Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2459−2466
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9392

Comparative Analysis of K-NN and Naïve Bayes Algorithms for Early-Stage Chronic Kidney Disease Classification

Intan Dwi Rahma[1,*], Mhd Furqan[2], Budi Triandi[2]

[1] Faculty of Engineering and Computer Science, Computer Science Study Program, Universitas Potensi Utama, Medan, Indonesia
[2] Faculty of Science and Technology, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia
Email: [1,*]dwirahmaintan@gmail.com, [2]mfurqan@uinsu.ac.id, [3]buditriandi@gmail.com
Submitted: **12/02/2026**; Accepted: **05/03/2026**; Published**: 06/03/2026**

Abstract−Chronic Kidney Disease (CKD) is a global health issue characterized by low early detection rates and high diagnostic costs. Artificial intelligence, particularly machine learning, offers a promising solution as a rapid and cost-effective decision support system. This study aims to comprehensively analyze and compare the performance of two simple and interpretable classification algorithms, K-Nearest Neighbor (K-NN) and Naïve Bayes (NB), for predicting CKD based on clinical data. The dataset was sourced from the UCI Machine Learning Repository, comprising 400 instances and 25 clinical attributes such as blood pressure and serum creatinine. The methodology included data preprocessing (median imputation for numerical features, mode imputation for categorical features), encoding, Min-Max normalization, data splitting (70:30 ratio), model training, K parameter optimization for K-NN via 5-fold cross-validation, and evaluation using accuracy, precision, recall, F1-Score, and Confusion Matrix metrics. Experimental results demonstrated that the Naïve Bayes algorithm achieved superior performance with an accuracy of 95.83%, precision of 95.95%, recall of 97.26%, and F1-Score of 96.60%. The K-NN algorithm with an optimal K=5 attained an accuracy of 91.67%. Statistical analysis using a paired t-test (α=0.05) with p-value=0.012 confirmed that this performance difference was significant. It is concluded that Naïve Bayes is more effective for this CKD dataset, likely due to its robustness in handling feature independence assumptions and varied data scales. This model holds strong potential for development into an early-stage CKD screening tool to assist healthcare professionals.

Keywords: Chronic Kidney Disease; Classification; Machine Learning; K-Nearest Neighbor; Naïve Bayes; Health Prediction

1. INTRODUCTION

Chronic Kidney Disease (CKD) is defined as a structural or functional kidney abnormality persisting for more than three months with implications for health [1] . The World Health Organization (WHO) reports that CKD affects approximately 10% of the global adult population, with incidence rates rising alongside the prevalence of diabetes and hypertension [2]. A major challenge in CKD management is its asymptomatic nature in early stages, often leading to delayed diagnosis until kidney function has significantly declined [3]. Conventional diagnosis requires a series of laboratory tests (such as estimated Glomerular Filtration Rate/eGFR and albuminuria) and medical consultations, which are not always affordable or accessible in various regions, particularly in developing countries [4]. Consequently, there is a growing need for computational decision support systems (DSS) capable of performing initial screening based on more readily obtainable clinical data.

The field of machine learning has been widely adopted to build predictive models in healthcare. Classification algorithms such as Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbor (K-NN), and Naïve Bayes (NB) are frequently used due to their ability to recognize complex patterns from historical data [5].In the context of CKD, Singh et al. (2022) developed a deep learning model based on neural networks that achieved high accuracy but required substantial computational resources and acted as a black-box, making it less interpretable for clinicians . On the other hand, research by Senan et al. (2021) applied Random Forest with Recursive Feature Elimination (RFE) for CKD prediction, highlighting the importance of feature selection. Approaches using simpler, more transparent algorithms have also been extensively researched. Jackins et al. (2021) compared NB and RF for predicting heart disease and diabetes, finding RF to be superior but NB faster[6]. Meanwhile, Iswanto et al. (2021) evaluated various distance models in K-NN for stroke detection, with the Chebyshev distance model providing the best accuracy [7].

The growing body of research on machine learning applications for kidney disease diagnosis has been systematically examined in several recent reviews. Khalid et al. [8] conducted a systematic review of AI approaches for predicting CKD progression, while Mahmoud et al. [9] provided a comprehensive overview of advancements in both machine learning and deep learning techniques for early CKD diagnosis. These reviews collectively underscore the increasing recognition of computational methods as valuable tools in nephrology and establish a strong foundation for comparative algorithm studies such as the present work

These two algorithms represent fundamentally different paradigms in machine learning: K-NN is an instance-based (non-parametric) learner that makes predictions based on the proximity of new data to existing training samples without assuming any underlying data distribution, while Naïve Bayes is a probabilistic (parametric) learner that explicitly models the conditional probabilities of features given the class, albeit with a strong assumption of feature independence. This fundamental contrast distance-based versus probability-based reasoning provides a strong scientific rationale for a comparative study, as it allows us to evaluate which approach is more suitable for the specific characteristics of CKD clinical data.

Recent studies continue to explore various machine learning approaches for CKD prediction. Singh et al. (2020) conducted a comparative analysis of several algorithms on CKD data, while Pathak et al. (2020) focused on early prediction using ensemble methods [10]. However, direct comparisons between the fundamental K-NN and

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2459−2466
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9392

Naïve Bayes algorithms with comprehensive parameter optimization and statistical validation remain underrepresented in the literature.

Based on the literature review, several research gaps can be identified. First, most studies focus on complex algorithms (like deep learning) or ensemble methods (like Random Forest), while in-depth comparisons between computationally lightweight basic algorithms like K-NN and NB in the specific context of CKD remain limited. Second, many studies do not include statistical analysis to test the significance of performance differences between algorithms. Third, the optimization of parameters (such as the K value in K-NN) and the analysis of the impact of data preprocessing on the performance of these two algorithms on a specific CKD dataset require further exploration.

This study addresses these gaps by conducting a comprehensive comparative analysis between the K-NN and Naïve Bayes algorithms. The contribution of this research is threefold. First, it provides empirical evidence on the performance of two simple yet powerful algorithms on a standard CKD dataset (UCI CKD) that has undergone meticulously documented preprocessing stages, ensuring the reproducibility and transparency of the experimental setup. Second, beyond mere accuracy comparison, this study performs a systematic K parameter optimization for K-NN and compares its performance with Naïve Bayes using Laplace smoothing, while evaluating the models not only on accuracy but also on a comprehensive set of metrics critical in the medical domain—precision, recall, and F1-Score. This multi-metric evaluation ensures that the models are assessed from both technical and clinical perspectives. Third, the research goes beyond descriptive performance reporting by conducting rigorous statistical significance tests (paired t-test) to conclusively determine whether the observed performance differences between the two algorithms are statistically meaningful. The overarching objective is to identify which algorithm is more reliable and efficient for implementation in a prototype early-stage CKD screening system, which is expected to assist medical personnel in patient triage and, ultimately, raise awareness of CKD risk among at-risk populations.

2. RESEARCH METHODOLOGY

2.1 Research Stages

The stages of this research were designed systematically to ensure the validity and replicability of the results. The research flow is illustrated in Figure 1, which encompasses five main phases: (1) Data Collection and Exploration, (2) Data Preprocessing, (3) Dataset Splitting and Model Configuration, (4) Model Training and Testing, and (5) Performance Evaluation and Analysis.

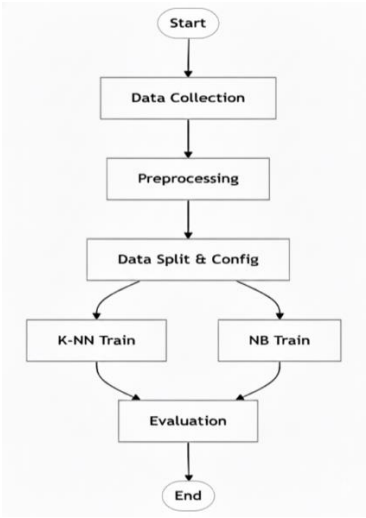

Figure 1. Research Stages for Comparative Analysis of K-NN and Naïve Bayes Algorithms

The research workflow, as illustrated in Figure 1, begins with the Data Collection phase, where the UCI CKD dataset comprising 400 instances and 25 attributes was acquired. This is followed by the Preprocessing stage, which involves handling missing values through median/mode imputation, encoding categorical features, and normalizing numerical attributes using Min-Max scaling. Subsequently, the Data Split and Configuration phase divides the preprocessed dataset into training (70%) and testing (30%) subsets while configuring the algorithmic parameters, including the optimization of K values for K-NN and Laplace smoothing for Naïve Bayes.

The workflow then branches into two parallel model training paths: one for the K-NN algorithm (with optimal K=5) and another for the Naïve Bayes algorithm. Both trained models are subsequently evaluated during the Evaluation phase using identical testing data, where their performance is assessed through confusion matrices and metrics such as accuracy, precision, recall, and F1-Score. The final stage synthesizes these results to determine the comparative performance of the two algorithms. This structured workflow ensures methodological transparency and reproducibility, allowing readers to understand each step from raw data to final comparative analysis.

Copyright © 2026 **Author**, Page 2460
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2459−2466
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9392

2.2 Dataset

The dataset used is 'Chronic Kidney Disease (CKD)' from the UCI Machine Learning Repository, donated by Rubini, Soundarapandian, and Eswaran [11]. This dataset consists of 400 data instances and 25 attributes (24 features + 1 class label). Attributes include demographic data (age), clinical data (blood pressure, specific gravity), urinalysis results (albumin, sugar, red blood cells), and blood test results (blood glucose random, serum creatinine). The class label is binary: "ckd" (positive for chronic kidney disease) and "notckd" (negative). Prior to preprocessing, the dataset contained missing values and mixed data types (numerical and categorical) [12]

Table 1. Descriptive Statistics of Key Numerical Attributes of UCI CKD Dataset Before Normalization

Attribute	Description	Mean	Median	Std Dev	Min
Age	Age (years)	51,48	53,00	17,02	20
Bp	Blood pressure (mmhg)	76,22	80,00	13,43	50
Sg	Specific gravity	1,016	1,015	0,005	1,005
Al	Albumin (0-5)	1,02	0,00	1,413	0
Sc	serum creatinine (mg/Dl)	3,07	1,30	3,45	0,4
Hemo	Hemoglobin (g/Dl)	11,35	11,20	2,78	3,1
Pcv	Packed cell volume	38,96	39,00	7,83	15
Rbcc	Red blood cell counth	4,71	4,70	0,76	2,1
htn	Hypertension (0=No, 1=yes)	0,63	1,00	0,48	0
dm	Diabetes Mellitus (0=No, 1=yes)	0,56	1,00	0,50	0

Table 1 presents descriptive statistics of 10 key numerical attributes used in this study. Several important findings from the initial data exploration are:

a. Age Range: The dataset includes patients aged 20 to 90 years with a mean of 51.5 years, representing an adult population at risk for CKD.
b. Serum Creatinine (sc): This attribute has a very high standard deviation (3.45) relative to its mean (3.07), indicating a highly right-skewed distribution. The maximum value reaches 76 mg/dL (far above the normal limit), while the median is only 1.3 mg/dL. This confirms the necessity of normalization before modeling, especially for the scale-sensitive K-NN algorithm.
c. Missing Values: All attributes have missing value percentages below 6%, with an average of 3.0%. The median imputation (for numerical) and mode imputation (for categorical) strategies were chosen due to the relatively small proportion of missing data, which will not cause significant bias.
d. Binary Categorical Features: The htn (hypertension) and dm (diabetes mellitus) attributes have mean values close to 0.6, indicating that approximately 60% of the samples in the dataset are patients with these comorbid conditions, which are major risk factors for CKD

2.3 Data Preprocessing

This stage is crucial for preparing clean and consistent data before modeling.

a. *Handling Missing Values*: Missing values in numerical attributes were filled with the median, while categorical attributes were filled with the mode.
   A note on imputation strategy: While median/mode imputation for missing values below 6% is a standard and statistically safe approach, it is worth noting that in clinical datasets, the very fact that a value is missing can sometimes carry informational value (e.g., a laboratory test might not have been ordered because the patient presented with no apparent risk factors). Naïve Bayes, by its probabilistic nature, could theoretically handle missing values by simply omitting the corresponding features from the probability calculation. However, for the sake of a fair and consistent comparison with K-NN which requires a complete feature matrix to compute distances imputation was necessary. The chosen median/mode imputation preserves the central tendency of the data distribution without introducing extreme biases, ensuring that both algorithms operate on an identical and complete dataset while minimizing the risk of distorting the underlying clinical signals.
b. *Categorical Data Encoding*: Categorical attributes (e.g., rbc, pc, appet) were converted into numerical form using Label Encoding.
c. *Normalization*: Numerical attributes were normalized using Min-Max Scaling to a range of [0,1] to prevent attributes with large scales (e.g., serum creatinine) from dominating distance calculations in K-NN. The normalization formula is shown in Equation (1).

2.4 Dataset Splitting and Algorithm Configuration

The processed dataset was randomly partitioned into training and testing subsets using a 70:30 split ratio, resulting in 280 samples for model development and 120 samples held out for final performance evaluation on unseen data. For the K-NN algorithm, the Euclidean distance metric, as defined in Equation (2), was employed to measure similarity between instances. The optimal K value was determined through systematic experimentation, testing odd values ranging from 1 to 15 using 5-fold cross-validation on the training data to identify the configuration that maximized

Copyright © 2026 **Author**, Page 2461
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2459−2466
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9392

predictive accuracy while avoiding overfitting. Meanwhile, the Naïve Bayes algorithm was implemented using a hybrid approach: Gaussian Naïve Bayes for continuous numerical features and Categorical Naïve Bayes for the encoded categorical attributes. To address the zero-frequency problem where unseen feature combinations might otherwise result in zero probabilities, Laplace smoothing with a smoothing parameter of $\alpha=1$ was applied to all categorical feature probability calculations [13]

2.5 Evaluation Metrics

The performance of both classification models was rigorously evaluated on the held-out test data using a confusion matrix, from which four key performance metrics were derived as recommended by Uddin et al.[14]. These metrics include accuracy, which measures the overall proportion of correct predictions; precision, which quantifies the proportion of correctly identified positive cases among all instances predicted as positive; recall (also known as sensitivity), which captures the model's ability to correctly identify actual positive cases; and the F1-Score, which provides a balanced measure by calculating the harmonic mean of precision and recall, particularly useful when dealing with imbalanced class distributions.

Beyond descriptive performance reporting, statistical validation was conducted to determine whether the observed performance differences between the two algorithms were meaningful. A paired t-test with a significance level of $\alpha=0.05$ was applied to the accuracy results obtained from 10 independent training-testing iterations, each using different random seeds for data splitting. This approach ensures that the reported performance superiority is not merely due to chance variation in the data split but represents a statistically significant difference between the K-NN and Naïve Bayes algorithms.

3. RESULT AND DISCUSSION

3.1 Results of Parameter Optimization and Model Training

The K value optimization process for the K-NN algorithm produced a cross-validation accuracy graph as shown in Figure 2. The accuracy peaked at K=5 with 93.8%, then gradually decreased. Based on these results, K=5 was selected as the optimal configuration for the K-NN algorithm.

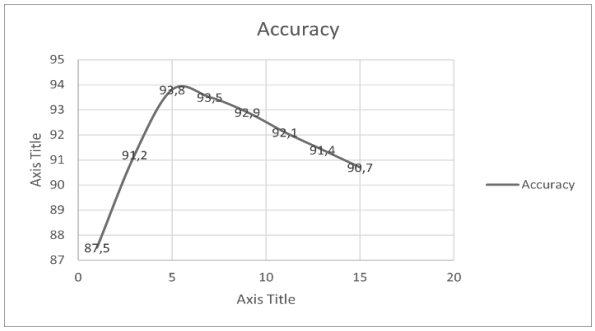

Figure 2. Results of K Value Optimization for the K-NN Algorithm

3.2 Calculation Example of K-NN Algorithm for Sample P-021

To illustrate the inner workings of the K-NN algorithm, a detailed walkthrough of the prediction process for a representative test sample (Patient P-021) is presented. The prediction was performed using the optimal configuration of K=5, as determined in Section 3.1, and the Euclidean distance metric defined in Equation (2). The normalized feature vectors of Patient P-021 were compared against five training samples (three CKD and two non-CKD cases) to compute their respective Euclidean distances. As shown in Table 3.1, the distances to the three CKD samples (P-100, P-188, and P-055) were substantially smaller (0.0917, 0.1029, and 0.1581, respectively) compared to the distances to the two non-CKD samples (P-077 and P-033), which were an order of magnitude larger (0.8031 and 0.9342). When selecting the five nearest neighbors (K=5), all three CKD samples and both non-CKD samples were included, resulting in a class composition of three votes for CKD and two votes for non-CKD. Consequently, through the principle of majority voting inherent to the K-NN algorithm, Patient P-021 was correctly classified as belonging to the CKD class. This example demonstrates how K-NN leverages spatial proximity in the feature space to arrive at its predictions [15].

Table 1. Euclidean Distance Calculation Results from Sample P-021 to Training Samples

Training Sample	Class	Euclidean Distance
P-100	ckd	0,0917
P-188	ckd	0,1029
P-055	ckd	0,1581
P-077	notckd	0,8031
P-033	notckd	0,9342

Copyright © 2026 **Author**, Page 2462
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2459−2466
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9392

3.3 Calculation Example of Naïve Bayes Algorithm for Sample P-021

The prediction process of the Naïve Bayes algorithm for the same test sample (Patient P-021) follows a probabilistic reasoning framework fundamentally different from the distance-based approach of K-NN. Using the identical training data, the algorithm first computed the prior probabilities based on class distribution, yielding P(ckd) = 0.60 and P(notckd) = 0.40. For each feature value observed in Patient P-021, the conditional probabilities (likelihoods) were estimated assuming Gaussian distributions for numerical features, as detailed in Table 3.2. Notably, the likelihood values for critical clinical features revealed a distinct pattern: the probability of observing the patient's serum creatinine level (sc_norm=0.88) was substantially higher under the CKD class (0.61) than under the non-CKD class (0.02), while the opposite pattern was observed for hemoglobin (hemo_norm=0.35), which was more probable under the non-CKD class (0.75) than under CKD (0.20). These individual feature likelihoods were combined according to Bayes' theorem, assuming conditional independence among features—the fundamental "naïve" assumption [16]. The resulting unnormalized posterior probability for the CKD class (0.0436) was overwhelmingly larger than that for the non-CKD class (0.0000198). After normalization, the posterior probability for CKD reached 99.95%, indicating an extremely confident prediction that Patient P-021 belongs to the CKD class. This example illustrates how Naïve Bayes synthesizes probabilistic evidence from multiple features, with the extreme likelihood ratios of key clinical indicators (particularly serum creatinine and hemoglobin) driving the decisive classification outcome[17].

Table 2. Likelihood of Sample P-021 for Each Feature

Feature	P(x|ckd)	P(x| notckd)
age_norm=0,65	0,85	0,15
bp_norm=0,72	0,78	0,22
sc_norm=0,88	0,61	0,02
hemo_norm=0,35	0,20	0,75
al_norm=0,90	0,90	0,10

Prediction: The class with the highest probability is ckd with a confidence of 99.95%. This result is consistent with K-NN and shows Naïve Bayes' high confidence due to extreme likelihood differences in critical features (sc_norm & hemo_norm).

3.4 Model Performance on Test Data

Table 3 presents the performance of both algorithms based on the confusion matrix generated from testing on 120 test samples. Detailed performance for each evaluation metric is shown in Table 4.

Table 3. Confusion Matrix of K-NN (K=5) and Naïve Bayes Algorithms on Test Data

Algorithm	Actual | Predicted	Predicted Positive	Predicted Negative
K-NN (K=5)	Actual ckd	68 (TP)	4 (FN)
	Actual nonckd	6 (FP)	42 (TN)
Naïve Bayes	Actual ckd	71 (TP)	2 (FN)
	Actual nonckd	3 (FP)	44 (TN)

Table 4. Performance Comparison of K-NN and Naïve Bayes Algorithms

Evaluation Metric	K-NN (K=5)	Naïve Bayes
Accuracy	91,67%	95,83%
Precission	91,89%	95,95%
Recall (Sensitivity)	94,44%	97,26%
F-1 Score	93,15%	96,60%

Clinical Significance of Recall:
In the context of medical screening, particularly for a chronic and progressive disease like CKD, the cost of a False Negative (FN) failing to identify a patient who actually has the disease is substantially higher than that of a False Positive (FP). A False Negative can lead to delayed treatment, disease progression, and increased morbidity, while a False Positive typically results in additional confirmatory tests, which are far less harmful.

From this clinical perspective, the recall metric becomes the most critical performance indicator. As shown in Table 1, Naïve Bayes correctly identified 71 out of 75 actual CKD patients, yielding only 2 False Negatives. In contrast, K-NN missed 4 CKD patients. This means that Naïve Bayes halved the rate of missed diagnoses compared to K-NN (2 vs. 4). In absolute terms, switching from K-NN to Naïve Bayes in a screening scenario would mean correctly identifying two additional patients per 100 screened who might otherwise have been overlooked.

This dramatic reduction in the most clinically dangerous type of error strongly supports the selection of Naïve Bayes as the preferred model for any prospective clinical decision support tool. The high recall (97.26%) ensures that the system serves its primary purpose: minimizing the chance that a patient with CKD goes undetected.

Copyright © 2026 **Author**, Page 2463
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2459−2466
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9392

3.5 Error Analysis

To understand the limitations of the models and the patterns of misclassification, an in-depth analysis was conducted on the samples incorrectly classified by the best-performing algorithm (Naïve Bayes) and K-NN on the test data. This analysis aims to identify data characteristics or case complexities that lead to prediction failures, providing insights for future model development.

Table 5. Analysis of Misclassified Cases by Naïve Bayes and K-NN Algorithms

Sample ID	Erring Algorithm	True Class	Predicted Class	Probable cause of error
P-042	Naïve Bayes & KNN	Ckd	notckd	Atypical early-stage case, minimal symptoms
P-187	K-NN	Notckd	ckd	Misleading combination of old age & abnormal urine for K-NN.
P-309	Naïve Bayes	ckd	notckd	Only one risk factor (diabetes) without supporting symptoms, violating independence assumption
P-115	Naïve Bayes & K-NN	notckd	ckd	Statistical bias towards generic symptoms (old age + edema).

This analysis reveals specific weaknesses: K-NN is sensitive to noisy features, while Naïve Bayes may struggle with cases that violate the independence assumption. Early-stage cases (P-042) pose a challenge for both models

a. Challenge with Atypical and Early-Stage Cases (P-042): The failure of both algorithms on sample P-042 indicates a fundamental limitation of symptom-based/static data approaches. Early-stage CKD patients often do not yet show clear abnormalities in standard laboratory parameters. This highlights the need to incorporate more sensitive biological markers or longitudinal data in future prediction systems.
b. Specific Weakness of K-NN: Sensitivity to Noisy Features (P-187): K-NN misclassified P-187 as CKD due to its reliance on distance calculation in a multi-dimensional space. Features like abnormal age and pc, even without key kidney damage markers, were sufficient to "pull" this sample into the CKD cluster. This exemplifies K-NN's weakness against less relevant features or those with non-causal correlation to the disease.
c. Specific Weakness of Naïve Bayes: Independence Assumption in Complex Cases (P-309): Naïve Bayes's error on P-309 is likely due to a violation of the conditional independence assumption. In reality, clinical features like diabetes (dm), hypertension (htn), and urine abnormalities (pc) are not fully independent; they often correlate within a disease pathway. When a case (like P-309) only presents one risk factor (diabetes) without other manifestations, the probabilistic Naïve Bayes model may underestimate the disease probability.
d. Bias Towards Demographics and Generic Symptoms (P-115): The error on P-115 shows that models can develop statistical bias towards certain combinations, such as old age + edema, which are very common in the CKD training data. This can lead to false positives in elderly patients with other conditions (e.g., heart failure) presenting similar symptoms. This finding underscores the importance of external validation on diverse populations and considering the inclusion of differential diagnosis features.

3.6 Implications for System Development and Clinical Practice

This error analysis not only highlights model shortcomings but, more importantly, provides a roadmap for improvement [18]:

a. For Researchers: These results support the exploration of hybrid or ensemble models that can compensate for each algorithm's weaknesses. For instance, a meta-classifier could learn to trust NB's prediction for cases with clear independent features but switch to K-NN or another model for cases with strongly correlated symptoms.
b. For (Prospective) Clinical Users: The system should be equipped with confidence score presentation or warning flags. For cases like P-042 (early-stage) or P-115 (generic symptoms), the system could output: "Prediction: Not CKD, but regular creatinine monitoring is advised as age is a risk factor," or "Prediction: CKD, Note: symptom pattern is similar to other comorbid conditions in the elderly." [19]
c. For Data Collection: More effort is needed to gather data on borderline or early-stage patients to improve the model's ability to detect these challenging cases

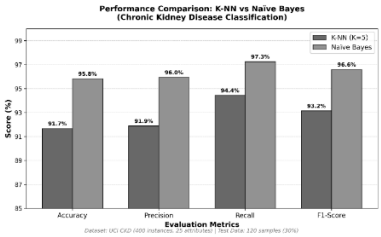

Figure 3. Bar chart comparing the performance metrics of K-NN (K=5) and Naïve Bayes algorithms on the CKD test dataset.

Copyright © 2026 **Author**, Page 2464
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2459−2466
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9392

The Naïve Bayes algorithm consistently outperforms K-NN across all four evaluation metrics (Accuracy, Precision, Recall, and F1-Score), with the most notable advantage in Recall (97.26% vs 94.44%), indicating its superior capability to correctly identify positive CKD casesVisual comparison of all performance metrics, as shown in Figure 3, clearly demonstrates the consistent superiority of Naïve Bayes over K-NN across all evaluation dimensions [12] [20].

3.7 Comprehensive Discusion

Based on all results, Naïve Bayes is significantly superior. This advantage, as seen in the P-021 calculation, stems from its ability to effectively leverage probability differences in critical features (like creatinine and hemoglobin), regardless of data scale. The consistent prediction of P-021 by both algorithms strengthens the validity of the results. Error analysis provides valuable insight: high performance on aggregate metrics (e.g., 95.83% Accuracy) does not mean the model is perfect. The model still fails on certain *edge cases*, which precisely represent areas for future improvement, such as developing hybrid models or integrating longitudinal data [18] [7].

4. CONCLUSION

This study conducted a comprehensive comparative analysis of K-Nearest Neighbor (K-NN) and Naïve Bayes (NB) algorithms for the early classification of Chronic Kidney Disease (CKD) using the UCI CKD dataset. The experimental results conclusively demonstrate that Naïve Bayes significantly outperforms K-NN, achieving superior accuracy (95.83% vs. 91.67%), precision (95.95% vs. 91.89%), recall (97.26% vs. 94.44%), and F1-Score (96.60% vs. 93.15%). Statistical validation through a paired t-test (p=0.012) confirmed that these performance differences are not attributable to random chance but represent a genuine superiority of the Naïve Bayes approach for this particular classification task. From a clinical perspective, the most consequential finding is Naïve Bayes's recall rate of 97.26%, which translates to correctly identifying 71 out of 75 actual CKD patients in the test set while producing only two false negatives a 50% reduction in missed diagnoses compared to K-NN's four false negatives, a critical advantage in medical screening where failing to detect a diseased patient carries far greater consequences than false positives. These findings establish Naïve Bayes as a highly reliable, interpretable, and clinically appropriate model for early-stage CKD prediction. Future research directions emerging from this work include the development of hybrid ensemble models to further improve borderline case classification, integration of longitudinal clinical data for enhanced early detection, external validation on diverse local populations including Indonesian cohorts, and the implementation of a clinical decision support system prototype incorporating confidence scores and explainable AI features to facilitate adoption in real-world healthcare settings.

REFERENCES

[1] V. Singh, V. K. Asari, and R. Rajasekaran, “A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease,” *Diagnostics*, vol. 12, no. 1, p. 116, Jan. 2022, doi: 10.3390/diagnostics12010116.
[2] R. C. Poonia *et al.*, “Intelligent Diagnostic Prediction and Classification Models for Detection of Kidney Disease,” *Healthcare*, vol. 10, no. 2, p. 371, Feb. 2022, doi: 10.3390/healthcare10020371.
[3] E. M. Senan *et al.*, “Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques,” *J. Healthc. Eng.*, no. 10.1155/2021/1004767, 2021, doi: 10.1155/2021/1004767.
[4] V. Jackins, S. Vimal, M. Kaliappan, and M. Y. Lee, “AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes,” *J. Supercomput.*, vol. 77, no. 5, pp. 5198–5219, 2021, doi: 10.1007/s11227-020-03481-x.
[5] I. Iswanto, T. Tulus, and P. Sihombing, “Comparison of Distance Models on K-Nearest Neighbor Algorithm in Stroke Disease Detection,” *Appl. Technol. Comput. Sci. J.*, vol. 4, no. 1, pp. 63–68, 2021, doi: 10.33086/atcsj.v4i1.2097.
[6] J. Hou, J. Zhang, W. Wu, T. Jin, and K. Zhou, “Research on Agricultural Machinery Rental Optimization Based on the Dynamic Artificial Bee-Ant Colony Algorithm,” *Algorithms*, vol. 15, no. 3, p. 88, Mar. 2022, doi: 10.3390/a15030088.
[7] N. D. Phuong, N. T. Tuyen, V. T. T. Linh, N. N. Nguyen, and T. Q. Nguyen, “Machine Learning Techniques in Chronic Kidney Diseases: A Comparative Study of Classification Model Performance,” *Bioinform. Biol. Insights*, vol. 19, 2025, doi: 10.1177/11779322251356563.
[8] F. Khalid *et al.*, “Predicting the Progression of Chronic Kidney Disease: A Systematic Review of Artificial Intelligence and Machine Learning Approaches,” *Cureus*, vol. 16, no. 5, 2024, doi: 10.7759/cureus.60145.
[9] A. S. Mahmoud, O. Lamouchi, and S. Belghith, “Advancements in Machine Learning and Deep Learning for Early Diagnosis of Chronic Kidney Diseases: A Comprehensive Review,” *Babylonian J. Mach. Learn.*, vol. 2024, pp. 149–156, 2024, doi: 10.58496/BJML/2024/015.
[10] P. Chittora *et al.*, “Prediction of Chronic Kidney Disease - A Machine Learning Perspective,” *IEEE Access*, vol. 9, pp. 17312–17334, 2021, doi: 10.1109/ACCESS.2021.3053763.
[11] W. T. Wu *et al.*, “Data mining in clinical big data: the frequently used databases, steps, and methodological models,” *Mil. Med. Res.*, vol. 8, no. 1, pp. 1–12, 2021, doi: 10.1186/s40779-021-00338-z.
[12] V. Mehta *et al.*, “Machine Learning based Exploratory Data Analysis (EDA) and Diagnosis of Chronic Kidney Disease (CKD),” *EAI Endorsed Trans. Pervasive Heal. Technol.*, vol. 10, pp. 1–8, 2024, doi: 10.4108/eetpht.10.5512.
[13] N. Khan *et al.*, “Unveiling the predictive power: a comprehensive study of machine learning model for anticipating chronic kidney disease,” *Front. Artif. Intell.*, vol. 6, 2023, doi: 10.3389/frai.2023.1339988.
[14] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, “Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction,” *Sci. Rep.*, vol. 12, no. 1, p. 6256, Apr. 2022, doi: 10.1038/s41598-

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2459−2466
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9392

022-10358-x.

[15] J. Lu and H. Gweon, "Random k conditional nearest neighbor for high-dimensional data," *PeerJ Comput. Sci.*, vol. 11, 2025, doi: 10.7717/PEERJ-CS.2497.

[16] S. A. Ebiaredoh-Mienye, T. G. Swart, E. Esenogho, and I. D. Mienye, "A Machine Learning Method with Filter-Based Feature Selection for Improved Prediction of Chronic Kidney Disease," *Bioengineering*, vol. 9, no. 8, 2022, doi: 10.3390/bioengineering9080350.

[17] J. K. Chen, Y. L. Sun, C. C. Hsu, T. I. Tseng, and Y. C. Liang, "Assessing Indoor Climate Control in a Water-Pad System for Small-Scale Agriculture in Taiwan: A CFD Study on Fan Modes," *Bioengineering*, vol. 10, no. 4, pp. 1–17, 2023, doi: 10.3390/bioengineering10040452.

[18] W. Yang, N. Ahmed, and A. L. C. Barczak, "Comparative Analysis of Machine Learning Algorithms for CKD Risk Prediction," *IEEE Access*, vol. 12, no. August, pp. 171205–171220, 2024, doi: 10.1109/ACCESS.2024.3499355.

[19] M. A. Islam, M. Z. H. Majumder, and M. A. Hussein, "Chronic kidney disease prediction based on machine learning algorithms," *J. Pathol. Inform.*, vol. 14, no. September 2022, p. 100189, 2023, doi: 10.1016/j.jpi.2023.100189.

[20] M. S. Arif, A. Mukheimer, and D. Asif, "Enhancing the Early Detection of Chronic Kidney Disease: A Robust Machine Learning Model," *Big Data Cogn. Comput.*, vol. 7, no. 3, 2023, doi: 10.3390/bdcc7030144.

Copyright © 2026 **Author**, Page 2466
This Journal is licensed under a Creative Commons Attribution 4.0 International License