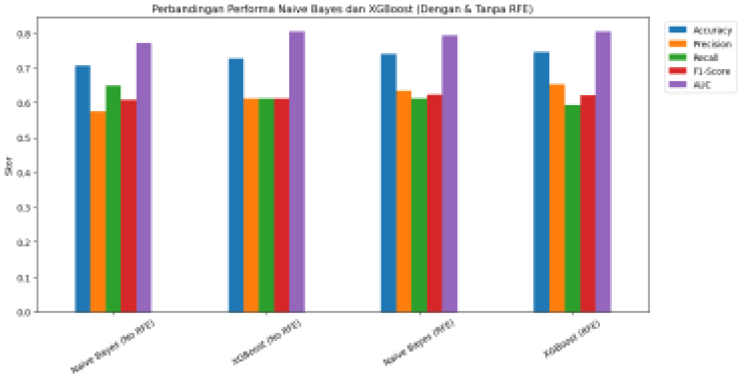

Figure 2. Comparison of Naive Bayes and XGBoost Performance With and Without Feature Selection (RFE)

Based on the results in Table 2 and Figure 2 above, which show a comparison of the performance of the Naive Bayes and XGBoost algorithms before and after the application of RFE. The results indicate that XGBoost performs better on most metrics. The application of RFE also improved accuracy, precision, and F1-Score, while recall remained stable, and AUC showed small fluctuations. Thus, XGBoost combined with RFE yields the best results among the models.

Confusion Matrix

A confusion matrix is used to describe a model's performance at distinguishing between positive and negative classes on test data. This matrix contains four main components:

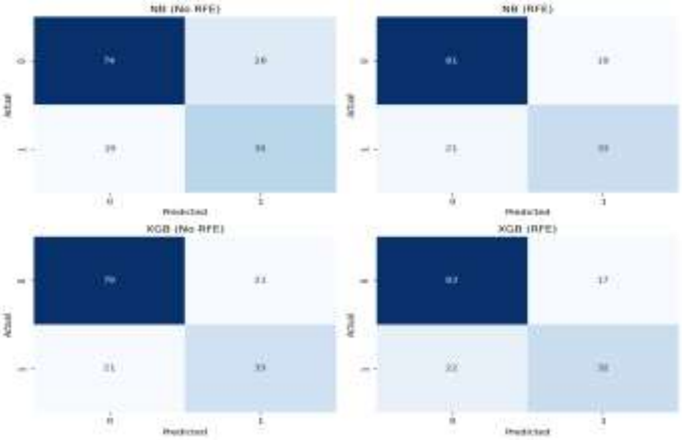

Figure 3. Confusion Matrix from Four Model Testing Scenarios

The results of the model performance evaluation are shown in the confusion matrix in Figure 3 above. In general, applying the Recursive Feature Elimination (RFE) method improves prediction quality, particularly by reducing the number of false positives.

In the Naive Bayes model without RFE, the number of true negatives (TN) was 74, true positives (TP) was 35, and false positives (FP) was 26. After applying RFE, the TN value increased to 81 and the FP decreased to 19, although the TP decreased slightly to 33. This shows that RFE helps improve the accuracy of negative predictions by reducing false positive classification errors.

Meanwhile, XGBoost shows more stable performance than Naive Bayes. Without RFE, the model produced 79 TN, 33 TP, 21 FP, and 21 FN. After using RFE, the performance improved with an increase in TN to 83 and a decrease in FP to 17. Although TP decreased slightly to 32, these results confirm that the combination of XGBoost and RFE provides the best balance, especially in minimizing false positive prediction errors.

Overall, it can be concluded that XGBoost with RFE provides the best performance, as it improves prediction reliability and reduces errors.

ROC Curve Model Comparison

Figure 4 shows the ROC curve comparing the performance of the Naive Bayes and XGBoost models, both before and after applying Recursive Feature Elimination (RFE). In general, the Area Under the Curve (AUC) indicates that XGBoost has better classification performance than Naive Bayes. The Naive Bayes model without RFE achieved an AUC of 0.773, indicating fairly good performance but still lower than that of other models. After feature selection with RFE, the AUC value increased to 0.793. This shows that RFE can improve Naive Bayes' discriminatory ability, though the improvement is relatively small.

Unlike Naive Bayes, the XGBoost model has shown high performance from the start. Without RFE, this model achieved an AUC of 0.804. The application of RFE did not significantly change the AUC, which remained at 0.804. However, based on the confusion matrix analysis, RFE still provided benefits, particularly in reducing false-positive prediction errors.

Thus, it can be concluded that XGBoost excels in class differentiation compared to Naive Bayes. The combination of XGBoost and RFE is the best choice because it maintains high AUC while reducing the potential for classification errors.

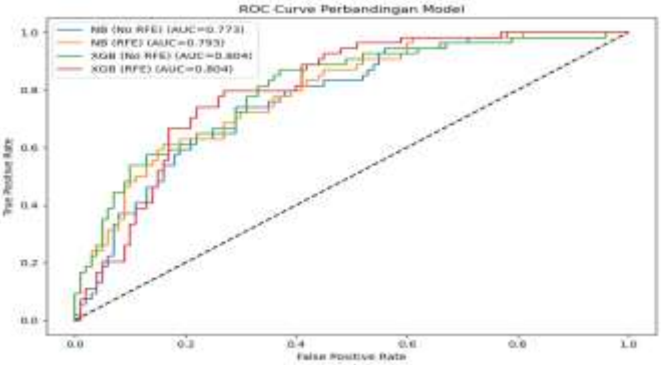

Figure 4. ROC Curve Comparison of the Performance of Four Model Configurations

Training vs. Validation Loss (XGBoost)

Figure 5 compares the training and validation losses of the XGBoost model after applying Recursive Feature Elimination (RFE). The graph shows that the training loss decreases steadily as the number of boosting rounds increases, indicating that the model effectively learns patterns in the training data. At the 50th iteration, the training loss approaches 0.1, indicating high accuracy on the training data.

However, the validation loss shows a different pattern. Instead of decreasing, the validation loss remains relatively stable at the beginning of training and then increases slowly, eventually exceeding 0.6. This phenomenon indicates overfitting, in which the model adapts too closely to the training data, leading to poor performance on new data.

Thus, although XGBoost + RFE can achieve optimal performance on the training data, these results underscore the importance of applying additional regularization techniques or early stopping to maintain the model's generalization ability.

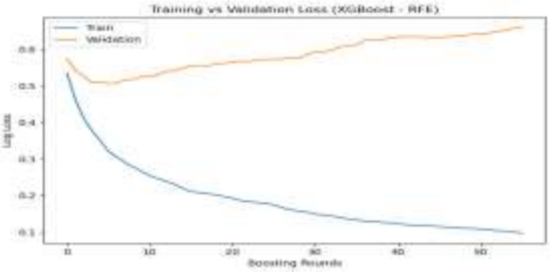

Figure 5. Training vs. validation loss (XGBoost)

Selected RFE Features

RFE is a feature selection algorithm that repeatedly filters attributes based on each feature's significance to the model. This method identifies the most influential feature combination that maximizes model performance while minimizing computational complexity. The application of RFE allows analysis to focus only on key features, thereby reducing the risk of the model becoming too complex (overfitting). The selected features are shown in Table 3.

Table 3. Selected Features Using RFE

Model	Selected Features
Naive Bayes	Pregnancies, Glucose, BloodPressure, BMI, DiabetesPedigreeFunction
XGBoost	Pregnancies, Glucose, Insulin, BMI, Age

The Recursive Feature Elimination (RFE) method selects five key features for Naive Bayes and five features for XGBoost. This confirms that they are the most relevant indicators in predicting type 2 diabetes. By using only selected features, the model can work more effectively, focusing on the variables that most influence classification results.

Model Training Time

Model performance reflects the total training and testing duration required for each algorithm. This measurement is essential for assessing computational efficiency, enabling a comprehensive evaluation of both accuracy and prediction speed. A comparison of the execution times of the two models can be seen in Table 4 below:

Table 4. Classification Metric Evaluation Results

Model	Training Time (s)	Prediction Time (s)
Naive Bayes (without RFE)	0.039428	0.003364
Naive Bayes (RFE)	0.001445	0.005156
XGBoost (without RFE)	0.311227	0.004337
XGBoost (RFE)	0.063147	0.002074

Based on temporal measurement results, implementing RFE reduced computation time in both algorithms, with a larger reduction in XGBoost, which previously had a longer training duration. In Naive Bayes, training time decreased, though a slight decrease followed this in prediction speed. In contrast, XGBoost showed consistent efficiency improvements in both the training and prediction phases after feature selection.

DISCUSSIONS

The results of the experiment show that the XGBoost algorithm consistently outperforms Naive Bayes across accuracy, precision, and F1-Score. This aligns with the characteristics of XGBoost as an ensemble boosting algorithm that can model nonlinear and complex relationships among variables, making it more effective at handling medical datasets such as the Pima Indians Diabetes dataset. Meanwhile, Naive Bayes, which assumes feature independence, tends to be less capable of capturing the complexity of attribute relationships, resulting in lower performance.

The application of Recursive Feature Elimination (RFE) has different effects across algorithms. In Naive Bayes, RFE has been proven to improve accuracy and precision compared to not using RFE. This indicates that selecting more relevant features helps reduce noise and improve prediction quality. Conversely, in XGBoost, although the accuracy increased to 74.67% and the precision to 0.65, the AUC decreased slightly compared to without RFE. However, the most significant advantage of RFE in XGBoost is computational efficiency. The training time, which was initially relatively long (0.311 seconds), can be reduced to 0.063 seconds, and the prediction time is also shorter than that of other models. This shows that RFE effectively speeds up the computational process without significantly compromising prediction quality.

When considering the trade-off between accuracy and efficiency, XGBoost with RFE is the most optimal choice. This model not only provides the best classification results but also has a more practical execution speed for real-world implementation, for example, in medical data-based early diabetes detection systems. Naive Bayes with RFE can still be considered if the research focus is on speed and model simplicity, even though it has lower accuracy than XGBoost.

CONCLUSION

Based on the research, it can be concluded that applying the Recursive Feature Elimination (RFE) feature selection method significantly improves model performance for predicting type 2 diabetes. In the Naive Bayes model, applying RFE increased accuracy from 70.7792% to 74.026% and AUC from 0.772778 to 0.793333. Meanwhile, the XGBoost algorithm showed the best performance after applying RFE, with an accuracy of 74.6753% and an AUC of 0.804259. These results show that XGBoost with RFE provides a better balance between performance on positive and negative classes than other models.

In addition to improving accuracy, applying RFE also improved computational efficiency. The training time for XGBoost decreased from 0.311227 seconds to 0.063147 seconds, and the prediction time decreased from 0.004337 seconds to 0.002074 seconds. This shows that RFE not only improves feature relevance but also speeds up model training and prediction.

However, this study still has limitations: it uses only one dataset and does not conduct in-depth parameter optimization. Therefore, further research is recommended to use a more diverse dataset, apply hyperparameter tuning, and compare the model with other machine learning algorithms. Overall, XGBoost with RFE can be recommended as an effective and efficient model to support early detection of type 2 diabetes in a clinical practice setting.

REFERENCES

Alqahtani, S. A. M., Alobaid, H. M., Alshammari, J., Alqarzae, S. A., Aloyouni, S. Y., Al-Eidan, A. A., Alhamad, S., Almiman, A., Alkhulaifi, F. M., & Alomar, S. (2024). Feature importance and model performance for prediabetes prediction: A comparative study. *Journal of King Saud University - Science*, *36*(11). https://doi.org/10.1016/j.jksus.2024.103583

Anasanti, M. D., Hilyati, K., & Novtariany, A. (2022). Exploring feature selection techniques on Classification Algorithms for Predicting Type 2 Diabetes at Early Stage. *Jurnal RESTI*, *6*(5), 832–839. https://doi.org/10.29207/resti.v6i5.4419

Aqmar, N., Wijayanto, H., & Afendi, F. M. (2025). *Performance Analysis of Machine Learning Models using RFE Feature Selection and Bayesian Optimization in Imbalanced Data Classification with Shap-Based Explanations*. *12*(3), 539–554. https://doi.org/10.15294/sji.v12i3.31459

Erkamim, M., Suswadi, S., Subarkah, M. Z., & Widarti, E. (2023). Komparasi Algoritme Random Forest dan XGBoosting dalam Klasifikasi Performa UMKM. *Jurnal Sistem Informasi Bisnis*, *13*(2), 127–134. https://doi.org/10.21456/vol13iss2pp127-134

Fitriyani, F. (2021). Prediksi Diabetes Menggunakan Algoritma Naive Bayes dan Greedy Forward Selection. *Jurnal Nasional Teknologi Dan Sistem Informasi*, *7*(2), 61–69. https://doi.org/10.25077/teknosi.v7i2.2021.61-69

Goyal, D., Singh, J., & Vashist, A. (2025). *Advanced Machine Learning Models Diabetes Risk Prediction Using Feature Selection Algorithms and Advanced Machine Learning Models*. 1–24.

Idris, N. F., Ismail, M. A., Jaya, M. I. M., Ibrahim, A. O., Abulfaraj, A. W., & Binzagr, F. (2024). Stacking with Recursive Feature Elimination-Isolation Forest for classification of diabetes mellitus. *PLoS ONE*, *19*(5 May), 1–18. https://doi.org/10.1371/journal.pone.0302595

Jawza, D. N., Mazdadi, M. I., Farmadi, A., Saragih, T. H., Kartini, D., & Abdullayev, V. (2025). Enhancing Diabetes Prediction Accuracy Using Random Forest and XGBoost with PSO and GA-Based Feature Selection. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *7*(2), 295–306. https://doi.org/10.35882/jeeemi.v7i2.626

Khurshid, M. R., Manzoor, S., Sadiq, T., Hussain, L., Khan, M. S., & Dutta, A. K. (2025). Unveiling diabetes onset: Optimized XGBoost with Bayesian optimization for enhanced prediction. *PLoS ONE*, *20*(1 January), 1–29. https://doi.org/10.1371/journal.pone.0310218

Maulana, I., & Ernawati, S. (2025). Meningkatkan Klasifikasi Penyakit Diabetes Menggunakan Metode Ensemble Softvoting Dengan SMOTE-ENN dan Optimasi Bayesian. *Evolusi: Jurnal Sains Dan Manajemen*, *13*(1), 71–86.

Naki, M. I., Tambengi, R. A., & Sumariangen, A. B. (2025). *Diabetes Mellitus Tipe 2 : Prevalensi , Etiologi , dan Pelaksanaannya*. *5*(1), 77–87.

Nemer, Z. N., Raheem, S. F., & Alabbas, M. (2025). Comparison of Classification of Different Machine Learning Algorithms in the Diagnosis and Detect of Diabetes. *International Journal of Computing and Digital Systems*, *18*(1). https://doi.org/10.12785/ijcds/1571016484

Parvez, A., & Mufti, M. J. (2025). *Generalizable Diabetes Risk Stratification via Hybrid Machine Learning Models*.

Siringoringo, R., Arisandi, D., Kurniawan, E., & Nababan, E. B. (2024). *Model Klasifikasi Dengan Logistic Regression Dan Recursive Classification Model Using Logistic Regression And Recursive*. *11*(4). https://doi.org/10.25126/jtiik.1148198

Susanto, E. R., Teknik, F., Indonesia, U. T., & Lampung, B. (2025). *Penerapan Algoritma XGBoost untuk Prediksi Diabetes : Analisis Confusion Matrix dan ROC Curve Agum Cahyana Abstrak*. *10*(1).

Syahputra, A. A., & Saputro, R. E. (2024). Application of the XGBoost Model with Hyperparameter Tuning for Industry Classification for Job Applicants. *Sinkron*, *8*(3), 1920–1931. https://doi.org/10.33395/sinkron.v8i3.13840

Wardhani, K. D. K., & Novayani, W. (2024). Principal Component Analysis for Prediabetes Prediction using Extreme Gradient Boosting (XGBoost). *Scientific Journal of Informatics*, *11*(3), 863–872. https://doi.org/10.15294/sji.v11i3.13416

Wijaya Kusuma, A., Mazdadi, M. I., Kartini, D., Farmadi, A., Indriani, F., & P., C. (2025). Improving Diabetes Prediction Using Feedforward Neural Network with Adam Optimization and SMOTE Technique. *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *7*(3), 539–548. https://doi.org/10.35882/ijeeemi.v7i3.127

Wiratama, M. A., & Pradnya, W. M. (2022). *Optimasi Algoritma Data Mining Menggunakan Backward Elimination Untuk Klasifikasi Penyakit Diabetes. Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI | 2. 11*, 1–12.