

MACHINE LEARNING-BASED EARLY PREDICTION OF KIDNEY FAILURE: A COMPARATIVE STUDY OF ARTIFICIAL NEURAL NETWORK AND RANDOM FOREST MODELS

Nur Mahmudah^{(1)*}, Alif Yuanita Kartini⁽²⁾, Muhammad Anshori⁽³⁾

^{1,2}Department of Statistics, Universitas Nahdlatul Ulama Sunan Giri, Indonesia

³Departement of Mechanical Engineering, Universitas Nahdlatul Ulama Sunan Giri, Indonesia

*Corresponding Author. E-mail: mudah15@gmail.com

ARTICLE INFO

Article History:

Received: 11-Sep. 2025

Revised: 16-Oct. 2025

Accepted: 21-Oct. 2025

Keywords:

Artificial Neural Network, Kidney Failure, Machine Learning, Prediction, Random Forest

ABSTRACT

Penyakit ginjal kronis merupakan kondisi progresif dengan angka kejadian yang terus meningkat di Indonesia, termasuk di Kabupaten Bojonegoro. Penyakit ini disebabkan oleh penurunan fungsi ginjal secara bertahap yang mengakibatkan akumulasi zat sisa metabolik dan racun dalam tubuh. Deteksi dini sangat penting untuk mencegah komplikasi serta meningkatkan efektivitas pengobatan. Namun, metode diagnostik saat ini masih bergantung pada pemeriksaan laboratorium dan anamnesis medis yang belum memberikan akurasi optimal. Penelitian ini bertujuan mengembangkan model prediksi gagal ginjal menggunakan algoritma *Artificial Neural Network* (ANN) dan *Random Forest* (RF). ANN memiliki keunggulan dalam mengenali pola nonlinier yang kompleks, sedangkan RF menunjukkan ketangguhan dalam tugas klasifikasi. Hasil penelitian menunjukkan bahwa model ANN mencapai akurasi sebesar 81%, dengan variabel dominan berupa *Blood Urea Nitrogen* (BUN), tekanan darah, kreatinin, dan diabetes. Sementara itu, model RF menghasilkan akurasi 80%, dengan tekanan darah sebagai variabel paling berpengaruh. Berdasarkan kinerja komparatif, ANN dipilih sebagai model optimal untuk prediksi gagal ginjal. Meskipun demikian, integrasi kedua algoritma dalam kerangka hibrida berpotensi meningkatkan cakupan prediksi serta mendukung deteksi dini dan pengambilan keputusan klinis dalam penanganan gagal ginjal.

Chronic kidney disease (CKD) is a progressive condition with an increasing incidence rate in Indonesia, including in Bojonegoro Regency. This disease results from a gradual decline in kidney function, leading to the accumulation of metabolic waste and toxins in the body. Early detection is crucial to prevent complications and enhance treatment effectiveness. However, current diagnostic methods rely heavily on laboratory tests and medical anamnesis, which may not provide sufficient accuracy. This study aims to develop a predictive model for kidney failure using *Artificial Neural Network* (ANN) and *Random Forest* (RF) algorithms. The ANN is particularly effective in recognizing complex nonlinear patterns, while RF demonstrates robustness in classification tasks. The results show that the ANN model achieved an accuracy of 81%, with blood urea nitrogen (BUN), blood pressure, creatinine, and diabetes identified as dominant predictors. Meanwhile, the RF model achieved an accuracy of 80%, with blood pressure emerging as the most influential variable. Based on comparative performance, ANN was selected as the optimal model for the kidney failure prediction. Nonetheless, integrating both algorithms into a hybrid



framework could further enhance predictive coverage and support early detection and clinical decision-making in kidney failure management.

This is an open access article under the CC-BY-NC-SA license



How to Cite: Mahmudah, N., Kartini, A. Y., & Anshori, M. (2025). Machine Learning-Based Early Prediction of Kidney Failure: A Comparative Study of Artificial Neural Network and Random Forest Models. *Journal of Mathematics Education and Science*, 8(2), 202-215. <https://doi.org/10.32665/james.v8i2.5598>

INTRODUCTION

Kidney failure is a chronic disease with a steadily increasing prevalence worldwide, including in Indonesia (Ardiantito et al., 2023). In Bojonegoro, the number of kidney failure patients continues to rise each year, second only to diabetes, with many cases diagnosed at advanced stages (Cahyani & Kartini, 2022). This condition occurs when the kidneys lose their filtration function, resulting in the accumulation of metabolic waste and excess fluids, which may lead to severe complications and death if left untreated (Amalia, 2018). Early detection is therefore essential to prevent further complications and optimize treatment outcomes. However, current diagnostic approaches, which mainly rely on laboratory testing and clinical anamnesis, are often insufficient for identifying early-stage risks (Khamidah et al., 2018). To address this gap, predictive models are needed to improve the accuracy and timeliness of kidney failure detection. In this regard, machine learning offers a promising approach to strengthen predictive capability and support medical decision-making in kidney failure management (Han et al., 2019).

RSUD Dr. R. Sosodoro Djatikoesoemo Regional General Hospital serves as the primary referral center for kidney failure cases in Bojonegoro Regency (Poonia et al., 2022). Early detection of this disease remains a major challenge due to the limitations of current screening approaches (Chittora et al., 2021). Diagnosis still largely relies on laboratory tests such as creatinine measurement (Singh et al., 2022). Which often detect the disease only after substantial renal impairment has occurred (Shaikhina et al., 2019). This underscores the urgent need for predictive approaches that enable earlier identification of at-risk patients (Arifin & Ariesta, 2019). Artificial Intelligence (AI) has shown remarkable potential in healthcare applications, particularly in disease prediction (Boyang et al., 2022). Among various AI techniques, the ANN is capable of modeling complex nonlinear relationships between clinical features, allowing it to capture subtle patterns that precede clinical manifestations of kidney disease. Meanwhile, the RF algorithm provides robust feature selection and reduces overfitting through ensemble learning, enhancing model stability and interpretability. The complementary strengths of ANN and RF make them highly effective for early detection tasks, where both sensitivity and reliability are essential. To address existing limitations in predictive modeling, this study proposes an integrated early detection model for kidney failure that combines ANN and RF. The methodological novelty of this research lies in the comparative evaluation and hybrid optimization of both algorithms using clinical datasets from Bojonegoro. By systematically analyzing their performance in terms of accuracy, interpretability, and computational efficiency, the study provides new empirical evidence for model selection in early kidney disease prediction. Furthermore, the findings are expected to guide local health authorities in developing evidence-based strategies and mobile-based diagnostic tools for wider community implementation (Ardiantito et al., 2023).

In the context of kidney disease classification, selecting an appropriate algorithm presents a considerable challenge (Billsus & Pazzani, 1999). The characteristics of medical datasets—such as limited sample size, class imbalance, and complex feature structures—significantly influence the performance of classification algorithms (Barua et al., 2024). Therefore, a comparative analysis is required to determine the most suitable algorithm for classifying medical data related

to kidney disease. This study focuses on comparing the performance of the Random Forest and Artificial Neural Network algorithms (Han et al., 2019). The comparison aims to evaluate the effectiveness and efficiency of each algorithm based on multiple evaluation metrics, thereby identifying the method that best aligns with the dataset characteristics and analytical objectives. Accordingly, this research, entitled “Machine Learning-Based Early Prediction of Kidney Failure: A Comparative Study of Artificial Neural Network and Random Forest Models”, seeks to assess and compare the performance of both algorithms in the context of medical data mining for kidney disease classification.

Machine learning techniques such as Random Forest (RF) and Artificial Neural Networks (ANN) have become widely adopted in medical data analysis due to their robustness and predictive accuracy (Hartono et al., 2023). Random Forest, an ensemble learning algorithm, enhances stability and performance by combining multiple decision trees trained on bootstrapped data samples with random feature selection at each node (Rahimi et al., 2023). This design minimizes overfitting and improves generalization, making RF effective for both classification and regression tasks. Through aggregation—majority voting for classification and averaging for regression—RF produces reliable predictions even in noisy and high-dimensional datasets (Singh et al., 2022).

In contrast, Artificial Neural Networks (ANN) are computational models inspired by the structure of biological neural systems (Almansour et al., 2019). By employing interconnected processing units and activation functions, ANN dynamically learn nonlinear relationships between input and output variables, enabling them to capture complex patterns in medical and biological data (Poonia et al., 2022). Common activation functions such as the binary and bipolar sigmoid govern the neuron’s response within hidden layers and contribute to the network’s adaptive learning process. Together, RF and ANN represent complementary approaches—RF offering interpretability and resilience to noise, and ANN providing strong nonlinear modeling capabilities—making their comparative analysis valuable in developing accurate and generalizable predictive models for health-related applications.

METHOD

Data Source. This study employed secondary data obtained from the medical records of kidney failure patients at Dr. R. Sosodoro Djatikoesoemo Regional Hospital, Bojonegoro. The dataset covers the period from January 2024 to February 2025. Two types of variables were considered: the response variable and predictor variables. The response variable, kidney failure status (Y), was determined based on the Glomerular Filtration Rate (GFR) values. The predictor variables in this study consisted of nine observed variables. Detailed information on all variables is summarized in Table 1 (Ardiantito et al., 2023).

Table 1. Research Variables

No	Variables	Description	Coding / Category
1	Kidney Failure (Y)	Response variable determined by GFR values	1 = Yes (GFR < 15) 0 = No (GFR > 15)
2	Gender (X ₁)	Patient sex	0 = Female 1 = Male

No	Variables	Description	Coding / Category
3	Age (X_2)	Age category	1 = Infant (0–5 years) 2 = Child (6–11 years) 3 = Early adolescent (12–16 years) 4 = Late adolescent (17–25 years) 5 = Early adult (26–35 years) 6 = Late adult (36–45 years) 7 = Early elderly (46–55 years) 8 = Late elderly (56–65 years) 9 = Senior (> 65 years)
4	Blood Urea Nitrogen / BUN (X_3)	BUN level	1 = Low (< 6 mg/dL) 2 = Normal (6–20 mg/dL) 3 = High (> 20 mg/dL)
5	Creatinine (X_4)	Serum creatinine level	1 = Low (M > 0.7 mg/dL, F > 0.6 mg/dL) 2 = Normal (M: 0.7–1.3 mg/dL, F: 0.6–1.1 mg/dL) 3 = High (M > 1.3 mg/dL, F > 1.1 mg/dL)
6	Haemoglobin (X_5)	Hb level	1 = Low (M < 13 g/dL, F < 12 g/dL) 2 = Normal (M: 13–18 g/dL, F: 12–16 g/dL) 3 = High (M > 18 g/dL, F > 16 g/dL)
7	Blood Pressure (X_6)	Blood pressure category	1 = Hypotension (< 90/60 mmHg) 2 = Normal (90–119/60–79 mmHg) 3 = Prehypertension (120–139/80–89 mmHg) 4 = Hypertension stage I (140–159/90–99 mmHg) 5 = Hypertension stage II ($\geq$ 160/100 mmHg) 6 = Hypertensive crisis (> 180/120 mmHg)
8	Anemia (X_7)	Presence of anemia during treatment	1 = Yes 0 = No
9	Diabetes (X_8)	Presence of diabetes during treatment	1 = Yes 0 = No
10	Hypertension (X_9)	Presence of hypertension during treatment	1 = Yes 0 = No

The development of the kidney failure prediction model using Artificial Neural Network (ANN) and Random Forest (RF) at Dr. R. Sosodoro Djatikoesoemo Regional Hospital, Bojonegoro, was conducted through the following steps:

1. Prediction of Kidney Failure Using Artificial Neural Networks and Random Forest Implemented with Python.
2. Collecting data on kidney failure patients at Dr. R. Sosodoro Djatikoesoemo Regional General Hospital, Bojonegoro.
3. Describe kidney failure patients' characteristics at Dr. R. Sosodoro Djatikoesoemo Regional General Hospital, Bojonegoro, based on the observed variables through descriptive statistical analysis and crosstabulation (data visualization) (Paper, 2013).

4. Data preprocessing is conducted through data cleaning and normalization, followed by feature selection using permutation importance, to enhance model performance, reduce complexity, and prevent overfitting (Dörpinghaus et al., 2022).
5. Splitting the dataset into two subsets, with 80% allocated for training and 20% for testing, ensures that the model is effectively trained on the training data. In contrast, the testing data is used to evaluate its predictive capability on previously unseen instances (Colin Campbell, 2011).
6. Checking for data imbalance and addressing it using the SMOTE technique if an imbalance is detected.
7. Determining kidney failure prediction using the Random Forest algorithm (Barua et al., 2024): Split the dataset into training and testing sets. Ensure data quality through preprocessing, such as handling missing values or converting categorical features into numerical form when necessary.
8. The first step in constructing a Random Forest is bootstrapping, which involves drawing random subsets from the training data with replacement, meaning that some observations may be selected multiple times while others may not be selected at all. Each decision tree is trained on these bootstrapped samples, thereby increasing diversity among trees and reducing the risk of overfitting. For each decision tree, the splitting of data is carried out using a randomly selected subset of the available features. By incorporating randomness at both the data and feature levels, this procedure enhances the heterogeneity across trees and, consequently, improves the robustness of the ensemble model. The splitting process is typically determined by well-established criteria, such as Gini impurity or Entropy in classification tasks, and Mean Squared Error (MSE) in regression tasks.
 - a. Gini Impurity
Gini impurity quantifies the probability that a randomly selected element from the dataset would be incorrectly classified if it were assigned according to the distribution of labels at a given node. A lower Gini value indicates a better separation achieved at that node.

$$Gini(t) = 1 - \sum_{i=1}^k p_i^2 \quad (1)$$
 - b. Entropy
Entropy is a measure of uncertainty or randomness within a node. A lower entropy value indicates greater homogeneity of the data contained in that node.

$$Entropy(t) = - \sum_{i=1}^k p_i \log_2(p_i) \quad (2)$$
 - c. Mean Squared Error (MSE)
Random Forest employs Mean Squared Error (MSE) as a splitting criterion in regression tasks. MSE quantifies the average squared difference between the predicted values generated by the model and the actual observed values.

$$MSE(t) = \frac{1}{n_t} \sum_{i=1}^{n_t} (y_i - \hat{y}_i)^2 \quad (3)$$

Once all trees have been trained, the final prediction is obtained through an aggregation process. In classification tasks, the outcome is determined by majority voting across the ensemble, where each tree contributes a class prediction, and the class most frequently selected is assigned as the final output. In regression tasks, the final prediction is computed as the arithmetic mean of the individual predictions produced by the constituent trees.
 - d. Model evaluation is conducted using established performance metrics. For classification tasks, measures such as accuracy, precision, recall, and the F1-score are employed to assess predictive capability.

- e. Cross-validation facilitates the determination of optimal parameter values, thereby reducing the risk of overfitting or underfitting and enhancing the model's generalization performance.
9. Determining kidney failure prediction using the Artificial Neural Network (ANN) algorithm Network (Chittora et al., 2021):
 - a. Determining kidney failure prediction using an Artificial Neural Network (ANN) through model initialization, consisting of: an Input Layer (receiving medical data as input, e.g., blood pressure and creatinine levels), Hidden Layers (processing information through weights and activation functions), and an Output Layer (producing a binary output, where 1 = kidney failure and 0 = healthy).
 - b. Specifying the activation function by assigning weights to the layers, as illustrated in the figure below.

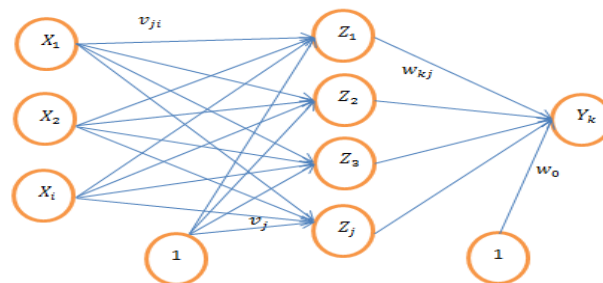


Figure 1. illustrates the architecture of the Artificial Neural Network

Figure 1 depicts a Neural Network composed of multiple layers that establish connection patterns within and across layers. Each layer consists of simple processing elements referred to as neurons. The network comprises an input layer with n neurons, a hidden layer with p units, and an output layer with q neurons. The resulting model can be denoted as follows:

$$Y_k = \psi_k + \left\{ w_k + \sum_{j=1}^p w_{kj} \psi_n \left(\sum_{i=1}^n v_{ji} X_i \right) \right\} \quad (4)$$

10. Performing model evaluation using K-fold Cross-Validation.
11. Evaluating the model by calculating the accuracy value as presented below.
12. Prediction Output, providing a comprehensive framework for developing an interpretable and accurate machine learning-based early prediction system for kidney failure.

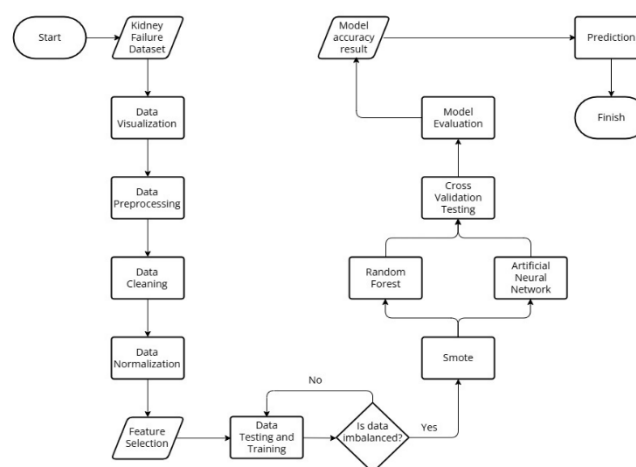


Figure 2. Flowchart of the Kidney Failure Prediction Model Based on ANN and RF

RESULTS

This study utilized secondary data from the medical records of kidney failure patients at Dr. R. Sosodoro Djatikoesoemo Regional General Hospital, Bojonegoro, covering the period from January 2024 to February 2025, with 172 patient records. The dataset comprises one response variable, namely kidney failure status (Y), and nine predictor variables: gender (X_1), age (X_2), blood urea nitrogen (BUN) level (X_3), creatinine level (X_4), Haemoglobin (X_5), blood pressure (X_6), anemia (X_7), diabetes (X_8), and hypertension (X_9). Before the model training process, a descriptive statistical analysis was conducted as the initial step. The following presents an overview of the research variables employed:

Table 2. Descriptive Statistics

Variables	Category	Frequency	Percentage
Kidney Failure (Y)	0 = No (GFR > 15)	33	19,2%
	1 = Yes (GFR < 15)	139	80,8%
Gender (X_1)	0 = Female	85	49,4%
	1 = Male	87	50,6%
Age (X_2)	1 = Infant (0–5 years)	0	0%
	2 = Child (6–11 years)	0	0%
	3 = Early adolescent (12–16 years)	0	0%
	4 = Late adolescent (17–25 years)	2	1,2%
	5 = Early adult (26–35 years)	13	7,6%
	6 = Late adult (36–45 years)	38	22,1%
	7 = Early elderly (46–55 years)	41	23,8%
	8 = Late elderly (56–65 years)	50	29,1%
	9 = Senior (> 65 years)	28	16,3%
Blood Urea Nitrogen / BUN (X_3)	1 = Low (< 6 mg/dL)	0	0
	2 = Normal (6–20 mg/dL)	13	7,6%
	3 = High (> 20 mg/dL)	159	92,4%
Creatinine (X_4)	1 = Low (M > 0.7 mg/dL, F > 0.6 mg/dL)	0	0
	2 = Normal (M: 0.7–1.3 mg/dL, F: 0.6–1.1 mg/dL)	7	4,1%
	3 = High (M > 1.3 mg/dL, F > 1.1 mg/dL)	165	95,9%
Haemoglobin (X_5)	1 = Low (M < 13 g/dL, F < 12 g/dL)	159	92,4%
	2 = Normal (M: 13–18 g/dL, F: 12–16 g/dL)	13	7,6%

Variables	Category	Frequency	Percentage
Blood Pressure (X_6)	3 = High (M > 18 g/dL, F > 16 g/dL)	0	0
	1 = Hypotension (< 90/60 mmHg)	0	0%
	2 = Normal (90–119/60–79 mmHg)	23	13,4%
	3 = Prehypertension (120–139/80–89 mmHg)	36	20,9%
	4 = Hypertension stage I (140–159/90–99 mmHg)	31	18,%
	5 = Hypertension stage II ($\geq$ 160/100 mmHg)	31	18%
Anemia (X_7)	6 = Hypertensive crisis (> 180/120 mmHg)	51	29,7%
	0 = No	100	58,1%
Diabetes (X_8)	1 = Yes	72	41,9%
	0 = No	149	86,6%
Hypertension (X_9)	1 = Yes	23	13,4%
	0 = No	147	85,5%
	1 = Yes	25	14,5%

Table 2 presents the dataset comprised of 172 patients diagnosed and treated at RSUD Dr. R. Sosodoro Djatikoesoemo Bojonegoro between January 2024 and February 2025. Among these, the majority (80.8%) were classified as having kidney failure, whereas only 19.2% did not present with kidney failure ($GFR > 15$). The gender distribution was nearly balanced, with 50.6% male and 49.4% female patients. Regarding age distribution, the largest subgroup was late elderly (56–65 years; 29.1%), followed by early elderly (46–55 years; 23.8%) and late adults (36–45 years; 22.1%). Senior patients aged over 65 years accounted for 16.3%, while smaller proportions were early adults (7.6%) and late adolescents (1.2%). Notably, no cases were recorded among infants, children, or early adolescents. Biochemical parameters demonstrated pronounced abnormalities. Most patients (92.4%) exhibited elevated blood urea nitrogen ($BUN > 20$ mg/dL), and 95.9% showed elevated creatinine levels ($M > 1.3$ mg/dL; $F > 1.1$ mg/dL), reflecting impaired renal function. Similarly, 92.4% presented with low Haemoglobin concentrations, indicating a high prevalence of anemia. Blood pressure analysis revealed that nearly one-third of patients were in hypertensive crisis ($>180/120$ mmHg; 29.7%), while additional cases were classified as prehypertension (20.9%), hypertension stage I (18.0%), and hypertension stage II (18.0%). Only 13.4% of patients exhibited normal blood pressure, and no hypotension were observed. In terms of comorbidities, 41.9% of patients had anemia, 13.4% had diabetes, and 14.5% had hypertension, underscoring the clinical complexity frequently associated with kidney failure. Collectively, these findings indicate that the study population was predominantly composed of elderly individuals with severe renal impairment, high anemia prevalence, and advanced hypertension—factors consistent with established risk profiles for chronic kidney disease progression.

The second stage of the analysis involved assessing correlations to characterize patient distribution and explore the interrelationships among the observed variables. Specifically, the

analysis examined the association between numerical predictors and the target variable (Y), providing preliminary insights into the relationships between clinical features and kidney failure status. The correlation heatmap revealed that diabetes exhibited the strongest positive correlation with kidney failure ($r = 0.52$), indicating that patients with a history of diabetes were more likely to develop the condition. Furthermore, anemia and Hemoglobin also demonstrated substantial associations. Anemia showed a moderate positive correlation ($r = 0.43$), while Haemoglobin displayed a negative correlation ($r = -0.44$), suggesting that lower Hemoglobin levels were strongly associated with an increased likelihood of kidney failure.

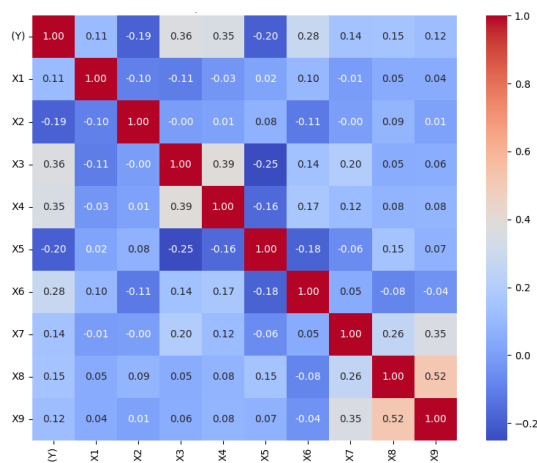


Figure 3. Correlation Heatmap of the Variables

Based on Figure 3, the findings reveal that blood pressure (X3), BUN (X4), and anemia (X6) demonstrate moderate positive correlations with kidney failure, indicating their potential role as dominant contributing factors. Conversely, creatinine (X5) exhibits a negative correlation, suggesting a potential protective pattern within this dataset. Moreover, a strong positive association was identified between diabetes (X8) and hypertension (X9), reflecting their well-established clinical interrelationship.

Table 3. Classification Accuracy Results for Kidney Failure

Random Forest			Artificial Neural Network		
Predicted Class	Actual Class		Predicted Class	Actual Class	
	Positive	Negative		Positive	Negative
Positive	6	4	Positive	5	2
Negative	6	36	Negative	7	21
Accuracy Value	80%		Accuracy Value	81%	

Table 3 summarizes the classification performance of the Random Forest (RF) and Artificial Neural Network (ANN) models in predicting kidney failure. The RF achieved 80% accuracy, while the ANN slightly outperformed with 81%. Despite comparable results, both models showed limitations in correctly identifying positive cases, highlighting the need for further optimization to reduce false negatives in clinical applications.

Table 4. Comparison of Feature Importance Across Classification Models

No	Random Forest		ANN	
	Feature	Importance	Feature	Importance
1	Blood Pressure	0.417	Blood Urea Nitrogen / BUN	0.09
2	Age	0.188	Blood Pressure	0.04
3	Anemia	0.092	Creatinine	0.03
4	Gender	0.082	Diabetes	0.018
5	Blood Urea Nitrogen / BUN	0.082	Age	0.005
6	Diabetes	0.042	Gender	0.000
7	Creatinine	0.041	Anemia	0.000
8	Hypertension	0.034	Hemoglobin	-0.003
9	Hemoglobin	0.022	Hypertension	-0.040

The analysis in Table 4 demonstrates that Random Forest identifies clinical factors as dominant predictors of kidney failure. Blood pressure showed the highest importance (0.417), indicating hypertension as the principal risk factor, followed by age (0.188), anemia (0.092), and sex (0.082). Laboratory indicators such as BUN (0.082) and creatinine (0.041) were less influential but remained clinically relevant. With an accuracy of 80%, Random Forest underscores the central role of clinical characteristics in diagnosis and early screening. Conversely, the Artificial Neural Network (ANN) model prioritized laboratory variables. BUN was the strongest predictor (0.09), followed by blood pressure (0.04) and creatinine (0.03), while diabetes (0.018) and age (0.005) had minor contributions. Anemia and sex showed no effect, and Hemoglobin (-0.003) and hypertension (-0.04) exhibited negative influences. With a slightly higher accuracy of 81%, ANN highlights the significance of laboratory parameters for classification. Comparative analysis indicates that the RF and ANN models utilize an identical set of predictor and response variables. The RF model demonstrates greater interpretability and places more emphasis on clinical features, whereas the ANN effectively captures complex nonlinear relationships among laboratory variables. This comparison underscores the complementary strengths of both algorithms in modeling kidney failure prediction. Rather than integrating the two methods, this study focuses on a systematic comparative evaluation to determine the most effective algorithm for early detection and clinical decision support in kidney failure management. Overall, the ANN model exhibited superior predictive performance, achieving the highest accuracy (81%) through the effective utilization of blood urea nitrogen (BUN) and creatinine variables, thereby enabling more precise classification and supporting earlier clinical intervention.

DISCUSSION

A comparative evaluation was conducted to assess the predictive performance of Random Forest (RF) and Artificial Neural Network (ANN) models in classifying kidney disease (Barua et al., 2024). The results of the model evaluation are presented in Table 5.

Table 5. Comparative Analysis of Random Forest (RF) and Artificial Neural Network (ANN) Models

Aspect	Random Forest (RF)	Artificial Neural Network (ANN)
Accuracy	80%	81%
Strengths	High interpretability and robustness; reduces overfitting through ensemble learning (Breiman, 2001).	Excellent in modeling nonlinear and complex feature interactions; adaptive

Key Predictors	Blood pressure and age identified as most influential variables (Shen et al., 2023)	learning structure (Rahimi et al., 2023) Greater emphasis on biochemical indicators such as BUN and serum creatinine (Poonia et al., 2022)
Interpretability	Provides variable importance measures offering transparent clinical insights.	Limited interpretability due to black-box nature of neural models.
Limitations	Performance decreases under class imbalance (Yaqin et al., 2025)	Requires parameter tuning and large datasets to avoid overfitting (Budiani & Mahmudah, 2025)
Clinical Implications	Highlights clinical predictors such as blood pressure, age, and anemia status for CKD progression (Darwanto et al., 2021).	Captures subtle biochemical patterns relevant to renal dysfunction for early prediction (Wijayanti et al., 2018).
Future Direction	Potential for integration with ANN to enhance nonlinear modelling (Ardiantito et al., 2023)	Potential hybridization with RF to improve interpretability (Hartono et al., 2023)

The findings indicate that RF and ANN exhibit complementary strengths in early kidney disease prediction. RF excels in interpretability, offering clearer insights into clinical risk factors such as blood pressure, anemia, and age. Conversely, ANN performs better in detecting intricate biochemical patterns, particularly in laboratory-based parameters such as BUN and creatinine. These results suggest that a hybrid approach integrating both RF and ANN could provide a more accurate and interpretable early kidney failure detection framework, particularly in medical settings with diverse data characteristics or limited diagnostic resources.

ANN prioritizes laboratory parameters, with BUN (importance = 0.09) and creatinine (0.03) being the most influential predictors, while clinical variables such as age and diabetes contribute minimally, and Hemoglobin and hypertension exert negative effects. This pattern is consistent with evidence that BUN and creatinine are key biomarkers for detecting renal function decline (Hasanah et al., 2023). Parallel international studies on predictive models, including explainable ensemble-tree approaches, also highlight laboratory markers and hypertension as critical features in early CKD detection. Accuracy comparisons reveal that ANN slightly outperforms RF (81% vs. 80%), reflecting its superior capacity to model non-linear numerical data from laboratory tests. Nonetheless, integrating RF and ANN may enhance diagnostic accuracy: RF enables rapid detection based on clinical conditions, whereas ANN validates predictions through laboratory indicators. Such a hybrid approach resonates with current trends in healthcare machine learning, which emphasize both interpretability and high accuracy. Theoretically, laboratory data offer greater diagnostic precision, while clinical factors provide practical utility for screening—making their integration an ideal strategy for early kidney failure detection.

CONCLUSION

Based on the findings presented, it can be concluded that (1) Based on the analysis of 172 medical records of kidney failure patients at RSUD Dr. R. Sosodoro Djatikoesoemo Bojonegoro from

January 2024 to February 2025, most patients were older adults, particularly in the late elderly group (56–65 years). The dominant clinical characteristics included elevated BUN and creatinine levels, low Hemoglobin, and abnormal blood pressure, with nearly one-third experiencing hypertensive crisis. Comorbidities such as anemia, diabetes, and hypertension further worsened patient conditions and must be considered in clinical management. (2) Correlation analysis indicated that diabetes had the strongest association with kidney failure, followed by anemia and Haemoglobin, while blood pressure and BUN also played significant roles. These findings confirm that clinical and laboratory factors are essential in detecting kidney failure risk. (3) Classification results demonstrated that Random Forest emphasized clinical factors with 80% accuracy, while Artificial Neural Network (ANN) focused on laboratory factors with 81% accuracy. Since ANN showed the highest accuracy, it is considered the best-performing model. Nevertheless, integrating both models in a hybrid approach may enhance predictive comprehensiveness and support early detection and medical decision-making in kidney failure management.

ACKNOWLEDMENT

The authors would like to express their gratitude to Dr. R. Sosodoro Djatikoesoemo, Regional General Hospital, Bojonegoro, for granting permission to use the research data. Sincere thanks are also extended to the Institute for Research and Community Service (LPPM), Nahdlatul Ulama University of Sunan Giri, through contract number 227/SKt/LPPM/071088/III/2025, for the funding support provided. The authors also acknowledge all parties who have offered guidance, support, and valuable contributions throughout the research process.

AI ACKNOWLEDMENT

The authors declare that generative AI or AI-assisted technologies were not used in any way to prepare, write, or complete this manuscript. The authors confirm that they are the sole authors of this article and take full responsibility for the content therein, as outlined in COPE recommendations.

INFORMED CONSENT

The authors have obtained informed consent from all participants.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCE

- Almansour, N. A., Syed, H. F., Khayat, N. R., Altheeb, R. K., Juri, R. E., Alhiyafi, J., Alrashed, S., & Olatunji, S. O. (2019). Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in Biology and Medicine*, 109(April), 101–111. <https://doi.org/10.1016/j.compbimed.2019.04.017>
- Amalia, H. (2018). Perbandingan Metode Data Mining Svm Dan Nn Untuk Klasifikasi Penyakit Ginjal Kronis. *Maret*, 14(1), 1.
- Ardiantito, W. S., Ramadhan, R. A., Steven Immanuel, R. S., William Iskandar Ps, J. V, Baru, K., Percut Sei Tuan, K., Deli Serdang, K., Utara, S., & penulis, K. (2023). Komparasi Algoritma Machine Learning dalam Memprediksi Penyakit Gagal Ginjal * Wahyu Ardiantito S. *Jurnal Penelitian Dan Karya Ilmiah*, 1(Desember), 363–374.
- Arifin, T., & Ariesta, D. (2019). Prediksi Penyakit Ginjal Kronis Menggunakan Algoritma Naive Bayes Classifier Berbasis Particle Swarm Optimization. *Jurnal Tekno Insentif*, 13(1), 26–30. <https://doi.org/10.36787/jti.v13i1.97>
- Barua, T., Hiran, K. K., Jain, R. K., & Doshi, R. (2024). Machine Learning with Python. In *Machine*

- Learning with Python*. <https://doi.org/10.1515/9783110697186>
- Billsus, D., & Pazzani, M. J. (1999). A hybrid user model for news story classification. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 407, 99–108. https://doi.org/10.1007/978-3-7091-2490-1_10
- Boyang, C., Yuexing, L., Yiping, Y., Haiyang, Y., Xufei, Z., Liancheng, G., & Yunzhi, C. (2022). Construction and analysis of heart failure diagnosis model based on random forest and artificial neural network. *Medicine (United States)*, 101(41), E31097. <https://doi.org/10.1097/MD.00000000000031097>
- Budiani, J. R., & Mahmudah, N. (2025). Comparison of Supervised Machine Learning Algorithms in Heart Failure Disease. *Barekeng*, 19(4), 2739–2750. <https://doi.org/10.30598/barekengvol19iss4pp2739-2750>
- Cahyani, N., & Kartini, A. Y. (2022). DLNN dan BPNN-GA Pada Prediksi Penyakit Diabetes di Bojonegoro. *Journal of Mathematics Education and Science*, 6(1), 1–9. <https://doi.org/10.32665/james.v6i1.416>
- Chittora, P., Chaurasia, S., Chakrabarti, P., Kumawat, G., Chakrabarti, T., Leonowicz, Z., Jasinski, M., Jasinski, L., Gono, R., Jasinska, E., & Bolshev, V. (2021). Prediction of Chronic Kidney Disease - A Machine Learning Perspective. *IEEE Access*, 9, 17312–17334. <https://doi.org/10.1109/ACCESS.2021.3053763>
- Colin Campbell, Y. Y. (2011). *Learning with Support Vector Machines (Synthesis Lectures on Artificial Intelligence and Machine Learning)* (p. 96).
- Darwanto, A. R. S., Taza Luzia Viarindita, & Yekti Widyaningsih. (2021). Analisis Regresi Logistik Binomial dan Algoritma Random Forest pada Proses Pengklasifikasian Penyakit Ginjal Kronis. *Jurnal Statistika Dan Aplikasinya*, 5(1), 1–14. <https://doi.org/10.21009/JSA.05101>
- Dörpinghaus, J., Düing, C., & Stefan, A. (2022). Biomedical Knowledge Graphs: Context, Queries and Complexity. In *Studies in Big Data* (Vol. 112, pp. 529–567). https://doi.org/10.1007/978-3-031-08411-9_20
- Han, X., Zheng, X., Wang, Y., Sun, X., Xiao, Y., Tang, Y., & Qin, W. (2019). Random forest can accurately predict the development of end-stage renal disease in immunoglobulin a nephropathy patients. *Annals of Translational Medicine*, 7(11), 234–234. <https://doi.org/10.21037/atm.2018.12.11>
- Hartono, A., Aska Dewi, L., Yuniarti, E., Tahta Hirani Putri, S., Surya Harahap, T., & Hartono, A. (2023). Machine Learning Classification for Detecting Heart Disease with K-NN Algorithm, Decision Tree and Random Forest. *Eksakta : Berkala Ilmiah Bidang MIPA*, 24(04), 513–522.
- Khamidah, F. S. N., Hapsari, D. P., & Nugroho, H. (2018). Implementasi Fuzzy Decision Tree Untuk Prediksi Gagal Ginjal Kronis. *INTEGER: Journal of Information Technology*, 3(1), 19–28. <https://doi.org/10.31284/j.integer.2018.v3i1.155>
- Paper, D. (2013). R Consistent data analysis. *R Package*, 53.
- Poonia, R. C., Gupta, M. K., Abunadi, I., Albraikan, A. A., Al-Wesabi, F. N., Hamza, M. A., & Tulasi, B. (2022). Intelligent Diagnostic Prediction and Classification Models for Detection of Kidney Disease. *Healthcare (Switzerland)*, 10(2). <https://doi.org/10.3390/healthcare10020371>
- Rahimi, M., Akbari, A., Asadi, F., & Emami, H. (2023). Cervical cancer survival prediction by machine learning algorithms: a systematic review. *BMC Cancer*, 23(1), 1–10. <https://doi.org/10.1186/s12885-023-10808-3>
- Shaikhina, T., Lowe, D., Daga, S., Briggs, D., Higgins, R., & Khovanova, N. (2019). Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation. *Biomedical Signal Processing and Control*, 52, 456–462. <https://doi.org/10.1016/j.bspc.2017.01.012>
- Shen, J., Li, J., Mao, Z., & Zhang, Y. (2023). First-principle study on the stability of Cd passivates in

- soil. *Scientific Reports*, 13(1), 1–8. <https://doi.org/10.1038/s41598-023-31460-8>
- Singh, V., Asari, V. K., & Rajasekaran, R. (2022). A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease. *Diagnostics*, 12(1), 1–22. <https://doi.org/10.3390/diagnostics12010116>
- Wijayanti, R. A., Furqon, M. T., & Adinugroho, S. (2018). Penerapan Algoritme Support Vector Machine Terhadap Klasifikasi Tingkat Risiko Pasien Gagal Ginjal. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(10), 3500–3507.
- Yaqin, A. A., Barata, M. A., & Mahmudah, N. (2025). *Implementation of the Random Forest Algorithm with Optuna Optimization in Lung Cancer Classification*. 14, 561–569.