Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

Improving Diabetes Prediction Performance Using Random Forest Classifier with Hyperparameter Tuning

Novita Lestari Anggreini*[1], Ade Yuliana[2], Dadan Saepul Ramdan[3], Wissam Al-Dayyeni[4]

[1,2,3]Informatics Engineering, Politeknik TEDC, Indonesia
[4]Electrical Engineering, ADA Univresity, Azerbaijan

Email: [1]novitalestari@poltektedc.ac.id

Received : May 19, 2025; Revised : Jun 25, 2025; Accepted : Jun 27, 2025; Published : Aug 18, 2025

Abstract

Diabetes mellitus is a chronic metabolic disorder that poses a serious challenge to global healthcare systems due to its increasing prevalence and the high costs associated with treatment. Although machine learning has been widely adopted to support early diagnosis, many predictive models still underperform due to limited preprocessing strategies and inefficient hyperparameter settings. This study proposes a comprehensive machine learning pipeline to enhance diabetes prediction accuracy by utilizing a Random Forest classifier optimized through systematic hyperparameter tuning. The novelty of this method lies in its integrated approach, which includes thorough preprocessing such as removing duplicate records, handling inconsistent unique values, addressing missing data, and applying the SMOTE technique to overcome class imbalance. Additionally, hyperparameter tuning is conducted using GridSearchCV combined with 5-fold cross-validation, and only the most influential features are selected to improve model interpretability and efficiency. The proposed model achieved an accuracy of 95 percent, with a recall of 0.88 and an F1-score of 0.85, indicating its robustness in identifying diabetic cases more effectively than previous studies using standard machine learning algorithms. This model contributes to the development of a reliable and scalable early detection system for diabetes, applicable in clinical decision support environments. Further refinement can be achieved by testing on larger and more diverse datasets or by implementing more efficient tuning techniques such as Bayesian optimization.

Keywords : *Data Mining, Diabetes Prediction, Hyperparameter Tuning, Model Enhancement, Random Forest Classifier.*

1. INTRODUCTION

Diabetes mellitus is one of the non-communicable diseases that poses a serious threat to global public health [1], [2], [3]. Based on data from the International Diabetes Federation (IDF), it is estimated that more than 500 million people worldwide are living with diabetes, and this number continues to increase every year [4], [5]. This disease not only causes a large economic burden on the health system but also has a significant impact on the quality of life of sufferers [6], [7]. As the number of cases increases, the need for a fast and accurate early detection system becomes increasingly important. In recent years, the use of artificial intelligence (AI) and machine learning (ML) technology in the medical field has shown great potential in improving the quality of disease diagnosis and prediction [8], [9], [10]. However, many studies still face challenges in terms of the accuracy and generalization of predictive models to new data, which are largely due to the suboptimal training and model configuration process [11]. This is the basis for the need to explore strategies to improve the performance of diabetes prediction models through a more systematic and adaptive technical approach.

Although various artificial intelligence and machine learning algorithms have been proven to provide effective predictive solutions for diabetes diagnosis, the performance of the resulting models is

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

often inconsistent when applied to varying real data [12], [13]. Models such as logistic regression, support vector machines, and decision trees have shown promising levels of accuracy, but their optimal performance is greatly influenced by the internal parameters used during training [14], [15], [16]. In this context, the process of optimizing the model through hyperparameter tuning techniques is a crucial step to improve the accuracy, sensitivity, and specificity of predictions [17], [18], [19]. By systematically and data-driven parameter tuning, the model can better adapt to the complexity of health data and improve its ability to identify patterns that indicate diabetes risk [20], [21]. Therefore, this study proposes an approach to improve diabetes prediction performance by utilizing advanced tuning techniques that are more adaptive and efficient, as a form of improvement in the application of existing machine learning methods.

Ismail et al. (2020) [22] used several classification algorithms such as Logistic Regression, k-Nearest Neighbors (k-NN), Decision Tree, and Support Vector Machine (SVM) on the Pima Indians Diabetes dataset. The results showed that SVM provided the highest accuracy of 78.9%, compared to other algorithms. However, this study still uses the default parameter approach without in-depth tuning, so the results obtained do not necessarily reflect the maximum performance of each model. The main weakness identified is the absence of further evaluation of the sensitivity and specificity of the model, which is important in the context of medical diagnosis.

Ahamed et al. (2022) [23] evaluated the performance of two ensemble algorithms: Random Forest and Gradient Boosting. This study showed that Random Forest produced an accuracy of 82.4%, while Gradient Boosting achieved 84.1%. This study indicates that the ensemble method has advantages in managing the complexity of medical data. However, its weakness lies in the hyperparameter selection process which still uses the conventional grid search approach, which tends to be time-consuming and inefficient in large parameter search spaces.

Sudhakar et al. (2025) [24] applied an advanced tuning method based on Bayesian Optimization on the XGBoost model. The experimental results showed an increase in accuracy from 80.2% to 86.7% after tuning, with a significant increase in F1-Score and ROC-AUC values. This study shows that a systematic tuning approach can significantly improve model performance. However, the limitations of this study are the dataset coverage which is still limited to one source (Pima dataset) and the lack of cross-validation on external data to ensure model generalization.

Table 1. Review of Related Works on Diabetes Prediction

Researcher	Methods Novelty	Drawback
Ismail et al. (2020) [22]	SVM, KNN, DT, LR (comparative study of classical ML models)	No hyperparameter tuning; lacks emphasis on medically critical recal
Ahamed et al. (2022) [23]	Random Forest & Gradient Boosting (evaluation of classical ensemble methods)	Used only grid search for tuning; inefficient for large parameter space
Sudhakar et al. (2025) [24]	XGBoost with Bayesian Optimization (advanced hyperparameter tuning)	Limited dataset (Pima); lacks validation on external datasets

Based on various studies above, researcher have shown that machine learning algorithms are able to provide quite good results in diabetes prediction, the performance of the resulting model is still not optimal due to the less than optimal hyperparameter tuning process. Several previously used approaches tend to rely on default configurations or inefficient conventional tuning methods, thus limiting the maximum potential of the predictive model. This study proposes a solution to this problem by applying a more adaptive and systematic advanced hyperparameter tuning technique to improve model

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

performance. The main focus is directed at optimizing tree ensemble-based classification algorithms, especially Random Forest, to obtain more accurate, stable, and reliable prediction results in the context of diabetes diagnosis. Thus, this study is expected to provide a significant contribution in overcoming the limitations of previous models and encouraging the application of more precise machine learning in the health sector. This study aims to improve diabetes prediction accuracy by applying systematic hyperparameter tuning to the Random Forest model combined with comprehensive preprocessing and feature selection.

2. METHOD

Based on Figure 1, the proposed method flow in this study starts from Raw Data obtained from the diabetes prediction dataset. In the pre-processing stage, several important steps are taken to ensure data quality, namely: Handling Duplicate Data to remove repeated entries, Handling Unique Data to check and eliminate data with irrelevant identical values, and Handling Missing Value to handle empty or missing values with an appropriate approach (such as mean or mode imputation). Furthermore, in the data preparation stage, Handling Imbalance Class is carried out using techniques such as SMOTE to balance the class distribution, determining the hyperparameter tuning scheme to optimize the search for the best parameters, and determining Random Forest parameters such as the number of trees and maximum depth. In addition, feature selection is also applied to select the features that contribute most to the prediction. In the final stage, namely evaluation and model formation, the data is divided into two parts: 80% for training and 20% for validation, then the model building process is carried out using the Random Forest algorithm. Model performance is evaluated using the Confusion Matrix to calculate metrics such as accuracy, precision, recall, and F1-score as a benchmark for classification success.

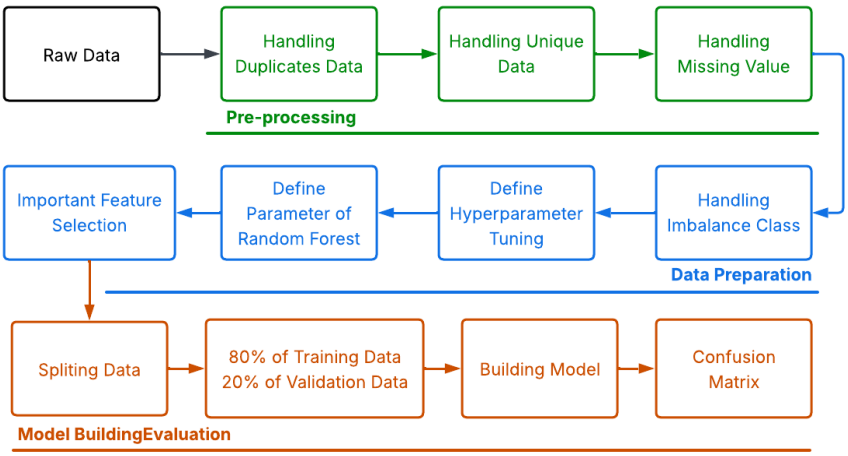


Figure 1. Proposed Methods

2.1. Data Pre-processing

Data preprocessing is a primary early step within any machine learning pipeline that serves to clean, organize, and prepare raw data for facilitating better prediction accuracy and reliability within a model. Within diabetic prediction, preprocessing includes duplicate entry identification and removal using duplicate data handling so that no individual record is over-weighted. Preprocessing also covers analysis and correction of features that feature skewed unique value distributions, which could point to irrelevant or inconsistent data. Missing value handling is also necessary to avoid biased learning or the occurrence of previously unseen defects within the model. By using these preprocessing methods, the

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

dataset becomes better ordered, consistent, and apt for accurate and reliable machine learning model training.

2.1.1. Handling Duplicate Data

The first step in the pre-processing stage is to identify and remove duplicate data that can affect the accuracy of the prediction model. In the dataset used, there were 3,854 duplicate entries that had identical values across columns. The presence of this repeated data has the potential to cause bias in the model training process, because it gives excessive weight to the same pattern. Therefore, all duplicate rows were removed to ensure that only unique data is used in model learning.

2.1.2. Handling Unique Data

The next step is to examine the features that have very high unique values or values that do not vary significantly. From the analysis results, it is known that columns such as bmi have 4,247 unique values, while the age feature has 102 unique values, which are still within reasonable limits for numeric features. No columns were found with a single unique value or almost no variation, which generally need to be removed because they do not provide discriminatory information. Thus, no feature removal was carried out at this stage because all features are still statistically relevant for use in model training.

2.1.3. Handling Missing Value

The next step is handling missing values, which is very important to ensure the integrity and consistency of the data used in model training. Based on the analysis results, a small number of missing values were found in several numeric features such as bmi and glucose, each with less than 0.5% missing values from the total data. To overcome this, the mean imputation technique was applied, which is to replace empty values with the average of each related feature. This approach was chosen because the proportion of missing values is very small and the data distribution is relatively normal. Thus, all data can still be utilized optimally without having to sacrifice a large amount of data due to row deletion. The following is the mean imputation equation used to replace missing values in a numeric feature:

$$x_i = \begin{cases} x_i, \ if \ x_i \neq 0 \ (not \ missing) \\ m, if \ x_i = 0 \ (mising) \end{cases} \quad (1)$$

$$m = \frac{1}{n} \sum_{j=1}^{n} x_i \quad (2)$$

Where x_i the value at the ith data point, m is mean (average) of all non-missing values in that feature, and n is total number of non-missing values in the feature.

2.2. Data Preparation

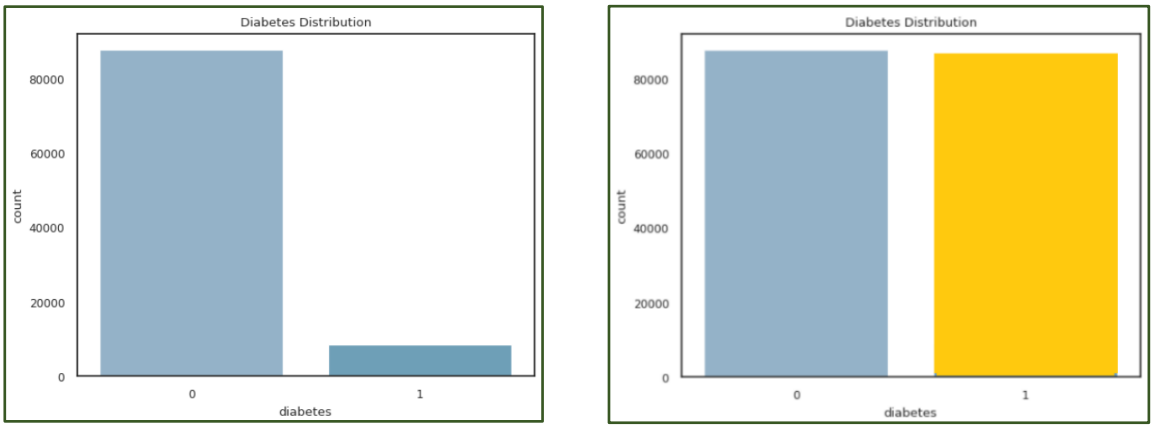
Data preparation is the step involving the converting of pre-processed data to a form that is ready for use within model training. It is one of the most important parts of the machine learning pipeline because it helps optimize the performance and generalization potential of the model. Data preparation for our research includes three major steps: handling imbalanced class distribution, defining the space for hyperparameter tuning, and determining the initial Random Forest classifier parameters.

2.2.1. Handling Imbalance Class

Class imbalance occurs when certain categories in the target variable have significantly more samples than others [25]. This condition may lead to biased learning where the model performs well only on the majority of classes. To address this, resampling techniques such as SMOTE (Synthetic Minority Oversampling Technique) are employed to synthetically generate new instances of minority

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

classes, ensuring a more balanced class distribution and reducing prediction bias [26]. The SMOTE processing results can be seen in Figure 2.



(a) Before SMOTE (b) After SMOTE

Figure 2. Oversampling using SMOTE

Based on the visualization in Figure 2, it can be seen that the condition before the SMOTE process (Figure 2a) shows a very unbalanced class distribution, where one or several classes dominate the number of samples significantly compared to other classes. This imbalance has the potential to cause the model to focus too much on the majority class and ignore the minority class. However, after oversampling with the SMOTE technique (Figure 2b), the distribution between classes becomes more balanced.

2.2.2. Define Hyperparameter Tuning

Hyperparameter tuning is a fundamental step in optimizing machine learning model performance [27], [28]. In this study, a Grid Search approach is employed to systematically explore combinations of parameters in the Random Forest classifier. The grid search is paired with 5-fold cross-validation, ensuring robustness and generalizability across different data splits. Each combination of parameters is evaluated, and the one yielding the highest cross-validated score is selected for final model training. The hyperparameters explored include the number of trees in the forest (n_estimators), the maximum tree depth (max_depth), the minimum number of samples required to split an internal node (min_samples_split), and the minimum number of samples required to be at a leaf node (min_samples_leaf) [29].

Table 2. Hyperparameter Tuning

Hyperparameter	Description	Function	Value Tested
Number of Estimators	Number of decision trees in the forest	$n\_estimators$	50, 100, 200
Maximum Tree Depth	Maximum depth of each tree	max_depth	None, 10, 20
Minimum Samples Split	Minimum samples required to split an internal node	$min\_samples\_split$	2, 5, 10
Minimum Samples Leaf	Minimum samples required to be at a leaf node	$min\_samples\_leaf$	1, 2, 4

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

2.3. Random Forest Classifier

Random Forest is an ensemble learning-based machine learning algorithm that combines several decision trees to produce more accurate and stable predictions [30], [31]. This method works by building many decision trees during the training process and combining the results of each tree to determine the output class based on the principle of voting (for classification) or average (for regression) [32], [33].

The main advantage of Random Forest is its ability to handle high-dimensional data and stability against overfitting, because its ensemble approach helps reduce model variance. In addition, this algorithm is able to handle correlated features and provides an estimate of the importance of each feature through the calculation of feature importance [34]. In this study, Random Forest is used as the main algorithm to predict diabetes incidence based on a number of health features. To achieve optimal prediction performance, a tuning process is carried out on several important parameters such as the number of estimators (trees), the maximum depth of the tree, and the minimum number of samples for splits and leaves.

2.4. Important Feature Selection

Feature selection is an important step in machine learning modeling because it can improve the efficiency, accuracy, and interpretability of the model [35], [36]. By selecting features that are truly relevant to the prediction target, the model can reduce complexity, speed up the training process, and avoid the risk of overfitting [37], [38]. In this study, the feature importance estimation method of the Random Forest algorithm was used to determine the relative contribution of each feature to model performance. Complete information regarding the level of importance of each feature can be seen in Table 3.

Based on the evaluation results, the HbA1c level and blood glucose level features showed the most dominant contribution to diabetes prediction, at 44% and 31%, respectively. Other features such as age and BMI contributed moderately, while variables such as smoking history and gender had an importance value close to zero, meaning they had minimal influence on classification decisions. The visualization can be seen in Figure 3.

2.5. Confusion Matrix Evaluation

Confusion matrix is an evaluation tool used to assess the performance of a classification model based on the comparison between the predicted results and the actual labels [39], [40]. This matrix displays information in the form of a two-dimensional table that shows the number of correct and incorrect predictions for each target class, making it easier to calculate various evaluation metrics. From these values, a number of evaluation metrics can be calculated to determine how well the model works, including:

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ Data} \quad (3)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (4)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (5)$$

$$F1\text{-}score = \frac{2 \times Precision \times Recall}{Precision \times Recall} \quad (6)$$

Where, True Positive (TP) is positive cases that are correctly predicted as positive, False Positive (FP) is negative cases that are incorrectly predicted as positive, True Negative (TN) is negative cases

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

that are correctly predicted as negative, False Negative (FN) is positive cases that are incorrectly predicted as negative.

Table 3. Feature Selection

Selected Feature	Importance
HbA1c level	0.44
Blood glucose level	0.31
Age	0.13
BMI	0.06
Hypertension	0.03
Heart disease	0.02
Smoking history (past smoker)	0.00
Smoking history (non-smoker)	0.00
Gender (Male)	0.00
Gender (Female)	0.00
Smoking history (current)	0.00

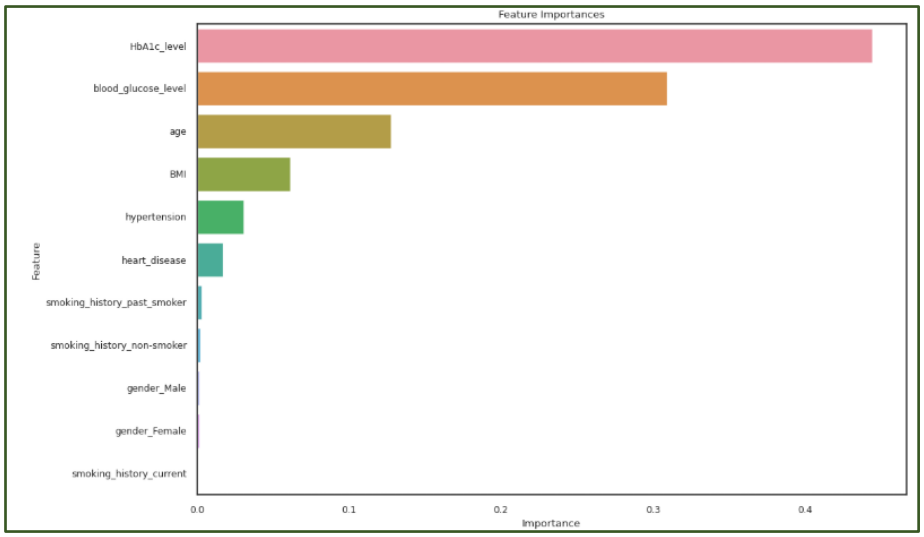


Figure 3. Feature Selection

3. RESULT

In this study, the model was implemented using Python, with the dataset split into training and validation sets in an 80:20 ratio. The 80% of the data was used for training the model, while the remaining 20% was reserved for validation. This split ensures that the model is trained on a substantial portion of the data, while also being evaluated on unseen data to test its generalizability and accuracy.

Before proceeding to model optimization, a correlation analysis was conducted to understand the statistical relationships between features and to identify which variables are most associated with the target class, namely diabetes. This analysis helps in detecting multicollinearity and selecting the most informative features for the classification task. The first correlation heatmap in Figure 4 presents the pairwise correlation matrix among all variables in the dataset

It visually highlights strong positive or negative correlations through a color gradient, which assists in identifying redundancy or relationships that may affect the model. Furthermore, to focus on the relationship between each feature and the diabetes target variable, a second heatmap in Figure 5 was generated. This plot shows the sorted correlation values of all independent features with respect to the

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

diabetes outcome. From this visualization, it is evident which features are more strongly correlated-either positively or negatively-with the presence of diabetes. These insights play a vital role in understanding the dataset structure and support further steps in feature selection and model development.

Next, for hyperparameter tuning, the GridSearchCV method was applied to find the best combination of parameters for the Random Forest classifier. The results of the grid search were converted into a DataFrame for better visualization. The plot generated (using seaborn) shows how the mean test score varies with the number of estimators and max depth. This visualization provides insights into the optimal values for these hyperparameters, helping to refine the model's performance. The tuning process, as seen in Figure 6, allows for a more informed decision regarding the configuration of the model to achieve the best results.

Following the hyperparameter tuning process illustrated in Figure 6, the model was trained using the optimal parameter configuration obtained from the GridSearchCV results. The performance of the model on the training dataset is summarized in Table 4, which includes evaluation metrics such as accuracy, precision, recall, and F1-score. These metrics provide a comprehensive assessment of the model's ability to correctly classify diabetic and non-diabetic cases. Meanwhile, the evaluation on the testing dataset is visualized through the confusion matrix shown in Figure 7.

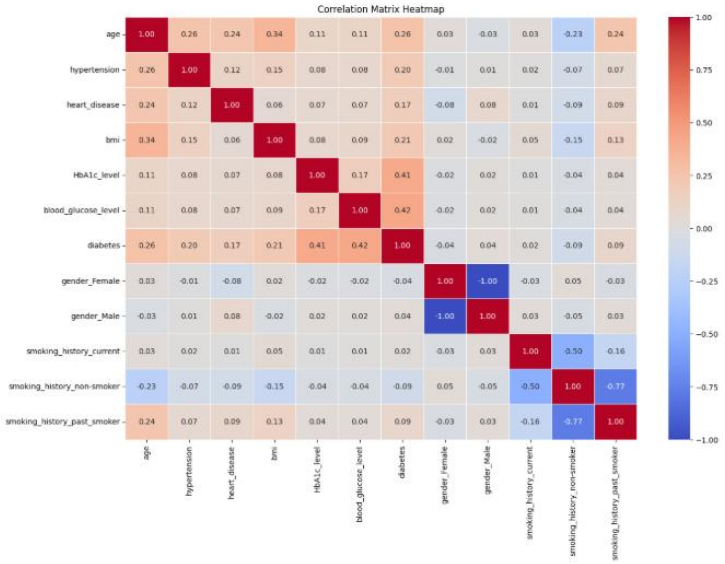


Figure 4. Heatmap Showing Pairwise Correlations Among All Features

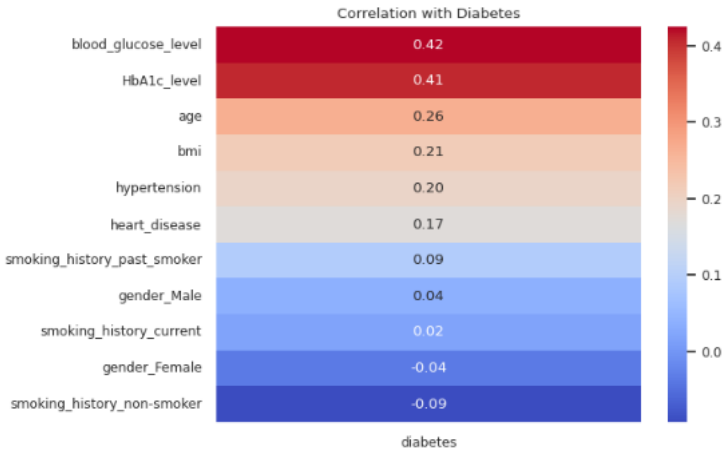


Figure 5. Correlation Between Each Independent Feature and The Diabetes Outcome

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

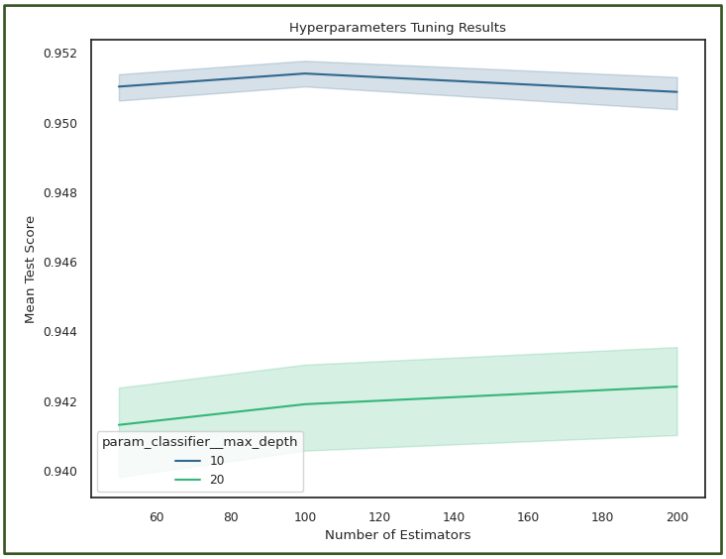


Figure 6. Hyperparameters Tuning Results

Table 4. Performance Model Evaluation of the Optimized Random Forest Model

Class	Accuracy	Precision	Recall	F1-Score
Normal (0)		0.98	0.96	0.97
Diabetes (1)	0.95	0.68	0.80	0.74
Average		0.83	0.88	0.85

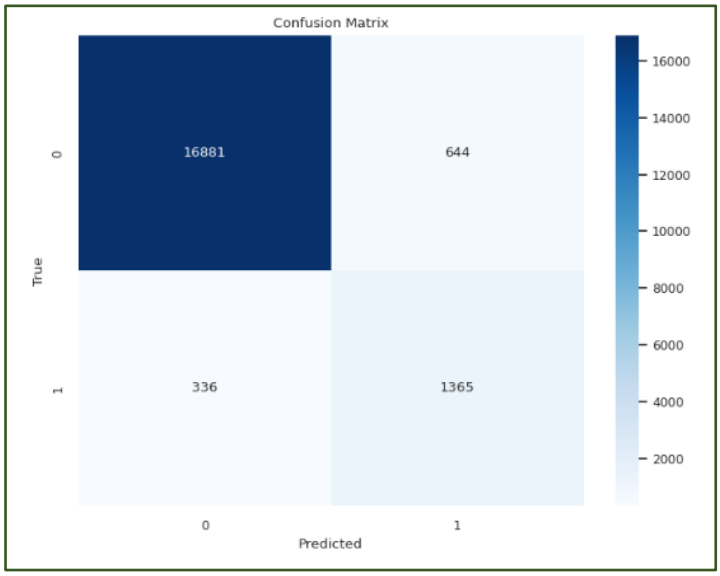


Figure 7. Confusion Matrix of the Final Model on the Test Dataset Showing Prediction Distribution Between Diabetic and Non-Diabetic Classes

The performance evaluation of the diabetes prediction model, as shown in Table 4 and the confusion matrix, indicates that the Random Forest Classifier performed effectively in classifying both diabetic and non-diabetic individuals. The model was optimized using GridSearchCV to tune several hyperparameters such as the number of estimators, tree depth, and minimum samples for split and leaf nodes. Prior to training, the dataset underwent thorough pre-processing steps including duplicate and

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

unique value handling, along with imputing missing data using the mean. During data preparation, class imbalance was addressed using SMOTE, which significantly balanced the representation between diabetic and non-diabetic cases, followed by the definition of hyperparameter tuning space and selection of important features to ensure model generalizability.

From the confusion matrix, it can be observed that the model correctly identified 16,881 TN and 1,365 tTP, with 644 FP and 336 FN. These results reflect the model's capability to distinguish diabetic conditions with high overall accuracy (94.9%) and robustness. The high number of TNs suggests that the model is especially reliable in ruling out non-diabetic cases, while the satisfactory TP rate indicates a strong ability to identify actual diabetic instances despite initial class imbalance. The enhancement process-including SMOTE and hyperparameter optimization-proved crucial in reducing misclassification and elevating performance across all key evaluation metrics.

4. DISCUSSIONS

Table 5. Comparison with Related Studies

Researcher	Novelty	Accuracy	Precision	Recall	F1-Score
Ismail et al. (2020) [22]	Support Vector Machine	0.79	0.76	0.74	0.75
Ahamed et al. (2022) [23]	Random Forest	0.84	0.82	0.83	0.83
Sudhakar et al. (2025) [24]	XGBoost with Bayesian Optimization	0.86	0.85	0.88	0.86
Our	Random Forest with Hyperparameter Tuning	0.95	0.83	0.88	0.85

Based on the comparison in Table 5, the proposed method shows significant improvement over existing approaches across all evaluation metrics. Previous studies that utilized models such as Support Vector Machine, standard Random Forest, and even XGBoost with optimization techniques yielded accuracies ranging from 0.79 to 0.86. In contrast, the model in this study achieved a much higher accuracy of 0.95, along with strong precision, recall, and F1-score values. Most notably, the recall score of 0.88 highlights the model's effectiveness in correctly identifying diabetic cases, which is crucial in medical prediction tasks to minimize false negatives.

The improvement achieved in this study is the result of a well-structured machine learning pipeline. This includes thorough preprocessing such as handling duplicates, unique entries, and missing values, followed by balancing the dataset using SMOTE to address class imbalance. The model also benefits from hyperparameter tuning with GridSearchCV and careful feature selection based on importance values. These steps collectively contribute to a more accurate, generalizable, and reliable predictive system for diabetes classification, surpassing the limitations identified in earlier approaches.

In addition to the technical improvements, this study contributes to the field of informatics by supporting the development of intelligent health analytics systems through structured and scalable machine learning pipelines. The integration of preprocessing, feature selection, and hyperparameter tuning offers a reusable framework that can be adapted for other medical prediction tasks, ensuring both interpretability and computational efficiency.

Despite the significant improvements, this study has limitations. The dataset used was obtained from a single source, which may restrict the diversity and representativeness of the input data. As a result, while the model shows strong performance within this context, its generalization to broader populations or other clinical datasets may vary. Future evaluations using multi-center, multi-ethnic

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

datasets are necessary to validate its robustness and extend applicability across diverse healthcare environments.

5. CONCLUSION

This study successfully improved the performance of diabetes prediction through the systematic application of hyperparameter tuning techniques to the Random Forest algorithm, with the uniqueness of the method in the form of a combination of deep data processing (handling duplicates, missing values, and imbalanced classes with SMOTE) and exploration of optimal parameters using GridSearchCV and 5-fold cross-validation. The experimental results showed a significant improvement compared to previous studies, with an accuracy of 95%, and F1-score and recall of 0.85 and 0.88, respectively, indicating the model's ability to reliably detect diabetes cases. The main contribution of this study is the implementation of a structured and adaptive machine learning pipeline, which can be used as an early decision support system in the healthcare sector. This work sets a foundation for developing clinically applicable AI-based early detection tools for diabetes, enabling smarter, scalable, and more timely interventions. For further research, testing on larger and more diverse clinical datasets from various geographic populations, as well as exploration of tuning methods based on Bayesian Optimization or Genetic Algorithm, are recommended to improve model generalization and tuning process efficiency.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

REFERENCES

[1] Y. Wang and D. J. Magliano, "Special Issue: 'New Trends in Diabetes, Hypertension, and Cardiovascular Diseases,'" Mar. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/ijms25052711.

[2] M. D. Butt *et al.*, "A systematic review of the economic burden of diabetes mellitus: contrasting perspectives from high and low middle-income countries," *J Pharm Policy Pract*, vol. 17, no. 1, Dec. 2024, doi: 10.1080/20523211.2024.2322107.

[3] W. Bielka, A. Przezak, P. Molęda, E. Pius-Sadowska, and B. Machaliński, "Double diabetes—when type 1 diabetes meets type 2 diabetes: definition, pathogenesis and recognition," *Cardiovasc Diabetol*, vol. 23, no. 1, p. 62, Feb. 2024, doi: 10.1186/s12933-024-02145-x.

[4] I. Hernar *et al.*, "Diabetes Distress and Associations With Demographic and Clinical Variables: A Nationwide Population-Based Registry Study of 10,186 Adults With Type 1 Diabetes in Norway," *Diabetes Care*, vol. 47, no. 1, pp. 126–131, Jan. 2024, doi: 10.2337/dc23-1001.

[5] N. Hermanns, B. Kulzer, and D. Ehrmann, "Person-reported outcomes in diabetes care: What are they and why are they so important?," *Diabetes Obes Metab*, vol. 26, no. S1, pp. 30–45, Mar. 2024, doi: 10.1111/dom.15471.

[6] I. P. Kamila, C. A. Sari, E. H. Rachmawanto, and N. R. D. Cahyo, "A Good Evaluation Based on Confusion Matrix for Lung Diseases Classification using Convolutional Neural Networks," *Advance Sustainable Science, Engineering and Technology*, vol. 6, no. 1, p. 0240102, Dec. 2023, doi: 10.26877/asset.v6i1.17330.

[7] R. J. Porter, M. J. Arends, A. M. D. Churchhouse, and S. Din, "Inflammatory Bowel Disease-Associated Colorectal Cancer: Translational Risks from Mechanisms to Medicines," *J Crohns Colitis*, vol. 15, no. 12, pp. 2131–2141, Dec. 2021, doi: 10.1093/ecco-jcc/jjab102.

[8] N. R. D. Cahyo and M. M. I. Al-Ghiffary, "An Image Processing Study: Image Enhancement, Image Segmentation, and Image Classification using Milkfish Freshness Images," *IJECAR) International Journal of Engineering Computing Advanced Research*, vol. 1, no. 1, pp. 11–22, 2024.

[9] F. Farhan, C. A. Sari, E. H. Rachmawanto, and N. R. D. Cahyo, "Mangrove Tree Species Classification Based on Leaf, Stem, and Seed Characteristics Using Convolutional Neural

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

Networks with K-Folds Cross Validation Optimalization," *Advance Sustainable Science Engineering and Technology*, vol. 5, no. 3, p. 02303011, Oct. 2023, doi: 10.26877/asset.v5i3.17188.

[10] M. M. I. Al-Ghiffary, N. R. D. Cahyo, E. H. Rachmawanto, C. Irawan, and N. Hendriyanto, "Adaptive deep learning based on FaceNet convolutional neural network for facial expression recognition," *Journal of Soft Computing*, vol. 05, no. 03, pp. 271–280, 2024, doi: https://doi.org/10.52465/joscex.v5i3.450.

[11] M. M. I. Al-Ghiffary, C. A. Sari, E. H. Rachmawanto, N. M. Yacoob, N. R. D. Cahyo, and R. R. Ali, "Milkfish Freshness Classification Using Convolutional Neural Networks Based on Resnet50 Architecture," *Advance Sustainable Science Engineering and Technology*, vol. 5, no. 3, p. 0230304, Oct. 2023, doi: 10.26877/asset.v5i3.17017.

[12] Z. Salahuddin, H. C. Woodruff, A. Chatterjee, and P. Lambin, "Transparency of deep neural networks for medical image analysis: A review of interpretability methods," Jan. 01, 2022, *Elsevier Ltd*. doi: 10.1016/j.compbiomed.2021.105111.

[13] M. Xiao, L. Zhang, W. Shi, J. Liu, W. He, and Z. Jiang, "A visualization method based on the Grad-CAM for medical image segmentation model," in *2021 International Conference on Electronic Information Engineering and Computer Science (EIECS)*, IEEE, Sep. 2021, pp. 242–247. doi: 10.1109/EIECS53707.2021.9587953.

[14] A. D. Amirruddin, F. M. Muharam, M. H. Ismail, N. P. Tan, and M. F. Ismail, "Synthetic Minority Over-sampling TEchnique (SMOTE) and Logistic Model Tree (LMT)-Adaptive Boosting algorithms for classifying imbalanced datasets of nutrient and chlorophyll sufficiency levels of oil palm (Elaeis guineensis) using spectroradiometers and unmanned aerial vehicles," *Comput Electron Agric*, vol. 193, p. 106646, Feb. 2022, doi: 10.1016/j.compag.2021.106646.

[15] N. R. D. Cahyo, C. A. Sari, E. H. Rachmawanto, C. Jatmoko, R. R. A. Al-Jawry, and M. A. Alkhafaji, "A Comparison of Multi Class Support Vector Machine vs Deep Convolutional Neural Network for Brain Tumor Classification," in *2023 International Seminar on Application for Technology of Information and Communication (iSemantic)*, IEEE, Sep. 2023, pp. 358–363. doi: 10.1109/iSemantic59612.2023.10295336.

[16] A. J. Albert, R. Murugan, and T. Sripriya, "Diagnosis of heart disease using oversampling methods and decision tree classifier in cardiology," *Research on Biomedical Engineering*, vol. 39, no. 1, pp. 99–113, Dec. 2022, doi: 10.1007/s42600-022-00253-9.

[17] X. Liu, M. Pedersen, and R. Wang, "Survey of natural image enhancement techniques: Classification, evaluation, challenges, and perspectives," *Digit Signal Process*, vol. 127, p. 103547, 2022, doi: https://doi.org/10.1016/j.dsp.2022.103547.

[18] M. J. Lakshmi and S. Nagaraja Rao, "Brain tumor magnetic resonance image classification: a deep learning approach," *Soft comput*, vol. 26, no. 13, pp. 6245–6253, Jul. 2022, doi: 10.1007/s00500-022-07163-z.

[19] Z. He, "Deep Learning in Image Classification: A Survey Report," in *Proceedings - 2020 2nd International Conference on Information Technology and Computer Application, ITCA 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 174–177. doi: 10.1109/ITCA52113.2020.00043.

[20] W. Yu, C. Y. Wong, R. Chavez, and M. A. Jacobs, "Integrating big data analytics into supply chain finance: The roles of information processing and data-driven culture," *Int J Prod Econ*, vol. 236, p. 108135, Jun. 2021, doi: 10.1016/j.ijpe.2021.108135.

[21] A. V. Poznyak, L. Litvinova, P. Poggio, V. N. Sukhorukov, and A. N. Orekhov, "Effect of Glucose Levels on Cardiovascular Risk," *Cells*, vol. 11, no. 19, p. 3034, Sep. 2022, doi: 10.3390/cells11193034.

[22] L. Ismail and H. Materwala, "Comparative Analysis of Machine Learning Models for Diabetes Mellitus Type 2 Prediction," in *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, IEEE, Dec. 2020, pp. 527–533. doi: 10.1109/CSCI51800.2020.00095.

[23] B. S. Ahamed, M. S. Arya, and A. O. V. Nancy, "Diabetes Mellitus Disease Prediction Using Machine Learning Classifiers with Oversampling and Feature Augmentation," *Advances in Human-Computer Interaction*, vol. 2022, pp. 1–14, Sep. 2022, doi: 10.1155/2022/9220560.

[24] A. Sudhakar, S. S, S. M., S. A, B. Subramanian, and V. Ramana K, "Bayesian Optimization for Hyperparameter Tuning in Healthcare for Diabetes Prediction," *Informing Science: The International Journal of an Emerging Transdiscipline*, vol. 28, p. 008, 2025, doi: 10.28945/5445.

[25] U. Hasanah, A. M. Soleh, and K. Sadik, "Effect of Random Under sampling, Oversampling, and SMOTE on the Performance of Cardiovascular Disease Prediction Models," *Jurnal Matematika, Statistika dan Komputasi*, vol. 21, no. 1, pp. 88–102, Sep. 2024, doi: 10.20956/j.v21i1.35552.

[26] M. F. Muzakki, R. D. Prayogo, and M. A. Rizky A, "Handling Imbalanced Data for Acute Coronary Syndrome Classification Based on Ensemble and K-Means SMOTE Method," *JOIV : International Journal on Informatics Visualization*, vol. 7, no. 3–2, p. 1989, Nov. 2023, doi: 10.30630/joiv.7.3-2.1429.

[27] A. H. Victoria and G. Maragatham, "Automatic tuning of hyperparameters using Bayesian optimization," *Evolving Systems*, vol. 12, no. 1, pp. 217–223, Mar. 2021, doi: 10.1007/s12530-020-09345-2.

[28] S. Rahman, M. Ramli, F. Arnia, R. Muharar, and A. Sembiring, "Performance analysis of mAlexnet by training option and activation function tuning on parking images," *IOP Conf Ser Mater Sci Eng*, vol. 1087, no. 1, p. 012084, Feb. 2021, doi: 10.1088/1757-899x/1087/1/012084.

[29] B. Arjmand *et al.*, "Machine Learning: A New Prospect in Multi-Omics Data Analysis of Cancer," Jan. 27, 2022, *Frontiers Media S.A.* doi: 10.3389/fgene.2022.824451.

[30] M. A. Rasyidi, T. Bariyah, Y. I. Riskajaya, and A. D. Septyani, "Classification of handwritten javanese script using random forest algorithm," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 3, pp. 1308–1315, Jun. 2021, doi: 10.11591/eei.v10i3.3036.

[31] E. H. Rachmawanto, D. R. I. M. Setiadi, N. Rijati, A. Susanto, I. U. W. Mulyono, and H. Rahmalan, "Attribute Selection Analysis for the Random Forest Classification in Unbalanced Diabetes Dataset," in *2021 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 2021, pp. 82–86. doi: 10.1109/iSemantic52711.2021.9573181.

[32] M. Daviran, M. Shamekhi, R. Ghezelbash, and A. Maghsoudi, "Landslide susceptibility prediction using artificial neural networks, SVMs and random forest: hyperparameters tuning by genetic optimization algorithm," *International Journal of Environmental Science and Technology*, vol. 20, no. 1, pp. 259–276, Jan. 2023, doi: 10.1007/s13762-022-04491-3.

[33] C. Umam, L. B. Handoko, and F. O. Isinkaye, "Performance Analysis of Support Vector Classification and Random Forest in Phishing Email Classification," *Scientific Journal of Informatics*, vol. 11, no. 2, pp. 367–374, May 2024, doi: 10.15294/sji.v11i2.3301.

[34] M. P. K. Dewi and E. B. Setiawan, "Feature Expansion Using Word2vec for Hate Speech Detection on Indonesian Twitter with Classification Using SVM and Random Forest," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 2, p. 979, Apr. 2022, doi: 10.30865/mib.v6i2.3855.

[35] O. Somantri, R. H. Maharrani, and S. Purwaningrum, "Coastal Sentiment Review Using Naïve Bayes with Feature Selection Genetic Algorithm," *Scientific Journal of Informatics*, vol. 10, no. 3, pp. 229–238, Jun. 2023, doi: 10.15294/sji.v10i3.43988.

[36] Priyanka and D. Kumar, "Feature Extraction and Selection of kidney Ultrasound Images Using GLCM and PCA," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 1722–1731. doi: 10.1016/j.procs.2020.03.382.

[37] F. Z. BOUKHOBZA, A. HACINE GHARBI, and K. ROUABAH, "A New Facial Expression Recognition Algorithm Based on DWT Feature Extraction and Selection," *The International Arab Journal of Information Technology*, vol. 21, no. 4, 2024, doi: 10.34028/iajit/21/4/6.

[38] N. A. Samee, G. Atteia, S. Meshoul, M. A. Al-antari, and Y. M. Kadah, "Deep Learning Cascaded Feature Selection Framework for Breast Cancer Classification: Hybrid CNN with Univariate-Based Approach," *Mathematics*, vol. 10, no. 19, Oct. 2022, doi: 10.3390/math10193631.

[39] B. Zhang, X. Chen, X. Cui, and M. Shen, "A Novel Bias-Adjusted Estimator Based on Synthetic Confusion Matrix (BAESCM) for Subregion Area Estimation," *Remote Sens (Basel)*, vol. 17, no. 7, p. 1145, Mar. 2025, doi: 10.3390/rs17071145.

Vol. 6, No. 4, August 2025, Page. 1847-1860
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4755

[40] C.-L. Fan, "Evaluation Model for Crack Detection with Deep Learning: Improved Confusion Matrix Based on Linear Features," *J Constr Eng Manag*, vol. 151, no. 3, Mar. 2025, doi: 10.1061/JCEMD4.COENG-14976.