International Journal of Computing and Artificial Intelligence

E-ISSN: 2707-658X
P-ISSN: 2707-6571
Impact Factor (RJIF): 5.57
https://www.computersciencejournals.com/ijcai/
IJCAI 2025; 6(2): 378-384
Received: 02-10-2025
Accepted: 05-11-2025

Dr. Archana Bhade
Assistant Professor, Department of Information Technology, Government College of Engineering, Chhatrapati Sambhajinagar, Maharashtra, India

Ashwini Maher
Department of Information Technology, Government College of Engineering, Chhatrapati Sambhajinagar Maharashtra, India

Dr. Smita Chavan
HoD, Department of Information Technology, Government College of Engineering, Chhatrapati Sambhajinagar, Maharashtra, India

Bhumika Pandhare
Department of Information Technology, Government College of Engineering, Chhatrapati Sambhajinagar, Maharashtra, India

Vaishnavi Walekar
Department of Information Technology, Government College of Engineering, Chhatrapati Sambhajinagar Maharashtra, India

Renuka Daud
Department of Information Technology, Government College of Engineering, Chhatrapati Sambhajinagar Maharashtra, India

Corresponding Author:
Dr. Archana Bhade
Assistant Professor, Department of Information Technology, Government College of Engineering, Chhatrapati Sambhajinagar, Maharashtra, India

Comparative analysis of chronic kidney disease prediction using machine learning algorithm

Archana Bhade, Ashwini Maher, Smita Chavan, Bhumika Pandhare, Vaishnavi Walekar and Renuka Daud

DOI: https://www.doi.org/10.33545/27076571.2025.v6.i2e.224

Abstract
Chronic kidney disease is a progressive medical condition that causes gradual and often irreversible damage to kidney function. Early detection is crucial because most patients remain asymptomatic until the disease reaches advanced stages, at which point treatment options become limited and costly. This study explores the use of machine learning techniques to predict CKD at an early stage using clinical patient data. The raw dataset is preprocessed through data cleaning, handling missing values, and feature encoding to ensure model readiness. Feature selection techniques are applied to identify the most relevant clinical parameters contributing to CKD prediction. Multiple machine learning algorithms, including Decision Tree, Random Forest, Support Vector Machine (SVM), Logistic Regression, and K- Nearest Neighbors (KNN), are trained and evaluated. Their performance is assessed using key metrics such as accuracy, precision, recall, and F1-score. The results highlight the most effective model for reliable CKD detection, demonstrating that machine learning can serve as a valuable decision-support tool for healthcare professionals, enabling timely diagnosis and better treatment planning.

Keywords: Psychiatric disorders, suicide, suicide attempt, first admission, recurrent admission, schizophrenia, bipolar disorder, depression, substance abuse disorder

Introduction
Chronic Kidney Disease (CKD) is a major global health concern that leads to progressive and irreversible kidney damage, eventually resulting in end-stage renal disease (ESRD). Studies indicate that CKD affects around 10-14% of the world's population, causing significant healthcare and economic burdens [3]. Early detection is crucial, as timely intervention can help delay disease progression and improve patient survival [1]. Traditional diagnostic practices, such as relying on serum creatinine and estimated glomerular filtration rate (eGFR), provide valuable insights but are often limited by patient-specific variability and inaccuracies [5].

To address these challenges, machine learning (ML) and deep learning (DL) techniques have been increasingly adopted for CKD prediction and diagnosis. Researchers have applied a wide range of algorithms, including decision trees, random forests, extra trees, and neural networks, to improve prediction accuracy [2]. Comparative studies have shown that ensemble methods like Random Forest and Extra Trees outperform traditional classifiers, achieving higher accuracy, precision, recall, and F1-scores [4]. Recent frameworks, such as Deep-Kidney, further demonstrate the potential of deep learning to provide more reliable predictions and assist clinicians in early intervention [6].

Literature Survey
Chronic kidney disease (CKD) is a significant public health issue worldwide, with increasing prevalence and serious complications if not detected early. Early diagnosis is crucial to prevent progression to end-stage renal failure, which often requires dialysis or kidney transplantation. Traditional diagnostic methods rely heavily on clinical expertise and laboratory tests, but these can be time-consuming and sometimes fail to detect early-stage disease. In recent years, machine learning (ML) and deep learning (DL) techniques have been increasingly applied to predict CKD using patient clinical records, laboratory investigations, and demographic information. These computational approaches aim to

enhance prediction accuracy, reduce the chances of misdiagnosis, and identify the most relevant clinical features that contribute to the disease, which can significantly aid in early intervention and treatment planning [1] [7].

Nagamani *et al.* [1] conducted a study to evaluate the effectiveness of multiple machine learning algorithms, including decision trees, random forests, and support vector machines, for predicting CKD. Their research emphasized that the careful selection and tuning of algorithms play a critical role in achieving optimal predictive performance. In particular, they highlighted the advantages of ensemble methods, which combine the predictions of multiple models to improve robustness and reduce overfitting. The study also demonstrated that, when applied to comprehensive clinical datasets, machine learning algorithms can achieve high accuracy in CKD prediction, indicating their potential as reliable decision-support tools for healthcare professionals Vashisth *et al.* [2] focused on the application of a multi-layer perceptron (MLP) neural network for diagnosing CKD. Their study illustrated that MLPs, despite being relatively simple neural network architectures, are highly capable of capturing complex nonlinear relationships among various clinical parameters such as blood urea, serum creatinine, and glomerular filtration rate. The model achieved promising predictive accuracy, showing that even basic neural network models, when properly trained, can serve as effective diagnostic tools in clinical settings. Moreover, their work highlighted the importance of preprocessing and normalization of clinical data to ensure that neural network models can learn patterns efficiently.

Chittora *et al.* [3] carried out a comprehensive comparative analysis of several machine learning algorithms, including logistic regression, random forest, k-nearest neighbors, and gradient boosting, to predict CKD. Their study emphasized that while simpler models may offer faster computation and easier interpretability, ensemble and tree-based methods consistently provide higher accuracy and robustness when dealing with complex clinical datasets. By systematically evaluating multiple algorithms, the study offered valuable insights into selecting appropriate models based on the trade- off between performance

Shukla *et al.* [4] investigated the significance of feature selection and identification of the most influential clinical attributes in CKD prediction. Their analysis revealed that parameters such as blood pressure, serum creatinine, albumin levels, and age significantly affect model performance. Prioritizing these features not only improves predictive accuracy but also enhances the interpretability of the models, which is crucial for gaining clinicians' trust. The study demonstrated that integrating domain knowledge with machine learning techniques allows for more precise and clinically meaningful predictions, bridging the gap between computational models and real-world medical applications.

Farjana *et al.* [5] explored hybrid and ensemble machine learning approaches to enhance CKD prediction accuracy. Their research showed that combining multiple algorithms through techniques like bagging, boosting, or stacking can significantly reduce prediction errors and improve the overall robustness of the models. The study highlighted that ensemble strategies are particularly effective in healthcare applications, where prediction reliability is critical. By leveraging the strengths of individual algorithms and minimizing their weaknesses, these hybrid models provide a more stable and accurate assessment of CKD risk, which can ultimately lead to better patient outcomes.

Saif *et al.* [6] proposed a deep learning framework called Deep-Kidney, which uses multi-layer neural networks to predict CKD. The model demonstrated superior performance compared to traditional machine learning techniques, particularly in terms of accuracy and sensitivity, when applied to larger datasets. The study emphasized that deep learning models are capable of automatically learning complex, nonlinear patterns from high-dimensional clinical data that may not be captured by conventional methods. This capability makes deep learning a powerful tool for early CKD detection, potentially enabling clinicians to identify high-risk patients more effectively and take timely preventive measures.

Hassan *et al.* [7] conducted a comparative study using comprehensive patient clinical records to develop and evaluate multiple machine learning models for CKD prediction. Their study highlighted the importance of using extensive and high- quality datasets to train models that can generalize well to diverse patient populations. They demonstrated that well-tuned machine learning algorithms, when combined with thorough data preprocessing and feature selection, can accurately predict CKD risk and assist clinicians in making informed decisions. The research further emphasized that predictive models are not only tools for diagnosis but can also support proactive healthcare management by identifying patients who require closer monitoring or early intervention.

Overall, the literature collectively indicates that while traditional machine learning models provide reliable performance for CKD prediction, ensemble and deep learning approaches often offer superior accuracy, robustness, and sensitivity. Feature selection, hybrid modeling, and the integration of domain knowledge enhance model effectiveness and interpretability, which are essential for practical clinical applications. These studies underscore the potential of computational intelligence techniques to revolutionize early CKD detection and improve patient care outcomes, paving the way for more personalized and data-driven healthcare solutions [1]-[7].

Methodology

The methodology adopted in this study is a structured, multi- step process designed to convert raw clinical data into actionable insights through the application of machine learning models. This systematic approach ensures that the resulting models are accurate, reliable, and capable of assisting in the early detection of Chronic Kidney Disease (CKD). The methodology comprises several key stages: data collection, data cleaning, preprocessing, feature selection, model building, evaluation, and model selection. Each stage is explained in detail below.

Data Collection

The dataset employed in this study is the Chronic Kidney Disease (CKD) dataset, which comprises 1659 patient records with 54 attributes, including both numerical and categorical variables. Numerical attributes include age, blood pressure, blood glucose, blood urea, serum creatinine, sodium, potassium, hemoglobin, and others, while categorical attributes consist of red blood cell count type, appetite, hypertension, diabetes mellitus, and presence of anemia. The target variable, located in the final column,

indicates whether a patient is classified as CKD or non-CKD. This dataset was selected for several reasons: it is publicly available and has been extensively used in prior research, facilitating reproducibility and benchmarking; it contains a diverse mix of clinical, demographic, and laboratory features suitable for building robust predictive models; and despite slight class imbalance, it provides sufficient samples from both CKD and non-CKD classes to train machine learning models effectively. The raw dataset was obtained and imported into the analysis environment for further preprocessing and analysis.

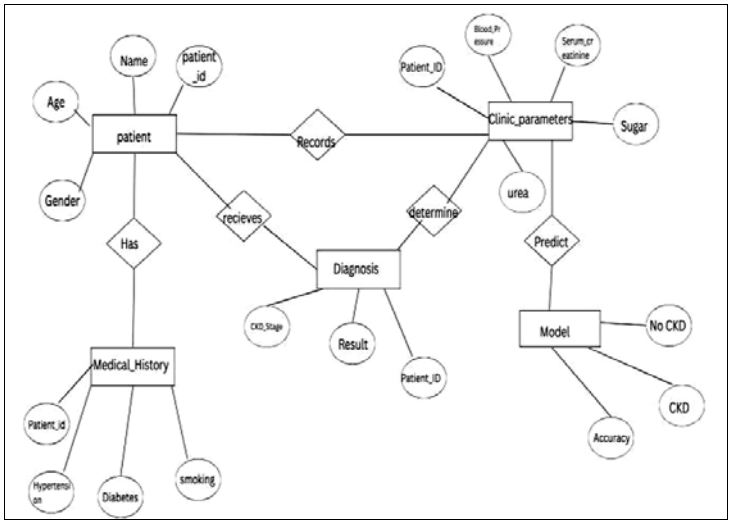

Fig 1: ER diagram

Data Preprocessing

Before applying machine learning algorithms, the raw Chronic Kidney Disease (CKD) dataset underwent an extensive preprocessing phase to ensure data quality, consistency, and suitability for modeling. Medical datasets, by nature, often contain inconsistencies, missing values, outliers, and noise, all of which can significantly impact the accuracy and reliability of predictive models if not handled properly. The first step in preprocessing involved addressing missing values. For numerical attributes such as blood pressure, serum creatinine, and blood urea, missing entries were imputed using the mean or median, with the choice depending on whether the feature distribution was normal or skewed. This approach prevents the introduction of bias and maintains the statistical characteristics of the data. For categorical attributes, such as red blood cell count type, appetite, or presence of hypertension, missing values were filled using the mode, i.e., the most frequently occurring category, ensuring that these essential clinical variables remained complete without compromising the overall dataset structure. Next, the dataset was checked for duplicate records, which could arise from repeated patient entries or data entry errors.

Duplicate rows were removed to prevent over-representation of certain records, which could bias the model and reduce its generalization ability. Following this, categorical variables were standardized to maintain consistency. Variations in text entries, such as ―yes,‖ ―Yes,‖ or ―Y,‖ were converted to a uniform format. This step is critical because machine learning algorithms cannot handle inconsistent categorical inputs and may treat different representations of the same category as distinct features. Outlier detection and treatment formed another important component of preprocessing. Extreme values were identified using box plots, histograms, and Z-score analysis. Outliers resulting from obvious data entry errors were removed, while valid extreme values were capped at acceptable thresholds to prevent distortion of the feature distribution, which is particularly important for algorithms sensitive to feature scales.

Once cleaning was complete, categorical features were encoded into numerical representations to make the dataset compatible with machine learning models. Binary features, such as ―yes/no‖ variables, were encoded using label encoding, while multi-class categorical features, such as red blood cell type, were converted using one-hot encoding, generating separate columns for each category to preserve information without introducing ordinal bias. Finally, all numerical attributes were scaled using min-max normalization, transforming their values into a [0,1] range.

Scaling ensures that features with larger ranges, such as blood urea or serum creatinine, do not dominate learning, which is especially crucial for distance-based algorithms like K-Nearest Neighbors (KNN) and Support Vector Machines (SVM). Additionally, the dataset exhibited minor class imbalance, with non-CKD cases slightly outnumbering CKD cases. To address this, the Synthetic Minority Over-sampling Technique (SMOTE) was applied, generating synthetic samples for the minority class. This process ensured a balanced dataset, reduced model bias toward the majority class, and improved the ability of classifiers to correctly detect CKD cases. Fig. 2. Architecture Representation of CKD.

Feature Selection

Feature selection plays a vital role in developing accurate and efficient machine learning models for Chronic Kidney Disease (CKD) prediction. Medical datasets often contain several attributes, some of which may be redundant, weakly related, or irrelevant to the target class. Using all available features without filtering can lead to overfitting, where the model learns unnecessary patterns from noise in the data, and it can also increase computation time.

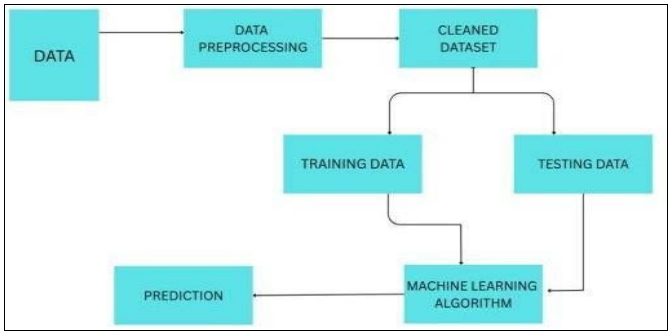

Fig 2: Architecture Representation of CKD

Therefore, feature selection is performed to identify the most significant clinical parameters that contribute to CKD classification, reduce dimensionality, and improve overall model generalization. In this study, feature selection was carried out using a combination of statistical, filter-based, and wrapper-based methods. Initially, a correlation matrix was generated to analyze the relationships between features and the target class. Features with very low or negligible correlation were removed, while highly correlated variables were carefully examined to avoid redundancy.

For categorical features such as diabetes mellitus, hypertension, red blood cell type, and appetite, chi-square tests were performed to measure their dependence on the class label. For continuous features like blood pressure, hemoglobin, blood urea, and serum creatinine, statistical significance tests such as ANOVA (Analysis of Variance) and Mutual Information were used to measure how much each feature contributed to predicting CKD. After statistical filtering, wrapper-based methods like Recursive Feature Elimination (RFE) were applied to further refine the feature set. RFE works by training a model repeatedly, ranking features based on their importance, and removing the least significant ones in each iteration until the optimal subset is obtained.

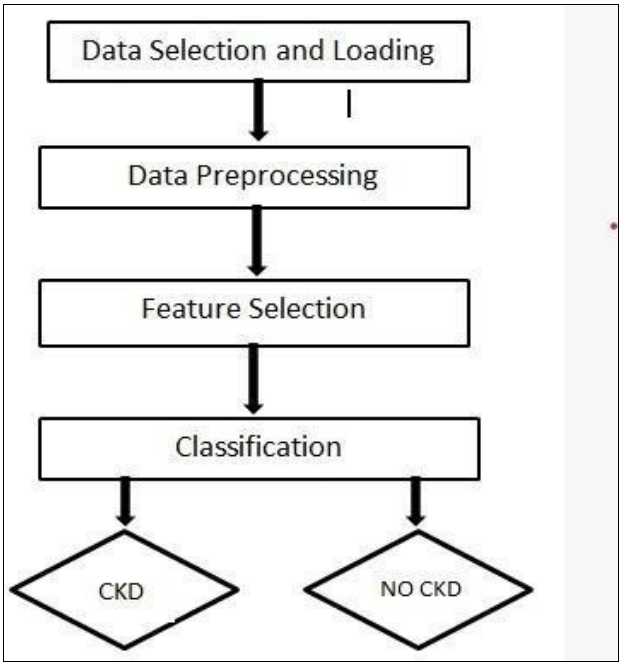

Fig 3: Flowchart of proposed System

This ensures that only the most informative features remain in the final dataset. The resulting subset contained clinically significant attributes such as blood pressure, specific gravity, albumin, serum creatinine, hemoglobin, packed cell volume, diabetes status, and hypertension—all of which are known medical indicators of kidney function. By reducing the number of irrelevant or redundant features, this process improved model training efficiency, reduced computational cost, and enhanced the predictive performance of the machine learning models. Additionally, it improved interpretability, as doctors can focus on the key parameters most strongly associated with CKD, making the model's results easier to explain in a clinical setting.

Compared Algorithms

In this work, we focus on the comparative analysis of six supervised machine learning algorithms—K-Nearest Neighbors (KNN), Logistic Regression, Decision Tree, Extra Trees, Random Forest, and Support Vector Machine (SVM)— for the prediction of kidney disease. Both linear and nonlinear classification approaches were considered to evaluate their ability to accurately classify patients into CKD (Chronic Kidney Disease) and non-CKD classes. To ensure a fair comparison, all models were trained on the same preprocessed dataset and tuned using a systematic hyperparameter optimization process. The performance of these models was compared using common metrics such as accuracy, precision, recall, F1-score, and AUC (Area Under the Receiver Operating Characteristic Curve). The evaluation was carried out using an extensive 10-fold cross-validation procedure to mitigate bias and variance in model assessment. Confusion matrices were also generated for each model to provide a detailed breakdown of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). These matrices offer granular insights into how well the algorithms differentiate CKD-positive and CKD-negative cases. Misclassified instances were further analyzed to investigate whether they resulted from overlapping feature distributions or model limitations.

Evaluation Metrics

The performance of the machine learning models was assessed using the following metrics: Accuracy: Measures the proportion of correctly classified instances out of the total number of instances, providing an overall indication of model performance. Precision: Represents the fraction of true CKD cases among all cases predicted as CKD by the model, reflecting its ability to avoid false positives. Recall (Sensitivity): Measures the fraction of correctly identified CKD cases among all actual CKD cases, indicating the model's capacity to capture true positives. F1- Score: A harmonic mean of precision and recall, used to provide a balanced measure of classification performance. AUC (Area

Under ROC Curve): Quantifies the model's ability to distinguish between CKD and non-CKD classes, with a value closer to 1 indicating superior discrimination capability

$$\text{Precision} = \frac{TP}{TP + FP}$$

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

$$F\text{-measure (F1-score)} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F\text{-measure (F1-score)} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Table 1: Confusion Matrics

Confusion Matrix	CKD (Predicted)	non-CKD (Predicted)
CKD	TP	FN
Non-CKD	FP	TN

Dataset Details

The dataset used in this study consists of 1,659 patient records with 54 attributes, encompassing demographic, lifestyle, medical history, clinical, and laboratory parameters. Demographic features include patient ID, age, gender, ethnicity, socioeconomic status, and education level. Lifestyle and behavioral factors such as smoking habits, alcohol consumption, physical activity, diet quality, and sleep quality are also recorded. Clinical measurements include systolic and diastolic blood pressure, fasting blood sugar, HbA1c, serum creatinine, blood urea nitrogen (BUN) levels, glomerular filtration rate (GFR), urinary protein levels, albumin-to- creatinine ratio (ACR), serum electrolytes (sodium, potassium, calcium, phosphorus), hemoglobin levels, and lipid profiles (total cholesterol, LDL, HDL, triglycerides). Medical history indicators, including family history of kidney disease, hypertension, diabetes, past acute kidney injury, and urinary tract infections, are provided along with medication usage such as ACE inhibitors, diuretics, NSAIDs, statins, and antidiabetic medications. Additional attributes capture clinical symptoms
(edema, fatigue, nausea/vomiting, muscle cramps, itching) and quality-of-life scores, as well as environmental and adherence factors such as heavy metal exposure, occupational chemical exposure, water quality, frequency of medical checkups, medication adherence, and health literacy. The target variable, Diagnosis, indicates the presence or absence of kidney disease and serves as the primary outcome for prediction. The dataset contains no missing values, making it well-suited for preprocessing and modeling without the need for imputation techniques.

Evaluation of Algorithms

In this research, six machine learning algorithms were employed to predict kidney disease, namely K-Nearest Neighbors (KNN), Logistic Regression, Decision Tree, Random Forest, Extra Trees, and Support Vector Machine (SVM). The evaluation of these algorithms was performed using standard classification metrics, including accuracy, precision, recall, and F1-score, which collectively assess both the predictive performance and reliability of the models.

Logistic Regression: Logistic regression is a linear model commonly employed for binary classification tasks. The model utilizes a logistic (sigmoid) function to estimate the probability of class membership, and the coefficients offer insight into the contribution of each individual feature. In this study, logistic regression was chosen as the baseline approach due to its simplicity in implementation and ease of interpretation. While it effectively highlights the influence of each predictor on the likelihood of kidney disease, the model has limited capacity to capture complex non-linear interactions among clinical and lifestyle variables.

Decision Tree: Decision trees classify instances by recursively partitioning the feature space into subsets based on feature thresholds. This model is interpretable and handles both numerical and categorical variables effectively. However, single decision trees are prone to overfitting, especially in datasets with noise or redundant features, which can affect their predictive stability.

Random Forest: Random Forest is an ensemble learning technique that constructs multiple decision trees and aggregates their predictions using majority voting. This approach reduces overfitting compared to a single decision tree and improves predictive performance by leveraging the diversity of multiple trees. Random Forest can handle high-dimensional data and complex feature interactions, making it particularly suitable for clinical datasets with mixed feature types.

Extra Trees (Extremely Randomized Trees): Extra Trees are similar to Random Forest but introduce additional randomness by selecting split thresholds at random for each feature. This further reduces variance and overfitting while maintaining high accuracy. In this study, Extra Trees provided slightly better generalization than Random Forest in some cases, due to increased randomness that prevents the model from relying heavily on any particular feature.

Support Vector Machine (SVM): SVM aims to find the optimal hyperplane that maximizes the margin between classes. By using kernel functions (linear, polynomial, or RBF), SVM can handle non-linear relationships in the data. It is effective for high-dimensional datasets and can achieve high accuracy, although it may require careful tuning of parameters such as the regularization coefficient (C) and kernel hyperparameters. SVM is also computationally intensive for large datasets.

Comparative Evaluation: Comparative analysis showed that ensemble methods, particularly Random Forest and Extra Trees, outperformed other models in terms of accuracy, precision, recall, and F1-score, indicating their robustness and superior predictive capability for kidney disease detection. Logistic Regression and KNN, while easier to interpret, had slightly lower performance but were valuable for understanding feature influence and local patterns, respectively. SVM performed well in separating classes in high-dimensional feature space but required

careful parameter tuning. Overall, the evaluation highlights a trade-off between interpretability and predictive performance, suggesting that ensemble-based models are optimal for clinical decision support in kidney disease prediction.

K-Nearest Neighbors (KNN): KNN is a non-parametric, instance-based learning algorithm that classifies a data point based on the majority class among its k-nearest neighbors in the feature space. In this study, KNN was applied after feature scaling to ensure uniform distance calculation. While KNN captures local similarity patterns effectively, it is sensitive to the choice of k and the presence of irrelevant features, which may reduce its generalization ability. It provides moderate accuracy and interpretable results by analyzing neighbor relationships.

Results and Discussion

After training and evaluating six machine learning models—K- Nearest Neighbors (KNN), Logistic Regression, Decision Tree, Random Forest, Extra Trees, and Support Vector Machine (SVM)—on the kidney disease dataset, their performance was compared using accuracy, precision, recall, and F1-score. Among these models, Extra Trees emerged as the optimal predictive model, demonstrating the highest accuracy and robustness. Its additional randomness in selecting split thresholds helped minimize overfitting, resulting in better generalization on unseen data.

Random Forest also performed strongly, achieving high predictive performance due to ensemble learning and majority voting across multiple decision trees, effectively capturing nonlinear interactions among features. Decision Tree, while slightly less accurate than the ensemble methods, provided clear interpretability and rule-based insights, which are valuable for understanding clinical decision pathways.

KNN and logistic regression showed moderate performance, useful for understanding local patterns and feature influences, whereas SVM required careful tuning to achieve reasonable classification results.

Based on these results, Extra Trees is recommended as the primary model for kidney disease prediction, followed by Random Forest and Decision Tree as alternative models. The study demonstrates that ensemble-based approaches, particularly those introducing additional randomness, are highly effective for medical datasets with diverse feature types and complex relationships.

These findings indicate that the selected models can serve as reliable clinical decision support tools, aiding healthcare professionals in early detection and management of kidney disease.

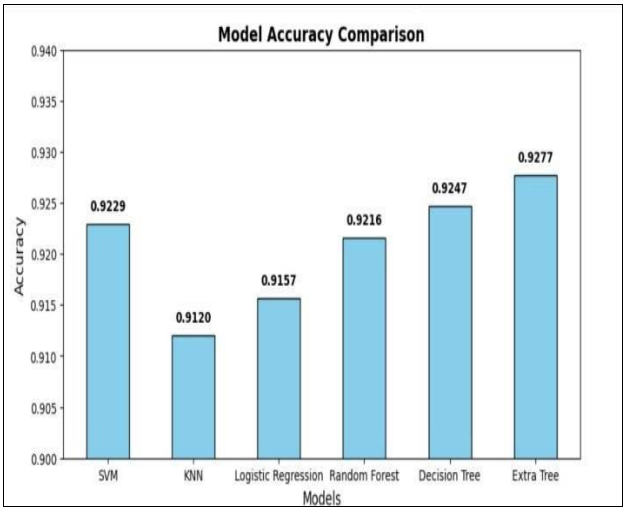

Fig 4: Comparison of Algorithms using Accuracy

Table 2: Comparison of classifier algorithms

Metrics	Accuracy	Precision	Recall	F1 score
Logistic Regression	0.9157	0.92	1.00	0.96
Support vector machine (SVM)	0.9229	0.92	1.00	0.96
k-Nearest Neighbor (KNN)	0.912	0.92	0.99	0.95
Decision Tree	0.9247	0.92	1.00	0.96
Random Forest	0.921	0.92	1.00	0.96
Extra Tree	0.9277	0.93	1.00	0.96

The experimental evaluation was carried out using multiple machine learning algorithms, namely Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression, Random Forest, Decision Tree, and Extra Tree classifiers. The primary objective was to determine the most effective model for the early detection and prediction of Chronic Kidney Disease (CKD). The performance of each model was assessed using accuracy as the evaluation metric. The results of the comparative analysis are illustrated in Figure X (Model Accuracy Comparison).

From the results, it is evident that all models achieved accuracy values above 91%, highlighting the strong predictive capability of machine learning methods in CKD classification tasks. The KNN classifier recorded the lowest accuracy at 91.20%, followed closely by logistic regression at 91.57%. These results suggest that while KNN and logistic regression are useful baseline models, they may not capture the complex patterns present in the dataset as

effectively as tree-based ensemble methods.

On the other hand, ensemble-based models demonstrated superior performance compared to linear approaches. The Random Forest classifier achieved an accuracy of 92.16%, indicating its robustness in handling feature variability and noise in the dataset. Similarly, SVM performed competitively with an accuracy of 92.29%, showcasing its ability to effectively separate classes in higher-dimensional spaces. Both models provide a strong balance between generalization and predictive accuracy.

Tree-based classifiers, namely Decision Tree and Extra Tree, exhibited the highest predictive accuracies. The decision tree achieved an accuracy of 92.47%, while the extra tree classifier outperformed all other methods with an accuracy of 92.77%. This performance gain can be attributed to the ability of tree- based algorithms to model non-linear relationships and interactions between features, which are often present in medical datasets. In particular, the Extra Tree classifier leverages randomized splits, which enhances generalization and reduces overfitting.

Based on the comparative analysis, the Extra Tree classifier was identified as the most suitable model for CKD prediction in this study. Although the margin of improvement over other classifiers is relatively small, its consistent performance highlights its potential as a reliable predictive model in healthcare applications. Nevertheless, further validation using larger and more diverse datasets, along with additional metrics such as precision, recall, and F1-score, is recommended to ensure fairness and robustness in clinical settings.

In conclusion, the study demonstrates that machine learning classifiers can effectively support medical practitioners in the early detection of CKD, with Extra Tree emerging as the best- performing model in this experimental setup. Future work will focus on integrating explainable AI techniques to provide better interpretability of model predictions, thereby increasing trust and adoption in real-world clinical environments.

Conclusion

In this study, six machine learning algorithms—K-Nearest Neighbors, Logistic Regression, Decision Tree, Random Forest, Extra Trees, and Support Vector Machine—were evaluated for predicting kidney disease using a dataset of 1,659 patient records with diverse demographic, clinical, lifestyle, and laboratory features.

Comparative analysis based on accuracy, precision, recall, and F1-score revealed that Extra Trees achieved the highest predictive performance, followed by Random Forest and Decision Tree. Ensemble-based models proved particularly effective in handling non-linear relationships and complex interactions among features while maintaining robustness and reducing overfitting.

The results demonstrate that machine learning models, especially Extra Trees, can serve as reliable clinical decision support tools for early detection of kidney disease, potentially aiding healthcare professionals in timely diagnosis andmanagement. This study also highlights the importance of feature diversity and proper model selection in medical prediction tasks. Future work could focus on expanding the dataset, incorporating additional biomarkers, and exploring advanced ensemble or hybrid techniques to further enhance prediction accuracy and clinical applicability.

References

1. Nagamani T, Kavya R, Lavanya D, Miruthula S. Chronic kidney disease prediction using machine learning algorithms. In: Proceedings of the 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE); 2024; pp. 979-983. DOI:10.1109/AMATHE61652.2024.10582109.
2. Chronic kidney disease using machine learning on patients' clinical records. Human-Centric Intelligent Systems. 2023;3:92-104. DOI:10.1007/s44230-023-00017-3.
3. Vashisth S, Dhall I, Saraswat S. Chronic kidney disease diagnosis using multi-layer perceptron classifier. In: Proceedings of the 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA); 2020; Uttar Pradesh, India. pp. 346-351. DOI:10.1109/ICCSEA49143.2020.9132906.
4. Chittora P, Chaurasia S, Chakrabarti P, Kumawat G, Chakrabarti T, Leonowicz Z, *et al*. Prediction of chronic kidney disease: a machine learning perspective. IEEE Access. 2021;9:17312-17326. DOI:10.1109/ACCESS.2021.3053763.
5. Shukla G, Dhuriya G, Pillai SK, Saini A. Chronic kidney disease prediction using machine learning algorithms and the important attributes for the detection. In: Proceedings of the 2023 IEEE IAS Global Conference on Emerging Technologies (GlobConET); 2023 May 19-21; London, UK. pp. 1-6. DOI:10.1109/GlobConET56651.2023.10149900.
6. Farjana A, Tabassum F, Pandit LPP, Das MC, Hasan M, Hossen MH, *et al*. Predicting chronic kidney disease using machine learning algorithms. In: Proceedings of the 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC); 2023 Jan; Las Vegas (NV). pp. 1-7. DOI:10.1109/CCWC57344.2023.10099221.
7. Saif D, Sarhan AM, Elshennawy NM. Deep-Kidney: an effective deep learning framework for chronic kidney disease prediction. Health Information Science and Systems. 2024;12(3):1-15. DOI:10.1007/s13755-023-00261-8.