Improving Chronic Kidney Disease Prediction Performance Using AdaBoost Optimization: A Comparative Machine Learning Study

Sarwo Derta[1*], Solikhun[2], Putrama Alkhairi[3]

[1]Informatics, Science and Technology, UIN Sjech M. Djamil Djambek Bukittinggi, West Sumatra, Indonesia
[2]Informatics Engineering, STIKOM Tunas Bangsa, Pematangsiantar, North Sumatra, Indonesia
[3]Informatics Management, STIKOM Tunas Bangsa, Pematangsiantar, North Sumatra, Indonesia
E-mail: [1*]sarwoderta@uinbukittinggi.ac.id, [2]solikhun@amiktunasbangsa.ac.id, [3]putrama@amiktunasbangsa.ac.id

(Received: 2 Sep 2026, revised: 5 Oct 2026, 8 Oct 2026, accepted: 9 Oct 2026)

Abstract

Chronic Kidney Disease (CKD) is a major global health challenge characterized by progressive renal dysfunction, high mortality, and increasing prevalence. Early detection is essential to prevent disease progression and improve patient outcomes; however, accurate prediction remains challenging because CKD is associated with complex interactions among multiple clinical indicators. This study proposes an optimized Machine Learning framework based on AdaBoost with GridSearchCV hyperparameter optimization to improve CKD prediction performance. A comparative experimental approach was conducted using a clinical CKD dataset comprising 1,000 patient records and 22 clinical attributes. Data preprocessing involved data cleaning, missing value handling, categorical feature encoding, feature scaling, exploratory data analysis, and feature selection. The proposed method was compared with Logistic Regression, Decision Tree, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Random Forest. GridSearchCV combined with 10-fold Stratified Cross-Validation was employed to identify the optimal AdaBoost configuration and improve model generalization. Model performance was evaluated using Accuracy, Precision, Recall, F1-Score, Confusion Matrix, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). Experimental results demonstrated that the optimized AdaBoost model achieved the best performance, obtaining 100% Accuracy, Precision, Recall, and F1-Score, with an AUC of 1.000, outperforming all baseline models. Permutation Feature Importance further identified serum creatinine, anemia, diabetes mellitus, hypertension, blood urea, and random blood glucose as the most influential predictors of CKD. These findings demonstrate that systematic hyperparameter optimization can enhance the predictive performance and interpretability of AdaBoost, supporting its potential as a clinical decision-support tool for early CKD detection using routinely available clinical patient data.

Keywords: Chronic Kidney Disease, AdaBoost, GridSearchCV, Machine Learning, Hyperparameter Optimization, Classification, Clinical Decision Support.

I. INTRODUCTION

Chronic Kidney Disease (CKD) is a global health problem that continues to increase annually. This disease is characterized by a gradual decline in kidney function that occurs over a long period and can progress to terminal kidney failure if not detected and treated early [1]–[3]. According to various international health reports, CKD is among the ten leading causes of death worldwide and significantly impacts patients' quality of life and the economic burden on healthcare systems [4], [5]. The high incidence of CKD presents a serious challenge because in its early stages the disease often shows no specific symptoms, resulting in many cases being identified only after kidney function has already suffered significant damage [1], [6].

Advances in information technology and artificial intelligence have opened up new opportunities to support the diagnosis and prediction of various diseases, including chronic kidney disease [7], [8]. One widely used approach is machine learning, a computational technique that enables systems to learn patterns from historical data to generate accurate predictions. In the context of CKD, machine learning can analyze various clinical factors such as age, blood pressure, creatinine levels, blood urea nitrogen (BUN), glomerular filtration rate (GFR), haemoglobin, proteinuria, diabetes, hypertension, and various other medical indicators related to a patient's kidney condition. The use of Machine Learning is expected to assist medical personnel in early detection and minimize the risk of complications [9], [10].

Although various Machine Learning algorithms have been widely applied to CKD prediction, a major issue that remains frequently encountered is the difference in performance between algorithms in identifying patients at risk of developing chronic kidney disease [11]–[13]. Several previous studies have shown that conventional algorithms, such as Logistic Regression, have a fairly good ability to handle linear relationships between features, but often experience limitations when faced with complex and non-linear data patterns [14], [15]. The K-Nearest Neighbor (KNN) algorithm offers simple classification capabilities [16]–[18], but is sensitive to data distribution and the selection of nearest neighbor parameters. Meanwhile, Support Vector Machines (SVM) [19]–[21], Decision Trees [22]–[24], and Random Forests offer better capabilities in handling data complexity, but there are still opportunities to improve model accuracy and stability through ensemble learning approaches [25]–[27].

In previous research conducted by [28] in the article entitled "Advancing Chronic Kidney Disease Prediction: Comparative Analysis of Machine Learning Algorithms and a Hybrid Model" has several advantages that make it relevant in the field of chronic kidney disease (CKD) prediction. This study proposes a Hybrid Model based on a stacking classifier that combines Gaussian Naive Bayes, Naive Bayes, Decision Tree, and Random Forest as a meta-classifier, so that it can utilize the advantages of each algorithm to improve prediction performance. In addition, the study conducted a comparison with popular algorithms such as XGBoost, Random Forest, Logistic Regression, and AdaBoost using various comprehensive evaluation metrics, including Accuracy, Precision, Recall, F1-Score, and AUROC, where the hybrid model achieved the best performance with an accuracy of 94.99% and an AUROC of 95.56%. The use of preprocessing techniques, Pearson Correlation-based feature selection, data visualization, and k-fold cross-validation are also added values because they help improve model quality and reduce the risk of overfitting. However, this article also has several weaknesses. The dataset used is derived only from the UCI Chronic Kidney Disease Dataset with a relatively small data set (400 samples), so the model's generalizability to a broader clinical population is still limited. The explanation of the proposed hybrid architecture is also still lacking in detail, especially regarding the stacking process, parameter tuning, and optimization mechanisms used, resulting in less than optimal reproducibility of the study. Furthermore, the study has not conducted external validation using hospital data or independent datasets, has not compared the model with modern deep learning approaches, and does not present a statistical significance analysis of performance differences between models. Therefore, although this study successfully demonstrates that the hybrid approach can improve the accuracy of CKD prediction compared to a single model, its novel contribution is still at the level of integration with existing models and requires further validation to support implementation in real clinical settings. Meanwhile, the following research conducted by [29] in this article entitled "Comparing Machine Learning Techniques for Detecting Chronic Kidney Disease in Early Stage" has the advantage of a comparative approach that compares several popular machine learning algorithms such as XGBoost, Random Forest, Logistic Regression, AdaBoost, and Hybrid Model to detect Chronic Kidney Disease (CKD) in the early stages using the UCI CKD dataset. This study also applies feature selection using Pearson Correlation, exploratory data visualization, and stacking ensemble that combines Gaussian Naive Bayes, Naive Bayes, Decision Tree, and Random Forest as a meta-classifier to improve accuracy and reduce overfitting. The experimental results show that the hybrid model achieves the best performance with 95% accuracy, 96.2% precision, 95.1% recall, 96% F1-score, and 94.56% AUROC, thus providing evidence that the ensemble approach is able to improve performance compared to a single model. However, this article has several significant weaknesses. The dataset used only consisted of 400 samples with 14 features, thus limiting the model's generalizability to larger and more diverse clinical data sets. Furthermore, the research's novelty is relatively low because the stacking ensemble approach and algorithm used are methods that have been widely applied in previous CKD research. Explanations regarding hyperparameter tuning, model configuration, and computational complexity analysis are also very limited, thus reducing the study's reproducibility. Furthermore, there was no external validation using different hospital or population datasets, no statistical significance testing between models, and no comparison with modern deep learning methods currently widely used in disease prediction. Therefore, although this study successfully demonstrated that a hybrid model can improve CKD prediction performance compared to several conventional machine learning algorithms, its scientific contribution is more about implementing performance improvements than offering truly novel methodological innovations.

One ensemble learning method widely used to improve classification performance is Adaptive Boosting (AdaBoost) [30], [31]. AdaBoost works by iteratively combining several weak learners to produce a more robust and accurate prediction model. AdaBoost's main advantage lies in its ability to assign greater weight to data misclassified in previous iterations, allowing the model to learn more effectively from difficult-to-recognize patterns. Furthermore, AdaBoost's performance can be further improved through a hyperparameter optimization process aimed at finding the best combination of parameters to produce an optimal model [32], [33].

Based on preliminary experiments conducted on the Chronic Kidney Disease dataset, consisting of 1,000 patient data sets and 15 clinical attributes, each machine learning algorithm demonstrated varying levels of performance. The Logistic Regression model achieved 86% accuracy, K-Nearest Neighbor achieved 75.5% accuracy, Support Vector Machine achieved 86.5% accuracy, Random Forest achieved 95% accuracy, and Decision Tree achieved 98% accuracy. These results indicate that although some algorithms are capable of delivering high performance, there is still room for improving predictive capabilities through more effective

DOI: 10.34148/teknika.v15i3.1565
TEKNIKA, Volume 15(3), November 2026, pp. 514-528
ISSN 2549-8037, EISSN 2549-8045

approaches [34]–[36]. To address these issues, this study proposes the use of the AdaBoost algorithm optimized using a Grid Search-based hyperparameter tuning technique. Experimental results show that the best configuration was obtained at a learning rate of 0.1 and 150 estimators, with an average cross-validation accuracy of 99.63%. The optimized model achieved 100% accuracy, precision, recall, and F1-score on the test data. These results demonstrate a significant performance improvement compared to all baseline models used in the study. In addition to focusing on improving prediction accuracy, this research also contributes through a comparative analysis of various machine learning algorithms commonly used in CKD prediction. This comparative analysis is crucial for providing a deeper understanding of the characteristics, advantages, and limitations of each algorithm in handling medical data. Therefore, the research results are expected to serve as a reference for researchers and healthcare practitioners in selecting the most effective prediction model to support the early detection of chronic kidney disease [37]–[39].

Based on this description, this research aimed to improve the performance of chronic kidney disease prediction by optimizing the AdaBoost algorithm and comparing it with several other machine learning algorithms. It is hoped that the results of this research will produce a prediction model with high accuracy, good stability, and the ability to support faster and more precise medical decision-making in the prevention and treatment of chronic kidney disease.

II. METHOD

This research uses an experimental quantitative approach focused on developing a predictive model for chronic kidney disease (CKD) using machine learning algorithms. The experimental approach was chosen because the research aims to evaluate and compare the performance of several classification algorithms in predicting CKD status based on patient clinical characteristics. Furthermore, this study proposes optimizing the AdaBoost algorithm through hyperparameter tuning to obtain a model with better predictive ability than the baseline model. The research process begins with collecting the CKD dataset, followed by a preprocessing stage to improve data quality before use in model training. Once the data is ready, several baseline models are built, consisting of Logistic Regression, K-Nearest Neighbor (KNN), Decision Tree, Random Forest, and Support Vector Machine (SVM). AdaBoost is then optimized using Grid Search Cross-Validation to obtain the best parameter combination. All models are then evaluated using various performance metrics such as Accuracy, Precision, Recall, F1-Score, ROC-AUC, and Cross-Validation Accuracy to obtain a comprehensive overview of the effectiveness of each algorithm

A. Research Dataset

The dataset used in this study was obtained from the Kaggle platform through the publicly accessible notebook "Chronic Kidney Disease Prediction (98% Accuracy)" by Nitesh Yadav. The corresponding Kaggle resource identifies the input data as the Chronic Kidney Disease dataset and provides the dataset as a publicly accessible resource for machine learning-based chronic kidney disease prediction. The dataset was accessed from Kaggle at: https://www.kaggle.com/code/niteshyadav3103/chronic-kidney-disease-prediction-98-accuracy/input.

The dataset was used as a secondary data source and was not collected directly by the authors from patients or healthcare institutions. Therefore, this study does not claim that the data were independently collected from a particular hospital, clinic, or geographic population. The dataset contains clinical and demographic variables relevant to chronic kidney disease classification, with the target variable representing the CKD status. Before model development, the dataset was examined for missing values, categorical representations, variable types, and the distribution of the target classes. Data preprocessing was subsequently performed to prepare the variables for machine learning classification.

This study used the Chronic Kidney Disease dataset, which consists of 1,000 patient records with several clinical attributes related to kidney health. The dataset contains various variables commonly used in CKD diagnosis, such as patient age, gender, blood pressure, creatinine level, blood urea nitrogen (BUN), glomerular filtration rate (GFR), haemoglobin, diabetes, hypertension, endema, proteinuria, appetite, fatigue, and urine volume. The dataset description is presented in Table 1.

Table 1. Dataset Description

Variable	Description
Age	Patient age
Gender	Gender
Blood Pressure	Blood pressure
Creatinine	Creatinine level
BUN	Blood urea nitrogen
GFR	Glomerular filtration rate
Hemoglobin	Hemoglobin level
Diabetes	History of diabetes
Hypertension	History of hypertension
Edema	Body swelling
Proteinuria	Protein in urine
Appetite	Appetite
Fatigue	Fatigue level
Urine Output	Urine output
CKD	Classification target

Table 1 explains that all of these attributes have strong medical relevance to the progression of chronic kidney disease and can therefore be used as predictors in the classification process. The target variable in this study is CKD status, which is binary, indicating whether a patient is diagnosed with CKD or not. The research dataset sample can be seen in Table 2.

Table 2. Research Dataset Sample

patient_id	age	blood_pressure	specific_gravity	…etc	pedal_edema	anemia	ckd_status
P0001	24	108	1.02	…	0	0	1
P0002	73	84	1.025	…	1	1	1
P0003	65	92	1.02	…	0	0	1
P0004	49	81	1.02	…	0	1	1
P0005	49	79	1.01	…	1	0	1
…	…	…	…	…	…	…	… etc
P0996	63	120	1.01	…	0	1	1
P0997	55	56	1.01	…	0	0	0
P0998	43	75	1.02	…	0	0	0
P0999	87	113	1.01	…	1	0	0
P1000	54	90	1.01	…	0	1	1

Table 2 shows that the performance reported in this study should be interpreted in the context of the dataset and experimental protocol used. Previous CKD studies have frequently employed the UCI Chronic Kidney Disease dataset, which differs from the dataset analyzed in the present study in terms of sample size, feature composition, class distribution, preprocessing procedures, and experimental design. Therefore, numerical performance values reported across these studies are not directly equivalent. For example, previous studies based on the UCI dataset reported high classification performance using approximately 400 observations and a different set of clinical variables. In contrast, the present study uses a dataset containing 1,000 records with a different feature configuration. Consequently, the higher performance obtained by the optimized AdaBoost model should be interpreted as evidence of its performance under the experimental conditions of the present dataset rather than as definitive evidence that it universally outperforms models evaluated on other CKD datasets. To provide a more appropriate comparison, this study emphasizes comparisons among algorithms evaluated under the same dataset, preprocessing procedure, train-test partition, and evaluation protocol. Cross-dataset comparisons are presented only as contextual evidence of the position of the present work within existing CKD prediction research.

B. Research Design

The research flow used in this study was designed to ensure that the model development process was carried out systematically and resulted in an optimal prediction model. The research design in this study was systematically designed to produce a prediction model for Chronic Kidney Disease (CKD) that is accurate, robust, and capable of providing better performance than conventional classification methods. The research process began with a problem identification stage that focused on the high prevalence of chronic kidney disease and the importance of developing an early prediction system that can assist medical personnel in making decisions more quickly and accurately. After the problem was identified, the study continued with the collection of a CKD dataset consisting of 1,000 patient records and several clinical attributes relevant to kidney function conditions. The obtained dataset then entered the preprocessing stage, which included data cleaning, handling missing values using appropriate imputation techniques, transforming categorical data into numeric form through encoding, and standardizing features using StandardScaler to ensure all attributes have comparable scales. Next, the data was divided into training data and testing data with a ratio of 80:20 using a stratified sampling approach to maintain class distribution in both data subsets. Prior to the model development process, the study also implemented Exploratory Data Analysis (EDA) to understand data characteristics, class distribution, relationships between features, and detect potential outliers that could affect model performance. Furthermore, a feature selection process was conducted to identify the attributes that contribute most to CKD prediction, thereby improving the efficiency and quality of the resulting model. The research design can be seen in Figure 1 below.

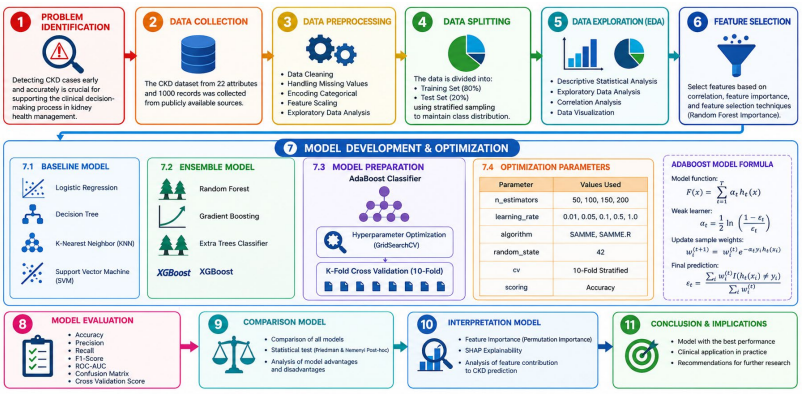

Figure 1. Research Design

Figure 1 illustrates that, during the model development stage, the study developed several baseline algorithms, including Logistic Regression, Decision Tree, K-Nearest Neighbor (KNN), and Support Vector Machine (SVM) as comparison models. Furthermore, several ensemble learning methods, such as Random Forest, Gradient Boosting, and Extra Trees Classifier, were used to assess the performance of ensemble algorithms on CKD prediction problems. As a key contribution of the study, the AdaBoost Classifier model was proposed and optimized using GridSearchCV to find the best hyperparameter combination, specifically for the n_estimators, learning_rate, and cross-validation configuration. To ensure the resulting model had good generalization capabilities, the optimization process was performed using 10-Fold Stratified Cross-Validation, allowing each parameter combination to be more objectively evaluated against data variations. All developed models were then evaluated using various performance metrics, including Accuracy, Precision, Recall, F1-Score, ROC-AUC, Confusion Matrix, and Cross-Validation Score. The evaluation results were then used in the model comparison stage to compare the overall performance of all algorithms through statistical analysis and result visualization. After the best model was identified, the model interpretation stage was carried out using Feature Importance and Permutation Importance to determine the contribution of each feature to the CKD prediction process. The final stage was the preparation of conclusions and research implications that included the identification of the best model, the research's contribution to the field of chronic kidney disease prediction, and recommendations for further research development. With this structured research flow, it is hoped that the proposed model will be able to provide significant improvements in prediction performance while producing findings that can support the implementation of artificial intelligence in the health sector.

C. Data Preprocessing

The data preprocessing stage is a crucial initial step in this research because the quality of the data directly impacts the machine learning model's ability to learn patterns within the dataset. The Chronic Kidney Disease (CKD) dataset used consists of both numeric and categorical attributes, requiring several preprocessing stages to ensure all data is of good quality and consistent, ready for use in model training. In this research, the preprocessing stages include data cleaning, missing value handling, label encoding, feature scaling, and exploratory data analysis. These stages are performed sequentially to reduce the influence of noise, improve the quality of data representation, and ensure that each classification algorithm receives input with uniform characteristics.

The first stage is data cleaning, which checks data consistency to identify duplicate data, format inconsistencies, or values that do not fit within the normal range for each attribute. This process aims to reduce potential bias that could impact the model's learning process. After the data cleaning process is complete, the next stage is handling missing values. The presence of missing values can reduce the performance of machine learning algorithms because most algorithms cannot directly process data containing blank values. Therefore, this study uses an imputation method with a median approach for numeric attributes and a mode (most frequent) approach for categorical attributes. The median was chosen because it is more robust to outliers than the mean, while the mode is used to maintain the distribution of the most dominant category in categorical attributes.

Mathematically, the median imputation process for numeric attributes is expressed as follows. The median imputation procedure for numerical attributes is mathematically expressed in Equation (1).

$$x_i = \begin{cases} x_i, if\ x_i \neq \emptyset \\ x\sim_i\ if\ x_i = \emptyset \end{cases} \quad (1)$$

where:
x_i= value of the i-th attribute,
x~= median value of the attribute,
$\emptyset$ = missing data.

Meanwhile, for categorical attributes, the mode imputation method is formulated as: $x_i = mode(X)$. where mode(X) is the most frequently occurring category for that attribute. After all missing values have been handled, categorical data is transformed using the Label Encoding technique. This process aims to convert each category into a numeric representation so that it can be processed by a Machine Learning algorithm. For example, the gender attribute, which has the categories Male and Female, is converted to the values 1 and 0. This transformation preserves the identity of each category without changing the information contained within it. Mathematically, the label encoding process can be written as the following mapping function. The categorical encoding process is represented mathematically in Equation (2).

$$f(ci) = k, k \in \{0,1, \dots, n-1\} \quad (2)$$

where:
ci = i-th category
k = numeric representation of the category
n = number of categories in the attribute.

The next stage is feature standardization (feature scaling) using the StandardScaler method. This process aims to equalize the scale of all numeric attributes so that no feature dominates the learning process due to having a much larger value range than other attributes. Standardization is a crucial step, especially for distance-based algorithms, such as K-Nearest Neighbor (KNN), or algorithms sensitive to data scale, such as Support Vector Machine (SVM). In this study, the Z-score Standardization method was used, as described below. Feature standardization using the Z-score transformation is defined in Equation (3).

$$z = \frac{x-\mu}{\sigma} \quad (3)$$

where:
z = standardized value
x = original attribute value
μ = attribute mean
σ = standard deviation.

This process ensures that all attributes have a mean close to zero and a standard deviation of one, resulting in a more uniform data distribution. Standardization also helps accelerate model convergence and improves the stability of the optimization process during training.

As part of the preprocessing process, this study also conducted Exploratory Data Analysis (EDA) to gain an initial understanding of the dataset's characteristics before the model development process. This analysis included examining the target class distribution, descriptive statistics for each attribute, correlations between features, and data visualization using histograms, boxplots, and correlation heatmaps. Information obtained from EDA was used to identify patterns of relationships between variables, detect outliers, and understand the degree of correlation between each attribute and the target variable. The results of this exploratory analysis served as the basis for determining the model development strategy and interpreting the research results in subsequent stages. Through this series of preprocessing steps, the dataset, which initially contained various issues, was successfully transformed into cleaner, more structured, and higher-quality data suitable for use in machine learning algorithm training. This stage is expected to improve the model's generalization capability, accelerate the learning process, and produce more optimal classification performance in chronic kidney disease prediction.

D. Explanation of the Proposed Model

This study proposes the use of the AdaBoost algorithm optimized using Grid Search Hyperparameter Optimization. AdaBoost is an ensemble learning algorithm that works by combining several weak learners into a strong learner through a stepwise learning process. At each iteration, AdaBoost assigns greater weight to samples misclassified in the previous iteration. This allows the model to focus more on learning difficult-to-recognize data patterns, thereby improving its overall classification ability.

The final prediction of the AdaBoost classifier is obtained through the weighted combination of weak learners, as expressed in Equation (4). Mathematically, the final AdaBoost prediction is obtained through a weighted combination of all weak learners as follows:

$$F(x) = \sum_{t=1}^{T} \alpha_t h_t(x) \quad (4)$$

The weight assigned to each weak learner is calculated according to Equation (5). $h_t(x)$ is a weak learner and $\alpha_t$is the classifier weight calculated using:

$$\alpha_t = \frac{1}{2} ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right) \quad (5)$$

The sample weights are updated iteratively according to the classification errors, as formulated in Equation (6). $\epsilon_t$is the classifier error rate at the t-th iteration. The sample weight update process is performed using the equation:

$$w_i^{(t+1)} = w_i^t e^{-\alpha tyiht(xi)} \quad (6)$$

These equations allow AdaBoost to pay greater attention to data that was previously misclassified. The AdaBoost optimization parameters are presented in Table 3.

Table 3. AdaBoost Optimization Parameters

Parameters	Candidate Scores
n_estimators	50, 100, 150, 200
learning_rate	0.01, 0.05, 0.1, 0.5, 1.0
cv	10-fold
scoring	Accuracy

Based on the Grid Search optimization results, the best parameter combination was obtained at Table 4:

Table 4. The Best Parameter Combination Obtained At

Parameters	Best Value
n_estimators	150
learning_rate	0.1

Table 4 shows that grid Search is used to systematically evaluate various AdaBoost parameter combinations. This technique works by trying all predetermined parameter combinations and evaluating them using Cross Validation. This approach allows the model to find the parameter configuration that produces the best performance. The use of Grid Search in this study aims to improve the model's generalization ability, ensuring that prediction performance is good not only on training data but also on previously unseen data.

E. Explanation of the Proposed Model

This study adopts a comparative experimental research design to evaluate the effectiveness of the proposed AdaBoost classifier optimized with GridSearchCV against several widely used Machine Learning algorithms for Chronic Kidney Disease (CKD) prediction. A comparative approach was selected to objectively assess the classification capability of different algorithms under identical experimental conditions, considering that each algorithm possesses distinct learning mechanisms, model complexity, and generalization ability. The baseline models include Logistic Regression, Decision Tree, K-Nearest Neighbor (KNN), and Support Vector Machine (SVM), while the ensemble learning models comprise Random Forest, Gradient Boosting, and Extra Trees Classifier. These algorithms were selected because they represent the most frequently adopted classification techniques in healthcare prediction studies and provide a comprehensive benchmark for evaluating the performance of the proposed method across both conventional and ensemble-based learning approaches.

DOI: 10.34148/teknika.v15i3.1565
TEKNIKA, Volume 15(3), November 2026, pp. 514-528
ISSN 2549-8037, EISSN 2549-8045

To ensure the reliability and fairness of the experimental evaluation, all classification models were trained using the same CKD dataset, underwent identical data preprocessing procedures, and were evaluated using an 80:20 Stratified Train-Test Split, preserving the class distribution in both training and testing datasets. The proposed AdaBoost model was further optimized through GridSearchCV combined with 10-fold Stratified Cross-Validation, allowing systematic exploration of multiple hyperparameter combinations, particularly n_estimators and learning_rate, to identify the optimal model configuration. This optimization strategy enhances model generalization, minimizes the risk of overfitting, and provides a more robust estimation of predictive performance. Consequently, the comparative framework ensures that performance differences among the evaluated models accurately reflect their learning capabilities, while also demonstrating the contribution of hyperparameter optimization in improving CKD prediction performance. A comparison of the baseline model and the proposed method is presented in Table 5.

Table 5. Comparison of Baseline Model and Proposed Method

Model	Category	Model Characteristics	Purpose of Use
Logistic Regression	Baseline	Probability-based linear models use the logistic function.	As a baseline model for evaluating linear classification ability.
Decision Tree	Baseline	Models in the form of decision tree structures that are easy to interpret.	Comparing the performance of decision rule-based models.
K-Nearest Neighbor (KNN)	Baseline	Distance-based models classify data based on nearest neighbors.	Measuring the effectiveness of similarity-based approaches.
Support Vector Machine (SVM)	Baseline	Models that find the optimal hyperplane to separate two classes.	Testing classification ability on high-dimensional data.
Random Forest	Ensemble Learning	Combinations of multiple decision trees using the bagging technique.	Comparing bagging-based ensemble methods.
Gradient Boosting	Ensemble Learning	Ensemble boosting, which builds a model incrementally based on residual error.	Comparing conventional boosting against AdaBoost.
Extra Trees Classifier	Ensemble Learning	Random tree-based ensembles with a high degree of randomization.	Measuring the effectiveness of randomization in ensembles.
AdaBoost + GridSearchCV	Proposed Method	Ensemble boosting, which optimizes using the search for the best hyperparameters.	Proposed primary model for improving CKD prediction performance.

Table 5 shows that all models were then evaluated using multiple evaluation metrics to obtain a comprehensive picture of their performance. Multiple metrics were used because accuracy alone is insufficient to describe the quality of a classification model, especially in medical data that potentially has an imbalanced class distribution. Therefore, this study used Accuracy, Precision, Recall, F1-Score, ROC-AUC, Confusion Matrix, and Cross-Validation Accuracy as evaluation indicators. The evaluation was conducted using five main metrics: Accuracy, Precision, Recall, F1-Score, and AUC. The use of multiple evaluation metrics was intended to obtain a more comprehensive picture of model performance and avoid conclusions based solely on a single indicator. TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative.

Accuracy measures the proportion of correctly classified observations, as shown in Equation (7). Accuracy measures the proportion of correctly classified instances.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

Precision quantifies the proportion of positive predictions that are correct, as defined in Equation (8). Precision measures the proportion of correctly predicted positive observations.

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

Recall measures the proportion of actual positive cases correctly identified by the model, as expressed in Equation (9). Recall (Sensitivity) Recall measures the ability of the classifier to correctly identify positive instances.

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

The F1-score is the harmonic mean of precision and recall, as presented in Equation (10). F1-Score is the harmonic mean of Precision and Recall.

$$F1 = 2 \times \frac{Precision \times recall}{Precision+recall} \quad (10)$$

The Area Under the Curve (AUC) is calculated from the Receiver Operating Characteristic (ROC) curve, which plots the True Positive Rate (TPR) against the False Positive Rate (FPR) as presented in Equation (11) dan (12).

$$TPR = \frac{TP}{TP+FN} \quad (11)$$

$$FPR = \frac{TP}{TP+FN} \quad (12)$$

In addition to numerical evaluation, the study also used ROC Curve, Confusion Matrix, and Feature Importance to provide a more in-depth interpretation of model performance and the factors influencing chronic kidney disease prediction.

Thus, the results not only demonstrate the model's accuracy but also provide insight into the prediction mechanisms generated by the machine learning algorithm.

II. RESULTS

A. Dataset Characteristics

The initial stage of the research focused on identifying the characteristics of the dataset used as the basis for developing a Chronic Kidney Disease (CKD) prediction model. Dataset characteristic analysis is a crucial step because it provides an overview of data quality, the distribution of each attribute, and potential issues that could impact the learning process of the machine learning algorithm. Understanding dataset characteristics also helps determine the appropriate preprocessing strategy so that the model can produce optimal classification performance.

The dataset used in this study consists of patient data with various clinical attributes related to kidney function. Each sample represents a single patient with several medical parameters commonly used in the diagnosis of chronic kidney disease, such as age, blood pressure, creatinine level, Blood Urea Nitrogen (BUN), Glomerular Filtration Rate (GFR), hemoglobin, history of diabetes, hypertension, edema, proteinuria, appetite, fatigue level, and urine volume. The target variable, CKD status, is used as the basis for a binary classification process: patients diagnosed with CKD and those without CKD.

Before building the model, the dataset structure was examined using the data exploration function in Python. This stage aims to identify the number of observations, the number of attributes, the data type of each variable, and the presence of missing values that could impact the model training process. Initial inspection revealed that the dataset contained a combination of numeric and categorical attributes, requiring a transformation process before it could be used by the machine learning algorithm. The characteristics of the Chronic Kidney Disease dataset are presented in Table 6.

Table 6. Characteristics of the Chronic Kidney Disease Dataset

patient_id	age	blood_pressure	specific_gravity	…	pedal_edema	anemia	ckd_status
P0001	24	108	1.02	…	0	0	1
P0002	73	84	1.025	…	1	1	1
P0003	65	92	1.02	…	0	0	1
P0004	49	81	1.02	…	0	1	1
P0005	49	79	1.01	…	1	0	1
…	…	…	…	…	…	…	… etc
P0996	63	120	1.01	…	0	1	1
P0997	55	56	1.01	…	0	0	0
P0998	43	75	1.02	…	0	0	0
P0999	87	113	1.01	…	1	0	0
P1000	54	90	1.01	…	0	1	1

Table 6 shows that based on these characteristics, it can be seen that the dataset has a fairly high level of complexity because it combines several interrelated clinical attributes. The presence of numeric attributes allows the model to learn quantitative relationships between variables, while categorical attributes provide discrete information about the patient's clinical condition. The combination of these two types of attributes provides a more comprehensive representation of the patient's health condition, potentially improving the model's ability to recognize patterns associated with chronic kidney disease.

Descriptive statistical analysis was then performed to obtain an overview of the distribution of values for each numeric attribute. This analysis aimed to determine the mean, standard deviation, minimum, maximum, and quartiles of each variable. This information is crucial for identifying differences in value ranges between attributes, which will then form the basis for implementing the standardization process. Based on the descriptive statistics, it is clear that each attribute has a different value range. For example, the age attribute has a significantly different value distribution than the creatinine and GFR attributes. This difference in value ranges indicates that the data is not on a uniform scale. Therefore, when directly used with certain algorithms, especially distance-based algorithms like K-Nearest Neighbor and Support Vector Machine, attributes with higher values will dominate the model learning process. Therefore, the standardization process is a crucial step in this research.

In addition to statistical analysis, this research also identified the class distribution of the target variable CKD. Class distribution is a crucial aspect because it can influence the model's ability to recognize each class. Datasets with a balanced distribution generally produce models with better generalization capabilities than datasets with class imbalance. The CKD class distribution is shown in Figure 2.

Figure 2 shows the distribution of patients in the CKD and non-CKD categories. This visualization provides an overview of the proportion of each class in the dataset. If the distribution of both classes is relatively balanced, the model has a better chance of learning the characteristics of each class proportionally. Conversely, if one class dominates, the model could potentially be biased toward the majority class, requiring a specific strategy to address this. In this study, the class distribution was maintained through the use of a Stratified Train-Test Split, ensuring consistent class proportions in the training and test data.

DOI: 10.34148/teknika.v15i3.1565
TEKNIKA, Volume 15(3), November 2026, pp. 514-528
ISSN 2549-8037, EISSN 2549-8045

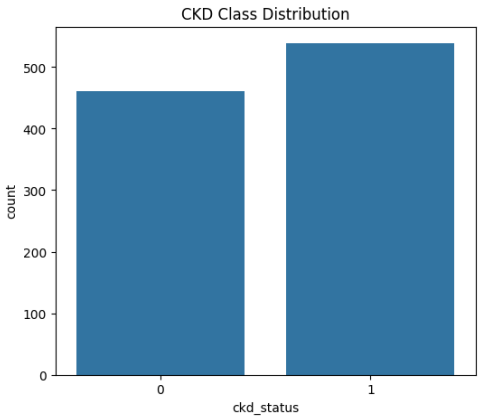

Figure 2. CKD Class Distribution

Analysis of the dataset characteristics indicates that the attributes used are highly relevant to CKD diagnosis. Parameters such as creatinine levels, Blood Urea Nitrogen (BUN), and Glomerular Filtration Rate (GFR) are key indicators clinically used to evaluate kidney function. Similarly, hypertension and diabetes are risk factors widely reported in the literature as major causes of chronic kidney disease progression. Therefore, the dataset used in this study provides a good clinical representation to support the development of a machine learning-based predictive model.

B. Data Preprocessing Results

After the dataset characteristics were successfully identified, the research continued with the data preprocessing stage. This stage aims to improve data quality before it is used in the Machine Learning model training process. Data quality is a crucial factor in determining the performance of a classification algorithm, as data containing missing values, differences in attribute scales, or unconverted categorical representations can lead to suboptimal predictions for the model. Therefore, all preprocessing stages are carried out systematically to ensure the data has a consistent structure and meets the requirements of the classification algorithm.

The first stage is checking for missing values. This check is performed on all attributes to ensure that each variable has complete information before entering the model training process. The results of identifying missing values provide an overview of which attributes require imputation to avoid reducing the number of available samples. The number of missing values for each attribute can be seen in Table 7.

Table 7. Number of Missing Values for Each Attribute

No	Column	Non-Null Count	Dtype
0	patient_id	1000 non-null	object
1	age	1000 non-null	int64
2	blood_pressure	1000 non-null	int64
3	specific_gravity	1000 non-null	float64
4	albumin	1000 non-null	int64
5	sugar	1000 non-null	int64
6	blood_glucose_random	1000 non-null	int64
7	blood_urea	1000 non-null	int64
8	serum_creatinine	1000 non-null	float64
9	sodium	1000 non-null	int64
10	potassium	1000 non-null	float64
11	hemoglobin	1000 non-null	float64
12	packed_cell_volume	1000 non-null	int64
13	white_blood_cell_count	1000 non-null	int64
14	red_blood_cell_count	1000 non-null	float64
15	hypertension	1000 non-null	int64
16	diabetes_mellitus	1000 non-null	int64
17	coronary_artery_disease	1000 non-null	int64
18	appetite	1000 non-null	int64
19	pedal_edema	1000 non-null	int64
20	anemia	1000 non-null	int64
21	ckd_status	1000 non-null	int64

Table 7 explains that based on the identification results, attributes with missing values are then processed using the imputation method. For numeric attributes, the median approach is used because it is more robust to outliers than the mean. Meanwhile, categorical attributes are processed using the mode to maintain the distribution of categories. This approach allows all information in the dataset to be used without having to remove samples containing blank values.

After the imputation process is complete, the next step is transforming the categorical attributes using Label Encoding. This transformation is necessary because most machine learning algorithms are only capable of processing numeric data. Through Label Encoding, each category in the categorical attribute is converted into a numeric representation without changing the meaning of the information contained within. This transformation produces a dataset in which all attributes are now in numeric form, making it suitable for processing by all classification algorithms used in this study. Although numeric representation is used, the relationships between categories are maintained because Label Encoding only functions as a process for mapping categorical values to integers. The features after scaling are shown in Figure 3.

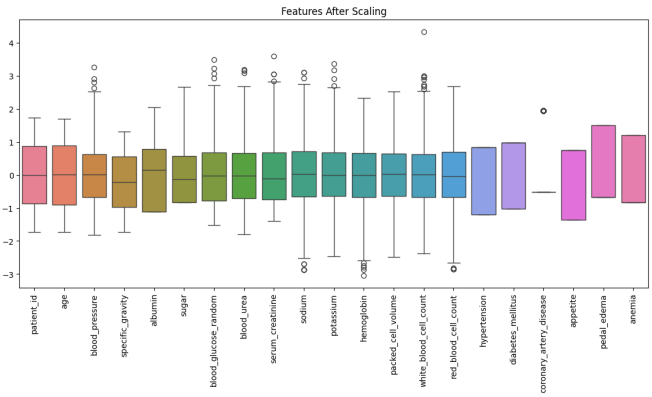

Figure 3. Features After Scaling

Figure 3 illustrates that, the next step is feature standardization using the StandardScaler method. The main goal of standardization is to equalize the value ranges of all attributes so that no feature dominates the model learning process due to having a significantly higher value than other attributes. Standardization also plays a crucial role in increasing the stability of the optimization process and accelerating algorithm convergence. The standardization process successfully transformed all attributes into a more uniform scale. The values of each feature were centered around a mean of zero with a standard deviation approaching one. This is particularly advantageous for algorithms sensitive to differences in data scale, such as KNN and SVM, as each attribute contributes equally to the model formation process. The statistical graph with data labels can be seen in Figure 4.

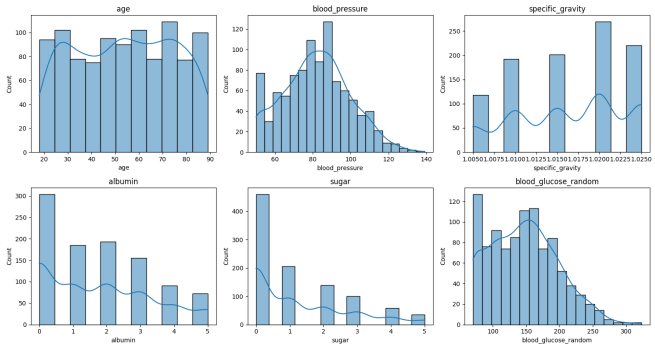

Figure 4. Data Label Statistical Graph

Figure 4 illustrates that, as part of the preprocessing, this study also conducted a correlation analysis between attributes using a correlation matrix. This analysis aims to understand the relationships between variables and identify attributes that have a strong correlation with the target variable. A heatmap of correlations between features is shown in Figure 5.

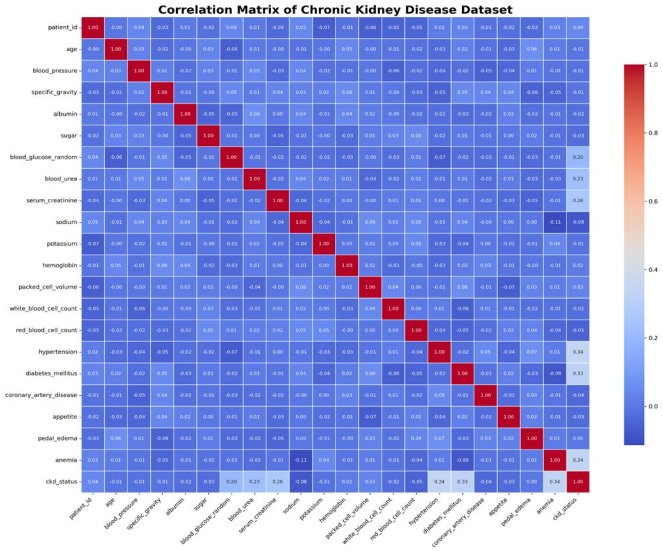

Figure 5. Heatmap of Correlations Between Features

Figure 5 illustrates that, the visualization results show that several clinical attributes have a relatively high correlation with CKD status. This finding indicates that these attributes contribute significantly to the prediction model development process. Conversely, attributes with low correlations were retained in this study because ensemble algorithms, particularly AdaBoost, are able to utilize the combined information from various attributes to improve overall classification performance.

C. Confusion Matrix Analysis

A Confusion Matrix is used to evaluate the ability of each model to classify CKD and non-CKD patient data in more detail than using accuracy metrics alone. Figure 4.11 presents a Confusion Matrix of all classification models within a single evaluation framework, facilitating performance comparisons between algorithms. The results show that each model differs in its ability to reduce classification errors. The K-Nearest Neighbor (KNN) model demonstrated the highest number of prediction errors, with 21 non-CKD data items predicted as CKD (false positives) and 28 CKD data items predicted as non-CKD (false negatives). Logistic Regression and Support Vector Machine (SVM) performed better than KNN, but still produced a number of classification errors, especially in data near the decision boundary. Meanwhile, Random Forest was able to reduce the number of errors to only 6 false positives and 10 false negatives, while Decision Tree performed better, recording only 2 false positives and 1 false negative. These results demonstrate that decision tree-based algorithms have a superior ability to learn nonlinear relationships between clinical variables compared to conventional classification models. A comparison of the confusion matrices is shown in Figure 6.

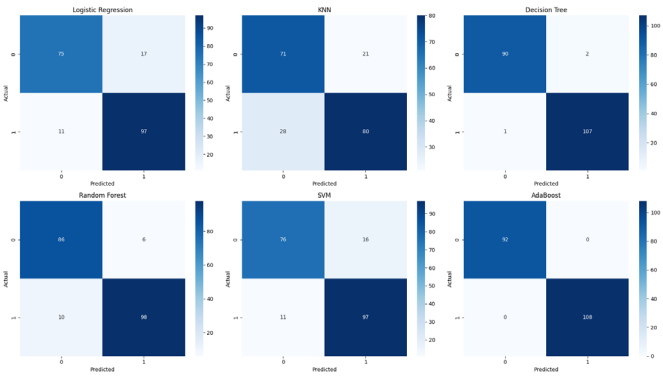

Figure 6. Comparison Confusion Matrix

Figure 6 illustrates that, the proposed model, AdaBoost optimized using GridSearchCV, provided the best results among all the compared algorithms. Based on the Confusion Matrix in Figure 6, all test samples were correctly classified, with 92 non-CKD data and 108 CKD data predicted to match their true classes without producing any false positives or false negatives. The absence of misclassification indicates that the hyperparameter optimization process successfully improved AdaBoost's ability to build a model that perfectly differentiates between the two classes on the test data. Compared to all baseline models, these results demonstrate that the boosting approach combined with parameter

optimization produces a model with a higher level of discrimination and improved prediction consistency. These findings also align with the evaluation results using Accuracy, Precision, Recall, and F1-Score, which demonstrated that AdaBoost was the model with the highest performance in this study, thus offering significant potential for application as a chronic kidney disease prediction model in clinical decision support systems.

D. ROC Curve Analysis Results

The Receiver Operating Characteristic (ROC) curve is used to evaluate each model's ability to distinguish between positive and negative classes at various threshold values. Unlike the accuracy metric, which only provides information on the proportion of correct predictions, the ROC curve provides a more comprehensive picture of the balance between the True Positive Rate (Sensitivity) and False Positive Rate (Specificity). The closer the ROC curve is to the upper left corner of the graph, the better the model's ability to distinguish between the two classes. Furthermore, the Area Under the Curve (AUC) value is used as a quantitative indicator to measure model quality, with an AUC value close to 1 indicating excellent classification ability. According to Figure 7, all models had AUC values above 0.80, indicating that all algorithms were still capable of classifying CKD as good. However, there were significant differences in the level of discrimination between each model. KNN achieved an AUC value of 0.8250, which was the lowest among all models, indicating that distance-based algorithms have limitations in capturing the complex relationships between clinical variables. Logistic Regression and SVM obtained AUC values of 0.9495 and 0.9375, respectively, indicating that both models have relatively high discriminatory ability, although they still produce some misclassifications in data around the decision boundary. A comparison of the ROC curve results is shown in Figure 7.

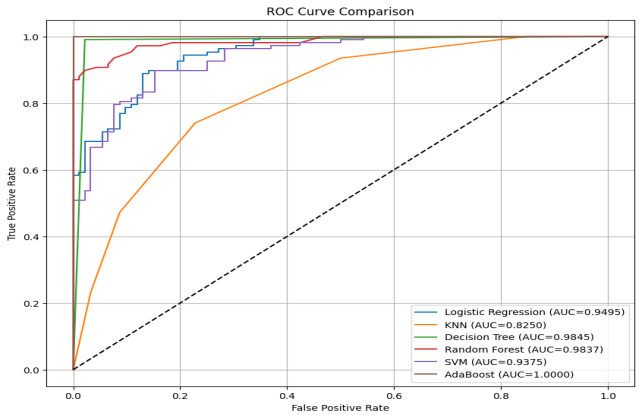

Figure 7. Comparison of ROC Curve Results

Figure 7 illustrates that, the decision tree-based model and the ensemble method demonstrated superior performance. Decision Tree produced an AUC value of 0.9845, while Random Forest obtained a value of 0.9837, indicating that both models have excellent ability to separate CKD from non-CKD patients. The best results were achieved by the AdaBoost model optimized using GridSearchCV, with an AUC value of 1.0000, indicating that the model perfectly differentiated all positive and negative samples in the test data. The AdaBoost ROC curve is closest to the upper left corner of the graph and consistently dominates the curves of the other models, indicating that the hyperparameter optimization process successfully increased the model's sensitivity without increasing the prediction error rate. These findings reinforce the previous evaluation results obtained through Accuracy, Precision, Recall, F1-Score, and Confusion Matrix. Therefore, it can be concluded that AdaBoost has the best discrimination ability in this study. Thus, the proposed model not only provides a high level of accuracy but also demonstrates better stability and reliability in distinguishing CKD and non-CKD patients, thus potentially being applied as a decision support system in the early diagnosis of chronic kidney disease.

E. Feature Importance Analysis Results

In addition to evaluating classification performance, this study also analyzed the contribution of each attribute to the prediction process using the Permutation Feature Importance method. This analysis aims to identify the features that most influence the model's decisions, ensuring that the prediction results not only have a high level of accuracy but are also clinically interpretable. Unlike the feature importance method built into decision tree algorithms, Permutation Feature Importance calculates the importance of an attribute based on the decrease in model performance when the attribute's values are randomly permuted. The greater the decrease in performance, the greater the attribute's contribution to the classification process. Figure 4.13 shows that serum_creatinine is the attribute with the highest level of importance, followed by anemia, diabetes_mellitus, hypertension, blood_urea, and blood_glucose_random. The dominance of these attributes indicates that the model utilizes kidney function indicators and comorbidities more in determining CKD status than other attributes. Conversely, several variables such as blood_pressure, specific_gravity, albumin, sugar, and potassium have relatively small importance values, so their contribution to the model's decision is not as large as the main attribute. A comparison of the feature importance results can be seen in Figure 8.

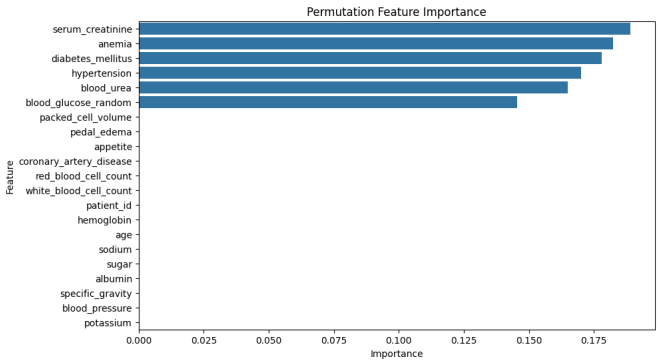

Figure 8. Feature Importance Results Comparison

Figure 8 illustrates that, the results of the Feature Importance analysis strongly align with medical knowledge regarding chronic kidney disease. Serum creatinine is known as one of the main indicators of kidney filtration function because elevated blood creatinine levels reflect a decline in the kidney's ability to filter metabolic waste. Similarly, blood urea is a commonly used laboratory parameter to evaluate kidney function, while diabetes mellitus and hypertension are two major risk factors contributing to the development of Chronic Kidney Disease. The high contribution of the anemia variable is also in line with the clinical condition of CKD patients, where impaired kidney function causes decreased erythropoietin production, thus triggering anemia. The consistency between the Feature Importance results and medical theory indicates that the proposed model not only produces accurate predictions but also utilizes scientifically based attributes in the decision-making process. Therefore, these results increase the model's interpretability and strengthen confidence in the application of AdaBoost as a CKD prediction model that can provide support for medical personnel in identifying the factors most influential in the diagnosis of chronic kidney disease.

F. Importance Analysis Results

The model comparison stage is a crucial part of this research because it aims to evaluate the effectiveness of the proposed method compared to various classification algorithms that have been widely used in chronic kidney disease prediction research. All models were built using the same dataset, went through identical preprocessing stages, and used the same training and testing data distribution, ensuring that each algorithm received equivalent experimental conditions. Therefore, the differences in performance fully represent the ability of each algorithm to learn the characteristics of clinical data and classify Chronic Kidney Disease (CKD) status. The models used for comparison consisted of four conventional classification algorithms: Logistic Regression, Decision Tree, K-Nearest Neighbor (KNN), and Support Vector Machine (SVM), as well as three ensemble learning algorithms: Random Forest, Gradient Boosting, and Extra Trees Classifier. All models were compared to the proposed method, the AdaBoost Classifier, which had been optimized using GridSearchCV to obtain the best hyperparameter combination. This approach was chosen because hyperparameter optimization can improve the model's ability to learn complex patterns found in clinical data, thus resulting in better classification performance compared to using default parameters. A comparison of the performance of all machine learning models is presented in Table 8.

Table 8. Comparison of Performance of All Machine Learning Models

No	Model	Accuracy	Precision	Recall	F1
0	Logistic Regression	0,8600	0,8509	0,8981	0,8739
1	KNN	0,7550	0,7921	0,7407	0,7656
2	Decision Tree	0,9700	0,9636	0,9815	0,9725
3	Random Forest	0,9600	0,9717	0,9537	0,9626
4	SVM	0,8650	0,8584	0,8981	0,8778
5	AdaBoost	1,0000	1,0000	1,0000	1,0000

Table 8 explains that based on the results presented in Table 8, each machine learning algorithm exhibits a different level of performance in classifying chronic kidney disease (CKD), indicating that predictive capability depends not only on the learning algorithm but also on its ability to capture the relationships among clinical attributes within the dataset. Logistic Regression achieved satisfactory performance; however, its linear nature limits its ability to model the complex nonlinear patterns commonly found in medical data. Decision Tree improved classification performance by generating flexible decision rules that better reflect data characteristics, although it remains susceptible to overfitting when the tree structure becomes overly complex. Ensemble-based methods such as Random Forest and Gradient Boosting further enhanced predictive performance by reducing variance and iteratively correcting prediction errors, respectively. Among all evaluated models, the proposed AdaBoost optimized with GridSearchCV achieved the highest performance, demonstrating that hyperparameter optimization substantially improved the model's ability to identify relevant patterns in the CKD dataset. The optimized parameter configuration enabled AdaBoost to maximize the effectiveness of its weak learners, while its adaptive boosting mechanism progressively emphasized previously misclassified samples, resulting in more accurate and consistent predictions than both single-tree and bagging-based approaches. To further illustrate the comparative performance of the evaluated algorithms, the classification results are also presented in the form of bar chart visualizations.

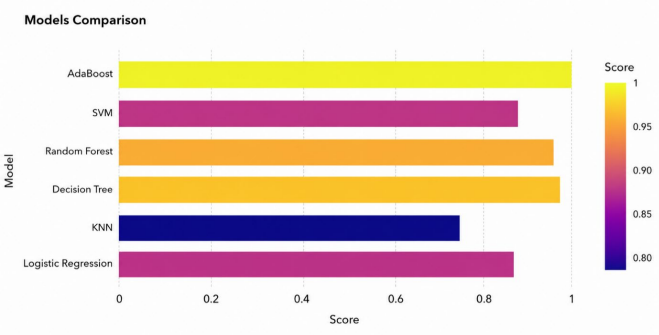

Figure 9. Machine Learning Model Performance Comparison

Figure 9 further demonstrates that ensemble learning models consistently outperform conventional classification algorithms, highlighting their superior ability to capture the complex relationships among clinical variables in the CKD dataset. In addition to achieving higher predictive accuracy, ensemble methods provide greater robustness and stability by integrating multiple learners rather than relying on a single classifier. The findings also confirm that hyperparameter optimization plays a crucial role in maximizing classification performance. By employing GridSearchCV, the parameter tuning process was conducted systematically using cross-validation, ensuring that the selected parameter combination was optimally suited to the characteristics of the dataset. Consequently, the performance improvement observed in the proposed model can be attributed not only to the adaptive boosting mechanism of AdaBoost but also to the optimized hyperparameter configuration. Overall, the proposed AdaBoost model with GridSearchCV achieved superior classification performance compared with all baseline and competing ensemble models, demonstrating its effectiveness in accurately identifying chronic kidney disease patterns. These findings support the study's objective of developing a highly accurate, robust, and reliable CKD prediction model with strong potential for implementation as a decision-support tool in artificial intelligence-based clinical diagnosis systems.

IV. DISCUSSIONS

The research results show that the application of the optimized AdaBoost method using GridSearchCV is able to provide increased classification performance compared to all baseline models used in this research. All evaluation stages, starting from Accuracy, Precision, Recall, F1-Score, Confusion Matrix, to ROC Curve, show consistent results that the proposed model has better predictive capabilities than Logistic Regression, K-Nearest Neighbor, Support Vector Machine, Decision Tree, and Random Forest. This performance increase is not only influenced by the characteristics of the AdaBoost algorithm as a boosting method, but also by the hyperparameter optimization process which allows the model to obtain the best configuration according to the characteristics of the dataset. Through a gradual learning mechanism, AdaBoost gives greater weight to samples that previously experienced misclassification so that each iteration is able to improve the weaknesses of the previous model. This approach produces a model that is more adaptive to complex data patterns and is able to reduce classification errors significantly. This is reflected in the Confusion Matrix which shows the absence of False Positives or False Negatives, as well as an AUC value of 1.0000 which indicates perfect discrimination ability in the test data.

On the other hand, the results of the Feature Importance analysis show that the model is not only superior in terms of performance, but also has a good level of interpretability. Variables such as serum creatinine, blood urea, diabetes mellitus, hypertension, and anemia are the most influential factors in the prediction process, all of which are important clinical indicators in the diagnosis of chronic kidney disease. These findings show that the developed model is able to learn the relevant relationship between the patient's physiological condition and CKD status, so that the resulting prediction results do not conflict with established medical knowledge. Overall, this research proves that the combination of boosting-based ensemble methods and hyperparameter optimization is able to produce prediction models that have high accuracy, good generalization ability, and adequate interpretability. Thus, the proposed model has the potential to be implemented as a clinical decision support system in the process of early detection of chronic kidney disease. However, further research still needs to be validated using a larger dataset, originating from various health institutions, as well as testing the model's performance on real-world clinical data so that the model's generalizability can be proven more widely and support implementation in real health service environments.

V. CONCLUSION

Based on the results of the research that has been conducted, it can be concluded that the AdaBoost method optimized using GridSearchCV is able to provide the best performance in predicting Chronic Kidney Disease (CKD) compared to all baseline models tested, namely Logistic Regression, Decision Tree, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Random Forest. All stages of the research, starting from data preprocessing, exploratory data analysis, feature selection, hyperparameter optimization, to model evaluation using Accuracy, Precision, Recall, F1-Score, Confusion Matrix, and ROC Curve, show that the proposed approach produces very superior classification capabilities with an accuracy of 100%, Precision of 100%, Recall of 100%, F1-Score of 100%, and an AUC value of 1,000, accompanied by a Confusion Matrix that does not produce False Positive or False Negative in the test data. In addition to producing very high predictive performance, the Permutation Feature Importance analysis also shows that the attributes serum_creatinine, anemia, diabetes_mellitus, hypertension, blood_urea, and blood_glucose_random are the most influential factors in the CKD identification process, which is in line with clinical indicators that have been recognized in the medical literature. The results of this study prove that the combination of boosting methods with hyperparameter optimization can improve the model's generalization ability, reduce classification errors, and produce a more accurate, stable, and interpretive prediction model compared to conventional classification algorithms or other ensemble methods. Thus, the proposed model has excellent potential to be implemented as a clinical decision support system to assist in the early detection of chronic kidney disease so that it can support faster and more accurate medical decision-making. However, further research is recommended to validate using larger and more diverse datasets from various healthcare institutions, as well as evaluate the model's performance on real-world clinical data so that its generalization ability and readiness for implementation in

healthcare environments can be demonstrated more comprehensively.

REFERENCES

[1] I. Lousa, "New potential biomarkers for chronic kidney disease management — A review of the literature," *Int. J. Mol. Sci.*, vol. 22, no. 1, pp. 1–37, 2021, doi: 10.3390/ijms22010043.

[2] H. Yanai, "Molecular biological and clinical understanding of the pathophysiology and treatments of hyperuricemia and its association with metabolic syndrome, cardiovascular diseases and chronic kidney disease," *Int. J. Mol. Sci.*, vol. 22, no. 17, 2021, doi: 10.3390/ijms22179221.

[3] H. N. Wee, "The Kynurenine Pathway in Acute Kidney Injury and Chronic Kidney Disease," *Am. J. Nephrol.*, vol. 52, no. 10, pp. 771–787, 2021, doi: 10.1159/000519811.

[4] Y. Liang, "Association of MAFLD with Diabetes, Chronic Kidney Disease, and Cardiovascular Disease: A 4.6-Year Cohort Study in China," *J. Clin. Endocrinol. Metab.*, vol. 107, no. 1, pp. 88–97, 2022, doi: 10.1210/clinem/dgab641.

[5] A. Sarraju, "Effects of canagliflozin on cardiovascular, renal, and safety outcomes in participants with type 2 diabetes and chronic kidney disease according to history of heart failure: Results from the CREDENCE trial," *Am. Heart J.*, vol. 233, pp. 141–148, 2021, doi: 10.1016/j.ahj.2020.12.008.

[6] M. Provenzano, "Albuminuria-Lowering Effect of Dapagliflozin, Eplerenone, and their Combination in Patients with Chronic Kidney Disease: A Randomized Cross-over Clinical Trial," *J. Am. Soc. Nephrol.*, vol. 33, no. 8, pp. 1569–1580, 2022, doi: 10.1681/ASN.2022020207.

[7] C. Bosch, "Tirzepatide and prevention of chronic kidney disease," *Clin. Kidney J.*, vol. 16, no. 5, pp. 797–808, 2023, doi: 10.1093/ckj/sfac274.

[8] E. M. Senan, "Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques," *J. Healthc. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/1004767.

[9] O. Gungor, "Effects of hormonal changes on sarcopenia in chronic kidney disease: where are we now and what can we do?," *J. Cachexia Sarcopenia Muscle*, vol. 12, no. 6, pp. 1380–1392, 2021, doi: 10.1002/jcsm.12839.

[10] H. L. Yang, "The Role of Dietary Fiber Supplementation in Regulating Uremic Toxins in Patients With Chronic Kidney Disease: A Meta-Analysis of Randomized Controlled Trials," *J. Ren. Nutr.*, vol. 31, no. 5, pp. 438–447, 2021, doi: 10.1053/j.jrn.2020.11.008.

[11] M. A. Khan, "Promoting Plant-Based Therapies for Chronic Kidney Disease," *J. Evid. Based Integr. Med.*, vol. 27, 2022, doi: 10.1177/2515690X221079688.

[12] H. Ogata, "Effect of Treating Hyperphosphatemia with Lanthanum Carbonate vs Calcium Carbonate on Cardiovascular Events in Patients with Chronic Kidney Disease Undergoing Hemodialysis: The LANDMARK Randomized Clinical Trial," *JAMA J. Am. Med. Assoc.*, vol. 325, no. 19, pp. 1946–1954, 2021, doi: 10.1001/jama.2021.4807.

[13] G. Song, "An ultrasensitive fluorescent breath ammonia sensor for noninvasive diagnosis of chronic kidney disease and helicobacter pylori infection," *Chem. Eng. J.*, vol. 440, 2022, doi: 10.1016/j.cej.2022.135979.

[14] M. F. Yacoub, H. A. Maghawry, N. A. Helal, S. V. Soto, and T. F. Gharib, "An Efficient 2-Stages Classification Model for Students Performance Prediction BT - Proceedings of the 8th International Conference on Advanced Intelligent Systems and Informatics 2022," 2023, pp. 107–122. doi: 10.1007/978-3-031-20601-6_9.

[15] M. F. Yacoub, H. A. Maghawry, N. A. Helal, S. V. Soto, and T. F. Gharib, "An Efficient 2-Stages Classification Model for Students Performance Prediction," in *Proceedings of the 8th International Conference on Advanced Intelligent Systems and Informatics 2022*, 2023, pp. 107–122. doi: 10.1007/978-3-031-20601-6_9.

[16] Z. Cen, "Improving fatigue properties, temperature stability and piezoelectric properties of KNN-based ceramics via sintering in reducing atmosphere," *J. Eur. Ceram. Soc.*, vol. 41, no. 8, pp. 4462–4472, 2021, doi: 10.1016/j.jeurceramsoc.2021.03.007.

[17] T. Gao, "Enhanced comprehensive energy storage properties of lead-free KNN based ceramics through composition optimization strategy," *Chem. Eng. J.*, vol. 500, 2024, doi: 10.1016/j.cej.2024.157560.

[18] F. Yu, "Developing 'smart windows' for optical storage and temperature sensing based on KNN transparent ceramics," *J. Eur. Ceram. Soc.*, vol. 43, no. 10, pp. 4408–4418, 2023, doi: 10.1016/j.jeurceramsoc.2023.03.034.

[19] P. Venkata, "Data Mining and SVM Based Fault Diagnostic Analysis in Modern Power System Using Time and Frequency Series Parameters Calculated From Full-Cycle Moving Window," *J. Oper. Autom. Power Eng.*, vol. 12, no. 3, pp. 206–214, 2024, doi: 10.22098/JOAPE.2023.10819.1789.

[20] A. P. Gopi, "Classification of tweets data based on polarity using improved RBF kernel of SVM," *Int. J. Inf. Technol. Singapore*, vol. 15, no. 2, pp. 965–980, 2023, doi: 10.1007/s41870-019-00409-4.

[21] B. S. Reddy, "An optimal superpixel segmentation based transfer learning using AlexNet–SVM model for weed detection," *Int. J. Syst. Assur. Eng. Manag.*, 2023, doi: 10.1007/s13198-023-02064-2.

[22] J. Sun *et al.*, "Synergistic construction of an annual 2 m mangrove species dataset from 2016 to 2023 using structural features and hybrid stacked deep learning models—Beibu Gulf, Guangxi, China," *ISPRS J. Photogramm. Remote Sens.*, vol. 230, pp. 754–778, 2025, doi: 10.1016/j.isprsjprs.2025.10.007.

[23] K. Chatterjee *et al.*, "PDD-ET: Parkinson's Disease Detection Using ML Ensemble Techniques and Customized Big Dataset," *Inf.*, vol. 14, no. 9, 2023, doi:

10.3390/info14090502.
[24] A. A. R. M. Abu Ebayyeh *et al.*, "Waveguide quality inspection in quantum cascade lasers: A capsule neural network approach," *Expert Syst. Appl.*, vol. 210, 2022, doi: 10.1016/j.eswa.2022.118421.
[25] P. Sadorsky, "A Random Forests Approach to Predicting Clean Energy Stock Prices," *J. Risk Financ. Manag.*, vol. 14, no. 2, 2021, doi: 10.3390/jrfm14020048.
[26] M. J. Jun, "A comparison of a gradient boosting decision tree, random forests, and artificial neural networks to model urban land use changes: the case of the Seoul metropolitan area," *Int. J. Geogr. Inf. Sci.*, vol. 35, no. 11, pp. 2149–2167, 2021, doi: 10.1080/13658816.2021.1887490.
[27] C. Chen, "Welding penetration monitoring for pulsed GTAW using visual sensor based on AAM and random forests," *J. Manuf. Process.*, vol. 63, pp. 152–162, 2021, doi: 10.1016/j.jmapro.2020.04.005.
[28] L. Algorithms, "Advancing Chronic Kidney Disease Prediction: Comparative Analysis of Machine Learning Algorithms and a Hybrid Model," pp. 15–21, doi: 10.32996/jcsts.
[29] E. Stage, "Comparing Machine Learning Techniques for Detecting Chronic Kidney Disease in Early Stage," pp. 20–32, doi: 10.32996/jcsts.
[30] P. Thamilselvan, "Lung Cancer Prediction and Classification Using Adaboost Data Mining Algorithm," *Int. J. Comput. Theory Eng.*, vol. 14, no. 4, pp. 149–154, 2022, doi: 10.7763/IJCTE.2022.V14.1322.
[31] N. Wu, "Cervical cell classification based on strong feature CNN-LSVM network using Adaboost optimization," *J. Intell. Fuzzy Syst.*, vol. 44, no. 3, pp. 4335–4355, 2023, doi: 10.3233/JIFS-221604.
[32] A. E. L. Bilali, "An integrated approach based on Gaussian noises-based data augmentation method and AdaBoost model to predict faecal coliforms in rivers with small dataset," *J. Hydrol.*, vol. 599, 2021, doi: 10.1016/j.jhydrol.2021.126510.
[33] C. W. Teoh, S. B. Ho, K. S. Dollmat, and C. H. Tan, "Ensemble-Learning Techniques for Predicting Student Performance on Video-Based Learning," *Int. J. Inf. Educ. Technol.*, vol. 12, no. 8, pp. 741–745, 2022, doi: 10.18178/ijiet.2022.12.8.1679.
[34] B. M. Kistler, "The International Society of Renal Nutrition and Metabolism Commentary on the National Kidney Foundation and Academy of Nutrition and Dietetics KDOQI Clinical Practice Guideline for Nutrition in Chronic Kidney Disease," *J. Ren. Nutr.*, vol. 31, no. 2, pp. 116–120, 2021, doi: 10.1053/j.jrn.2020.05.002.
[35] A. C. Lai, "A Personalized Approach to Chronic Kidney Disease and Cardiovascular Disease: JACC Review Topic of the Week," *J. Am. Coll. Cardiol.*, vol. 77, no. 11, pp. 1470–1479, 2021, doi: 10.1016/j.jacc.2021.01.028.
[36] M. Rashed-Al-Mahfuz, "Clinically Applicable Machine Learning Approaches to Identify Attributes of Chronic Kidney Disease (CKD) for Use in Low-Cost Diagnostic Screening," *IEEE J. Transl. Eng. Heal. Med.*, vol. 9, pp. 1–11, 2021, doi: 10.1109/JTEHM.2021.3073629.
[37] A. Mima, "Hypoxia-inducible factor-prolyl hydroxylase inhibitors for renal anemia in chronic kidney disease: Advantages and disadvantages," *Eur. J. Pharmacol.*, vol. 912, 2021, doi: 10.1016/j.ejphar.2021.174583.
[38] R. Sawhney, "A comparative assessment of artificial intelligence models used for early prediction and evaluation of chronic kidney disease," *Decis. Anal. J.*, vol. 6, 2023, doi: 10.1016/j.dajour.2023.100169.
[39] S. M. Kim, "Association of Chronic Kidney Disease With Atrial Fibrillation in the General Adult Population: A Nationwide Population-Based Study," *J. Am. Heart Assoc.*, vol. 12, no. 8, 2023, doi: 10.1161/JAHA.122.028496.