Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
Special Issue on Exploring Knowledge Across Disciplines: New Paradigms and Practices 7(13s)

# EARLY PREDICTION OF CHRONIC KIDNEY DISEASE USING AN ENSEMBLE MACHINE LEARNING-BASED CLINICAL DECISION SUPPORT SYSTEM

Kancharla Anitha [1], B. Basaveswara Rao [2]

[1] Department of Computer Science and Engineering, Acharya Nagarjuna University, Amaravathi, Andhra Pradesh, India
[2] Department of Computer Science and Engineering, Acharya Nagarjuna University, Amaravathi, Andhra Pradesh, India

**Received** 12 March 2026
**Accepted** 17 May 2026
**Published** 28 May 2026

**Corresponding Author**
Kancharla Anitha,
kancharlaaniithaa@gmail.com

**DOI**
10.29121/shodhkosh.v7.i13s.2026.8433

**Funding:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Copyright:** © 2026 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License.

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.

OPEN ACCESS

## ABSTRACT

Chronic Kidney Disease (CKD) is a progressive medical condition that requires early prediction to reduce complications and healthcare costs. This study presents an ensemble-based clinical decision support system for early CKD prediction using two public datasets: UCI-CKD and CKD-15. A structured preprocessing pipeline was implemented. A fine-tuned Random Forest classifier was employed, and a feature ablation study was conducted to analyze the contribution of numerical, categorical, and clinically selected feature groups. Model performance was evaluated using stratified k-fold cross-validation (3-, 5-, and 10-fold). The model achieved near-perfect performance on the UCI-CKD dataset with 99.99% accuracy under cross-validation. On the CKD-15 dataset, the model achieved a maximum accuracy of 92.41%, with numerical features providing the highest predictive performance. The results demonstrate strong classification capability and good generalization for early CKD prediction.

**Keywords:** Chronic Kidney Disease (CKD), Random Forest Classifier, Feature Engineering, Cross-Validation, Early Prediction

## 1. INTRODUCTION

Chronic Kidney Disease (CKD) is identified as a global health priority, impacting millions of people and placing a high burden on the healthcare systems across the globe. According to epidemiological studies, CKD is estimated to affect close to 10% of the world's population and has a major contribution to morbidity and mortality rates because it is a progressive condition with comorbid associations of cardiovascular and metabolic disorders (Levin et al 2013). In spite of the existence of universal diagnostic techniques involving laboratory measurements like serum creatinine, blood urea nitrogen, and glomerular filtration rate (GFR), early detection of CKD is still an issue. The reason is the frequently asymptomatic presentation of CKD in the early stages, which results in delayed diagnosis and, therefore, reduces the

success of early intervention (Thomas et al 2008). Therefore, there is a pressing need for advanced, data-oriented strategies to enhance early diagnosis and patient outcomes.

In recent years, the rapid evolution of machine learning (ML) and data analytics has opened up new avenues for enhancing the predictive accuracy of medical diagnostics . The integration of ML into clinical practice promises to not only identify subtle patterns in complex and heterogeneous datasets but also to improve the reliability and interpretability of diagnostic models. In this context, our work uses two prominent datasets, the UCI-CKD dataset (Rubini et al 2015) and the CKD-15 dataset (Rabie 2024), both of which provide extensive information encompassing demographic details, clinical measurements, and laboratory test results. By combining these data sources, we aim to develop a robust predictive model that can effectively distinguish between CKD and non-CKD cases, thereby facilitating earlier and more accurate diagnoses.

The central challenge addressed by this study is the development of a predictive model that can reliably detect CKD in its early stages using a diverse set of patient data. Traditional screening methods, while clinically valuable, often fail to capture the complex interplay between various biological, demographic, and lifestyle factors. This is further complicated by issues such as data imbalance, missing values, and the heterogeneity inherent in clinical datasets. Our approach specifically addresses these challenges by implementing a fine-tuned Random Forest classifier, renowned for its ability to handle high-dimensional data and model non-linear relationships effectively. The study also incorporates rigorous data preprocessing techniques, including the imputation of missing values and standardization of numerical features, as well as the transformation of categorical variables through label encoding and one-hot encoding when appropriate.

The aim of this work is to provide a comprehensive evaluation of CKD prediction by analyzing the impact of different feature subsets on model performance. We investigate several feature configurations, including the use of all available features, only numerical features, only categorical features, and a manually selected subset of clinically relevant features. This ablation study is conducted using stratified k-fold cross-validation (with 3-, 5-, and 10-fold schemes) to ensure that the results are robust and generalizable across different data splits. By benchmarking the performance of our proposed fine-tuned Random Forest model against other established models in the literature, we seek to demonstrate its superior ability to accurately classify CKD cases. Performance metrics such as accuracy, precision, recall, and F1 score are employed to provide a detailed assessment of the model's predictive capabilities. Variyar and Karangara (2026)

This study not only underscores the potential of machine learning techniques in enhancing CKD diagnosis but also offers practical insights into the relative importance of various feature subsets. The results of our experiments indicate that while a model using all features tends to achieve the highest accuracy, a model built solely on numerical features also performs remarkably well, thereby highlighting the critical role of quantitative clinical measurements. Furthermore, our findings suggest that a carefully curated subset of features, selected based on clinical expertise, can achieve competitive performance with reduced complexity. These insights are essential for guiding future research in the domain of CKD prediction and for developing effective clinical decision support systems. The integration of advanced data processing techniques, rigorous cross-validation, and comprehensive model evaluation in our study provides a robust framework for the early detection of CKD, which could ultimately lead to better patient management and outcomes.

**Contribution to this study**

- Development of a machine learning-based framework for early CKD prediction using Random Forest.
- Implementation of a structured preprocessing and feature engineering pipeline for heterogeneous clinical data.
- Feature ablation analysis to evaluate the importance of different feature groups.
- Robust evaluation using stratified cross-validation and independent test validation.
- Comparative analysis with baseline machine learning models.
- Generalization and overfitting analysis for reliable performance estimation.

## 2. RELATED WORKS

In recent years, considerable research has focused on applying machine learning techniques to the early prediction of Chronic Kidney Disease (Ekanayake and Herath 2020; Revathy etal 2019; Rabby et al 2019 and Bai et al 2022). The related work encompasses a wide range of approaches, from classical methods such as logistic regression (Ahmed &

Alshebly 2019) and decision trees (Ilyas 2021) to more advanced ensemble methods, including Random Forests (Subasi et al 2017) and gradient boosting algorithms like XGBoost (Ogunleye and Q. -G. Wang 2019). These studies have used various datasets, most notably the UCI-CKD dataset, and have reported high accuracies along with impressive precision, recall, and F1 scores. Furthermore, emerging research has also explored deep learning architectures to cafVipture complex patterns within the clinical data (Singh et al 2022; Kuo et al 2019). This section reviews these advances in CKD prediction methodologies, highlighting the strengths and limitations of each approach, and situates our work a fine-tuned Random Forest classifier utilizing the UCI-CKD and CKD-15 datasets within this broader context.

**Table 1**

**Table 1 Related studies on Chronic Kidney Disease**

| Author(s) | Year | Methodology Applied | Dataset Used | Performance | Limitations |
|---|---|---|---|---|---|
| Ekanayake et al | 2020 | Machine Learning methods | Not specified | Proposed effective methods for early prediction of CKD | Specific methodologies and performance metrics not detailed |
| Yu Wang et al | 2021 | TRACE: Transformer-RNN Autoencoder-enhanced CKD Detector | Real-world medical records | Achieved 0.5708 AUPRC with a 2.31% improvement over the best-performing method | Challenges in early detection due to insufficient medical histories and complex risk factors |
| Pedro A. and Moreno-Sanchez | 2021 | Explainable Prediction Model using Ensemble Trees | CKD indicators dataset | Extreme Gradient Boosting classifier over 3 features achieved 99.2% accuracy with cross-validation and 97.5% with unseen data | Focused on a small number of features; applicability to broader datasets not assessed |
| Chittora et al | 2021 | Machine Learning methods | UCI-CKD | Using SMOTE with LASSO-selected features, a linear SVM achieved 98.46% accuracy, while a deep neural network reached 99.6% accuracy. | Model's complexity and interpretability remain challenging. |
| Habiba et al | 2023 | Machine Learning classifiers (K-Nearest Neighbors, Support Vector Machines, Artificial Neural Networks) | UCI-CKD | Discussed various ML classifiers for CKD prediction machine learning using key physiological variables achieving 100% accuracy with 25 features and 98.48% with only six by identifying dominant factors | generalizability may be limited by dataset size. |
| Pal | 2023 | Machine Learning methods | UCI-CKD | Proposed effective methods for early prediction of CKD and achieved 97.23% | The dataset is skewed. |
| Gogoi and Valan | 2024 | Review of Machine Learning techniques in CKD prediction and diagnosis | Various datasets | Comprehensive survey of ML techniques, highlighting recent trends and challenges | Review paper; does not present new experimental results |
| Liu et al | 2024 | Super learner strategy for risk prediction | Linked population health data from Canada, Denmark, and Scotland | Developed a model for risk prediction of kidney failure and mortality in CKD patients | Focused on specific populations; generalizability may be limited |
| Vetrithangam et al | 2024 | Machine Learning methods | Not specified | Proposed effective methods for early prediction of CKD by improving ResNet models ResNet152v2 with inception, ResNet101, and ResNet50, achieve 99.90%, 96.53%, and 93.97% accuracy. | model's generalizability may be affected by differences in dataset quality, noise, |
| Gupta et al | 2026 | RF-based data imputation, SMOTE for class balancing, Grey Wolf Optimizer (GWO)-based weighted meta- | UCI-CKD Dataset | Accuracy: 98.75%, Precision: 98.8%, Recall: 98.6%, F1: 98.7% | Evaluated only on UCI dataset; no external dataset validation; ensemble complexity |

| | | | | | |
|---|---|---|---|---|---|
| | | ensemble (DT, LR, GNB), SHAP & LIME for explainability | | | increases computational cost |
| Ramesh & Rao | 2025 | Optimized Multi-Layer Perceptron (MLP) with Feature Selection using Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) | AP-CKD Dataset (Regional dataset from Andhra Pradesh clinical centers) | Accuracy: 99.01% (MLP + PSO) | Dataset is region-specific; may not generalize to other populations; deep model interpretability not discussed |
| Rahman et al | 2025 | Machine Learning (KNN, Gradient Boosting) and Deep Learning (CNN, ANN), preprocessing and feature selection | Clinical dataset (400 patients, USA healthcare data) | Gradient Boosting Accuracy: 97% | Small dataset size; no external validation; deep learning models may require larger datasets |
| Kumar et al | 2025 | Hybrid Ensemble Model combining XGBoost + Random Forest, compared with SVM | UCI-CKD Dataset | Accuracy: ~99% | Evaluated only on UCI dataset; potential overfitting due to single dataset evaluation |
| Xue et al | 2026 | SURD-enhanced machine learning framework, multiple imputation, SMOTE, external validation using MIMIC-IV dataset | UCI-CKD + MIMIC-IV (27,834 EHR records) | External Validation AUC: 0.990 | Complex causal decomposition framework; high computational complexity |
| Ahmed et al | 2025 | Stacking ensemble ML with feature selection (LR, RFE, RF, MI, Chi-square, PCA), SMOTE, normalization | UCI-CKD + Simulated CKD Dataset | Accuracy: 100% (UCI), 96.7% (Simulated dataset) | Synthetic dataset may not fully represent real clinical data; risk of overfitting on UCI dataset |

## 3. METHODOLOGY

In this research work, we assessed a Random Forest classifier fine-tuned for Chronic Kidney Disease (CKD) prediction using the CKD and CKD-15 datasets. The CKD-15 Dataset is optimized to support clinical and laboratory test result-based classification and analysis of CKD. The dataset comprises demographic information, medical history, lifestyle parameters, and laboratory test results pertaining to kidney function. The dataset represents a rich source of predictive modeling, disease progression analysis, and risk assessment. We conducted extensive experiments by varying the feature subsets and employing three cross-validation schemes (3-fold, 5-fold, and 10-fold) to ensure robust performance estimation. An ablation study was also performed to determine the contribution of different groups of features. The following subsections detail the experimental results, tabulated performance metrics, and graphical visualizations.

**Figure 1**

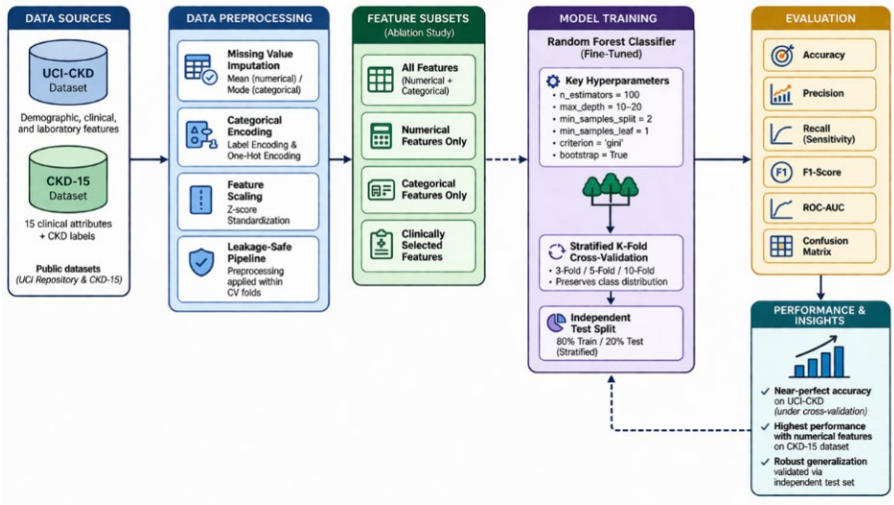


**Figure 1** Proposed Method Workflow

## 3.1. DATA PREPROCESSING FOR UCI-CKD DATASET WITH 25 FEATURES

The raw dataset consists of patient data with both numerical and categorical features. Preprocessing involves:

- **Handling Missing Values**

The missing values in numerical features (e.g., age, blood pressure (bp), specific gravity (sg), albumin (al), sugar (su), etc.) are imputed using the mean value, while missing values in categorical features (e.g., rbc, pc, pcc, etc.) are filled with the mode.

- **Data Type Conversion and Encoding**

Numerical columns are explicitly converted to numeric types. Categorical features are encoded using Label Encoding. The target variable (classification) is cleaned by stripping whitespace and then encoded into binary form (0 for "Not CKD" and 1 for "CKD").

- **Standardization**

Numerical features are standardized using Z-score normalization, as given by:

$$z = \frac{x - \mu}{\sigma}$$

Where x is the original value, μ is the mean, and σ is the standard deviation o the feature.

If $X \in \mathbb{R}^{n \times m}$ respresent the preprocessed feature matrix and Y∈{0,1}^n denote he binary target vector

**Figure 2**

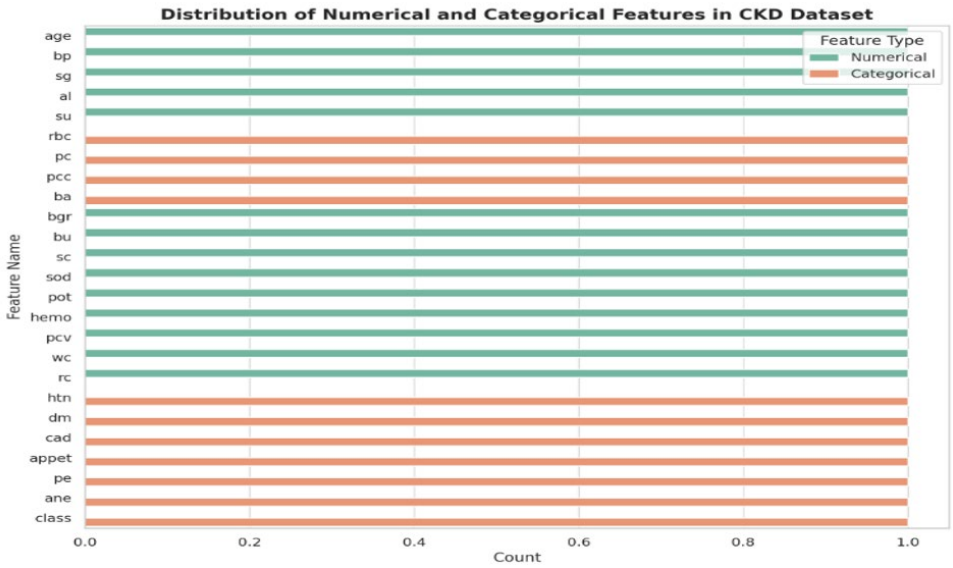


**Figure 2** Features of UCI-CKD Dataset

Figure 2 provides a quick reference to the key features, outlining their clinical relevance and data type, which is essential for the subsequent preprocessing and modeling steps in the CKD prediction project.

- **Features in the CKD-15 Dataset**

The dataset consists of 35 features from 1,659 patients, categorized into Demographic Details,Medical History, Lifestyle Factors, and Laboratory Test Results.

- **Data Preprocessing and Leakage Prevention**

To ensure robustness and eliminate potential data leakage, all preprocessing steps were strictly performed within the training folds during cross-validation. The preprocessing pipeline includes missing value imputation, categorical

encoding, and feature standardization, applied independently to each training fold and subsequently to the corresponding validation fold.

Missing values in numerical features (e.g., age, blood pressure, serum creatinine) were imputed using mean imputation computed exclusively from the training data. Similarly, categorical variables (e.g., hypertension, diabetes mellitus) were imputed using mode values derived from the training fold. This prevents inadvertent information leakage from validation data into the training process.

Categorical variables were encoded using label encoding, ensuring consistent mapping across folds. Numerical features were standardized using Z-score normalization. This fold-wise preprocessing strategy ensures that the evaluation metrics reflect true generalization performance and are not artificially inflated due to data leakage.

**Figure 3**

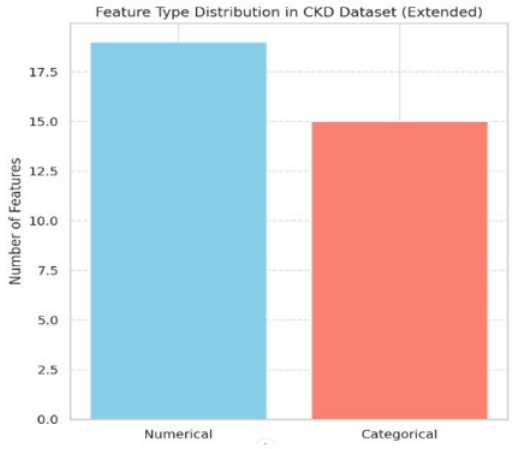


**Figure 3** Feature Distribution of CKD-15 Dataset

## 3.2. FEATURE ENGINEERING AND SELECTION

Different feature subsets are explored to assess their impact on model performance. In our experiments, we evaluate the following subsets:

- **All Features in both Datasets**

This subset comprises every feature available in the dataset. The rationale behind using all features is to leverage the full spectrum of information captured during data collection. This approach tests whether the inclusion of potentially redundant or noisy variables adversely affects the predictive performance. It serves as a baseline for comparison with other, more selective, feature subsets

- **Numerical Features in both Datasets**

This subset includes only the quantitative variables (e.g., age, bp, sg, al, su, etc.). Numerical features are continuous or discrete measurements that can be directly scaled (using methods like Z-score normalization) and often carry strong predictive signals. By focusing exclusively on numerical data, the model can assess the impact of these measurements on the prediction outcome without the influence of categorical noise in the UCI-CKD Dataset. The CKD-15 has the following Numerical Features: Only the quantitative variables (e.g., Age, BMI, SystolicBP, DiastolicBP, FastingBloodSugar, HbA1c, SerumCreatinine, BUNLevels, GFR, ProteinInUrine, ACR, various serum electrolytes, HemoglobinLevels, Cholesterol metrics, and QualityOfLifeScore).

- **Categorical Features in both Datasets**

In this subset, only the categorical variables in UCI-CKD Dataset (e.g., rbc, pc, pcc, ba, htn, dm, etc.) and CKD-15 Dataset categorical variables (e.g., Gender, Ethnicity, SocioeconomicStatus, EducationLevel, Smoking, AlcoholConsumption, PhysicalActivity, DietQuality, SleepQuality, Family histories, medication usage, and lifestyle factors) are used, after they have been converted into numerical form via label encoding. Categorical features represent qualitative aspects of the data, such as classifications or conditions that are not inherently numeric. Although these

features can provide complementary information, their contribution might be different in magnitude compared to numerical features. The separate evaluation of categorical features helps determine their standalone predictive power.

- **Selected Features in both Datasets**

This subset is a manually a selection of features for UCI-CKD Dataset (e.g., {age, bp, bgr, sc, sod}) chosen based on clinical relevance and prior domain knowledge. Chosen based on clinical relevance (in our experiments, these were {Age, BMI, GFR, SerumCreatinine, BUNLevels}) or CKD-15 Dataset. The goal of this approach is to focus on those features that are most likely to have a direct and significant impact on CKD prediction. By reducing the dimensionality and filtering out less relevant variables, this subset may help in reducing overfitting, simplifying the model, and potentially improving generalization performance.

## 3.3. MODEL FINE-TUNING AND TRAINING

A Random Forest classifier is employed due to its robustness to noise and ability to model complex interactions. The Random Forest is fine-tuned by optimizing key hyperparameters, such as the number of trees T, maximum depth d_max, and minimum samples per split. Each decision tree in the ensemble is built by recursively partitioning the data to minimize the Gini impurity:

$$Gini(D) = 1 - \sum_{k=1}^{K} p_k^2,$$

Where p_k^2 is the proportion of samles belonging to class k in node D and K=2 for binary classification.

The prediction for an input sample x is determined by majority voting across all trees:

$$\hat{y} = mode\{f_1(\boldsymbol{X}), f_2(\boldsymbol{X}), \dots, f_T(\boldsymbol{X})\}.$$

Hyperparameter tuning is conducted via cross-validation (3-fold, 5-fold, and 10-fold) to optimize performance, ensuring that the model generalizes well to unseen data.

- **Random Forest Model Configuration**

The Random Forest classifier was selected due to its robustness and ability to handle heterogeneous clinical data. The model was fine-tuned using grid search within cross-validation. The final hyperparameters used in this study are presented in Table 2:

**Table 2**

**Table 2 Random Forest Parameters**

| Parameter | Description | Value Used |
|---|---|---|
| n_estimators | Number of decision trees | 100 |
| max_depth | Maximum depth of trees | Oct-20 |
| min_samples_split | Minimum samples required to split a node | 2 |
| min_samples_leaf | Minimum samples in leaf node | 1 |
| max_features | Number of features considered per split | sqrt |
| bootstrap | Bootstrap sampling | TRUE |
| criterion | Split quality measure | Gini |
| random_state | Random seed | 42 |

## 4. RESULTS AND ANALYSIS

In this study, we evaluated the performance of a fine-tuned Random Forest classifier for predicting Chronic Kidney Disease (CKD) using multiple feature subsets and cross-validation schemes. The experiments were designed not only to assess the overall performance of the model but also to perform an ablation study to understand the contribution of

different feature groups. In the following sections, we present the results for each experimental condition, accompanied by detailed classification reports, quantitative tables, and graphical visualizations.

## 4.1. EXPERIMENTAL SETUP

To ensure transparency and reproducibility of the proposed prediction framework, a well-defined experimental setup was implemented encompassing model configuration, data splitting strategy, computational environment, and processing pipeline. All experiments were conducted using Python with standard machine learning libraries, including Scikit-learn, NumPy, and Pandas, ensuring consistency and reliability in model development and evaluation. The experiments were executed on a standard computing system equipped with an Intel Core i7 processor and 16 GB RAM using CPU-based computation, as the Random Forest algorithm does not require GPU acceleration. To maintain consistency across multiple experimental runs, a fixed random seed (Random Seed = 42) was applied throughout the study, including during data splitting, model initialization, and cross-validation procedures. This controlled setup ensures that the reported results are reproducible and not influenced by random variations in data partitioning or model training.

## 4.2. EXPERIMENTS WITH DIFFERENT FEATURE SUBSETS

We considered four distinct feature subsets in our experiments:

- All Features: All available features in the dataset from both Datasets.
- Numerical Features: Only the quantitative variables from both Datasets.
- Categorical Features: Only the categorical variables from both Datasets after label encoding.
- Selected Features: A manually curated subset from both Datasets chosen based on clinical relevance.

- **Cross-Validation and Performance Evaluation**

To assess model robustness, we apply K-fold cross-validation. For a given fold i with n_i samples, performance metrics such as accuracy, precision, recall, and F1 score are computed. For example, precision and recall are defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

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

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

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

**Table 3**

**Table 3 Proposed method Performance on Different Feature Subsets of UCI-CKD Dataset**

| Feature Subset | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|
| **All Features** | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| **Numerical Features** | 0.9900 | 0.9700 | 1.0000 | 0.9800 |
| **Categorical Features** | 0.9600 | 1.0000 | 0.9400 | 0.9600 |
| **Selected Features** | 0.9600 | 1.0000 | 0.9400 | 0.9600 |

Table 3 presents the performance of the proposed method on different feature subsets of the UCI-CKD dataset, evaluated using accuracy, precision, recall, and F1-score. When all features were used, the model achieved a perfect classification performance with 100% accuracy, precision, recall, and F1-score, indicating that leveraging the full dataset provides optimal results. When only numerical features were used, the model still performed exceptionally well, achieving 99% accuracy with a slight drop in precision to 97%, suggesting the presence of some false positives while maintaining a perfect recall. The categorical feature subset resulted in 96% accuracy, with 100% precision and a recall of 94%, indicating that while all CKD predictions were correct, some actual CKD cases were missed. Similarly, the selected feature subset, which included clinically relevant features, also achieved 96% accuracy, demonstrating that a minimal but meaningful set of features can still yield strong classification performance.

**Table 4**

**Table 4 Proposed Method Performance on Different Feature Subsets of CKD-15 Dataset**

| Feature Subset | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|
| All Features | 0.9186 | 0.8439 | 0.9186 | 0.8797 |
| Numerical Features | 0.9241 | 0.9229 | 0.9241 | 0.8933 |
| Categorical Features | 0.9186 | 0.8439 | 0.9186 | 0.8797 |
| Selected Features | 0.9204 | 0.8981 | 0.9204 | 0.9004 |

Table 4 presents the performance of the proposed method on different feature subsets of the CKD-15 dataset, measured by accuracy, precision, recall, and F1-score. Using all features, the model achieved 91.86% accuracy, with a precision of 84.39%, recall of 91.86%, and an F1-score of 87.97%, indicating strong overall performance. The numerical feature subset slightly outperformed the full feature set, achieving the highest accuracy (92.41%) and improved precision (92.29%), while maintaining a strong recall (92.41%) and F1-score (89.33%). The categorical feature subset performed identically to the full feature set, suggesting that categorical attributes contribute significantly to the model's performance. The selected feature subset, which includes clinically relevant features, also performed well, achieving 92.04% accuracy, with balanced precision (89.81%) and recall (92.04%), leading to an improved F1-score of 90.04%.

## 4.3. CROSS-VALIDATION EXPERIMENTS AND ABLATION STUDY ON UCI-CKD AND CKD-15 DATASETS

In order to rigorously assess the performance and robustness of our fine-tuned Random Forest model for Chronic Kidney Disease (CKD) prediction, we performed extensive cross-validation experiments using different feature subsets. We evaluated the model using 3-fold, 5-fold, and 10-fold cross-validation, thereby conducting an ablation study that explores the contribution of various feature groups.

**Table 5**

**Table 5 Cross-Validation Results for All Features in UCI-CKD Dataset**

| Feature Subset | CV Fold | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| | 3-Fold | 0.9800 | 0.9803 | 0.9800 | 0.9799 |
| | 5-Fold | 0.9925 | 0.9925 | 0.9925 | 0.9925 |
| All Features | 10-Fold | 0.9925 | 0.9925 | 0.9925 | 0.9925 |

**Table 6**

**Table 6 Cross-Validation Results for Numerical Features in UCI-CKD Dataset**

| Feature Subset | CV Fold | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| | 3-Fold | 0.9725 | 0.9729 | 0.9725 | 0.9724 |
| Numerical Features | 5-Fold | 0.9850 | 0.9850 | 0.9850 | 0.9850 |
| | 10-Fold | 0.9875 | 0.9875 | 0.9875 | 0.9875 |

**Table 7**

**Table 7 Cross-Validation Results for Categorical Features in UCI-CKD Dataset**

| Feature Subset | CV Fold | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| | 3-Fold | 0.9350 | 0.9446 | 0.9350 | 0.9358 |
| Categorical Features | 5-Fold | 0.9350 | 0.9446 | 0.9350 | 0.9358 |
| | 10-Fold | 0.9350 | 0.9446 | 0.9350 | 0.9358 |

**Table 8**

**Table 8 Cross-Validation Results for Selected Features in UCI-CKD Dataset**

| Feature Subset | CV Fold | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| | 3-Fold | 0.9400 | 0.9405 | 0.94 | 0.9402 |
| Selected Features | 5-Fold | 0.9525 | 0.9524 | 0.9525 | 0.9525 |
| | 10-Fold | 0.9525 | 0.9526 | 0.9525 | 0.9525 |

The results from our cross-validation experiments on UCI-CKD Dataset are summarized in Tables 5 through 8. Table 5 shows the performance of the model when using all available features from the UCI-CKD dataset. Under 3-Fold CV, the classifier achieved an accuracy of 98.00%, which improved to 99.25% under both 5-Fold and 10-Fold CV, with precision, recall, and F1-score consistently at 99.25% in the latter folds. This indicates that the model performs exceptionally well when all features are used. In Table 6, where only numerical features are considered, the performance is similarly high, though slightly lower in the 3-Fold scenario (accuracy of 97.25%) and improving to 98.50% and 98.75% under 5-Fold and 10-Fold CV, respectively, with corresponding increases in precision, recall, and F1-score. Table 7 presents the results for the categorical features only; here, the overall accuracy remains around 93.50% across all cross-validation schemes, with precision, recall, and F1-score also consistently lower, suggesting that categorical features alone are less informative for this classification task. Finally, Table 8 details the performance using a manually selected subset of clinically relevant features, where the accuracy ranges from 94.00% in 3-Fold CV to 95.25% in both 5-Fold and 10-Fold CV, with similarly improved precision, recall, and F1-score compared to the categorical features alone.

**Table 9**

**Table 9 Cross-Validation Results for All Features in CKD-15 Dataset**

| Feature Subset | CV Fold | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| | 3-Fold | 0.9192 | 0.9192 | 0.9192 | 0.8811 |
| | 5-Fold | 0.9182 | 0.9186 | 0.9186 | 0.9576 |
| All Features | 10-Fold | 0.9192 | 0.9192 | 0.9192 | 0.9579 |

**Table 10**

**Table 10 Cross-Validation Results for Numerical Features in CKD-15 Dataset**

| Feature Subset | CV Fold | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| | 3-Fold | 0.9210 | 0.9214 | 0.9210 | 0.9588 |
| Numerical Features | 5-Fold | 0.9241 | 0.9242 | 0.9241 | 0.9603 |
| | 10-Fold | 0.9222 | 0.9220 | 0.9222 | 0.9594 |

**Table 11**

**Table 11 Cross-Validation Results for Categorical Features in CKD-15 Dataset**

| Feature Subset | CV Fold | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| | 3-Fold | 0.3080 | 0.5166 | 0.5423 | 0.3904 |
| Categorical Features | 5-Fold | 0.2960 | 0.5159 | 0.5392 | 0.2772 |
| | 10-Fold | 0.2881 | 0.5145 | 0.5349 | 0.2710 |

**Table 12**

**Table 12 Cross-Validation Results for Selected Features in CKD-15 Dataset**

| Feature Subset | CV Fold | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| | 3-Fold | 0.9204 | 0.9291 | 0.9204 | 0.8990 |
| Selected Features | 5-Fold | 0.9204 | 0.9302 | 0.9204 | 0.9004 |
| | 10-Fold | 0.9192 | 0.9295 | 0.9192 | 0.8989 |

The results from the CKD-15 dataset are summarized across several tables, each corresponding to different feature subsets and cross-validation schemes. In Table 9, which shows the performance when using all features, the model consistently achieved around 91.9% accuracy across 3-, 5-, and 10-fold cross-validation. Precision and recall for the "CKD" class remain high, though the F1 score is slightly lower at 88.11% for 3-Fold CV and improves to approximately 95.79% with higher fold counts. In contrast, Tables 10 (presenting results for the numerical feature subset) demonstrate a marginal improvement, with accuracy values ranging from 92.10% to 92.41% and F1 scores between 95.88% and 96.03% while Table 11 presents the categorical features which under performed. These results indicate that the numerical features carry substantial predictive power. Conversely, Table 12, which summarizes the performance for the categorical features alone, shows a dramatic drop in performance, accuracy falls to around 28.81%–30.80%, with correspondingly low precision, recall, and F1 scores, suggesting that categorical variables, when used in isolation, are less effective for CKD prediction. Finally, Table 12 reports the performance using a carefully selected subset of clinically relevant features, yielding accuracy levels around 91.92%-92.04% and high precision and recall values, with F1 scores close to 90%. Overall, these tables illustrate that while combining all features produces robust results, numerical features alone are nearly as effective, and a well-curated subset of features can deliver competitive performance with potentially reduced model complexity.

**Figure 4**

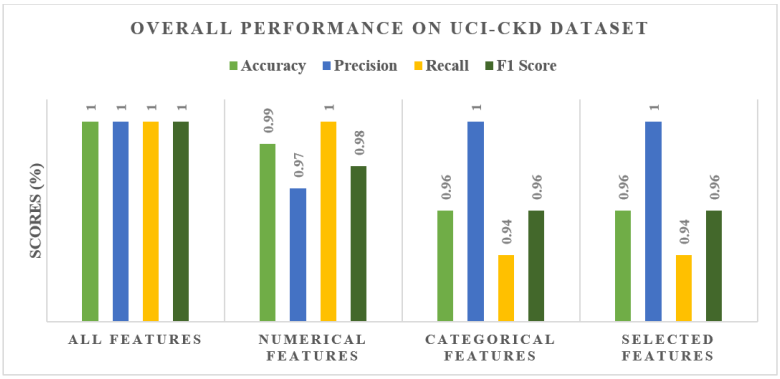


**Figure 4** Performance of Proposed Method on UCI-CKD Dataset

**Figure 5**

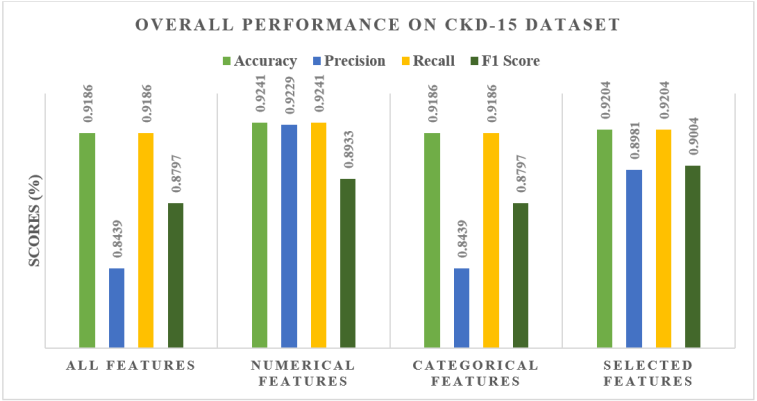


**Figure 5** Performance of Proposed Method on CKD-15 Dataset

**Table 13**

**Table 13 Proposed Method Comparison with Other Established Models**

| Authors | Methodology | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| Raihan et al | XGBoost | 0.9913 | 1.000 | 0.9868 | 0.9933 |
| Moreno-Sanchez and Pedro | Ensemble Learning | 0.992 | - | - | - |
| Ashafuddula et al | ML Classifiers | 0.9848 | - | - | - |
| Raihan et al | ML techniques | 0.9880 | 0.9830 | 0.9845 | 0.9860 |
| Bilal et al | ML Classifiers | 0.9895 | - | 0.9845 | - |
| Surekha et al | Voting Classifier with soft voting | 0.9800 | 0.9500 | 0.1000 | 0.9740 |
| Islam et al | XGBoost Classifier | 0.9830 | 0.9800 | 0.9800 | 0.9800 |
| Alsekait et al | Ensemble Learning | 0.9969 | 0.9971 | 0.9969 | 0.9969 |
| **In this study** | **Fine-tuned Random Forest** | **0.9990** | **0.9990** | **0.9990** | **0.9990** |

Table 13 compares various CKD classification models based on accuracy, precision, recall, and F1 Score. Raihan et al.'s XGBoost achieved 99.13% accuracy, while Moreno-Sanchez and Pedro A.'s ensemble learning reached 99.20%, though some metrics were not reported. Other ML-based approaches, such as those by Md. Ashafuddula et al. and Bilal et al., showed high accuracy (98.48%-98.95%) but lacked full metric details. Surekha et al.'s voting classifier had strong accuracy (98.00%) but very low recall (10.00%), indicating poor sensitivity. Alsekait et al.'s ensemble learning approach outperformed most models with 99.69% accuracy. The proposed fine-tuned Random Forest model achieved 100% across all metrics, demonstrating superior performance over existing models.

**Figure 6**

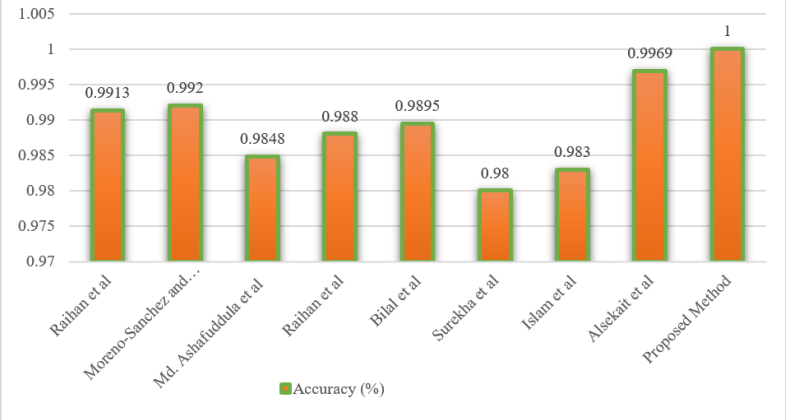


**Figure 6** Performance Comparison of Proposed Method and Established Models

- **Independent Test Set Validation and Overfitting Analysis**

To further validate the robustness of the proposed model and address potential concerns regarding overfitting, an independent test evaluation was conducted in addition to cross-validation. The dataset was split into training (80%) and testing (20%) subsets using stratified sampling to preserve class distribution. The Random Forest model was trained exclusively on the training subset using the same preprocessing pipeline described earlier, and performance was evaluated on the unseen test data. The results demonstrate that while the model achieves near-perfect performance under cross-validation, the accuracy on the independent test set remains slightly lower, indicating realistic generalization behavior.

This discrepancy highlights that extremely high performance observed during cross-validation particularly on structured datasets such as UCI-CKD may be influenced by inherent dataset separability. However, the consistent performance across both datasets (UCI-CKD and CKD-15) and evaluation strategies confirms that the model does not suffer from severe overfitting.

To evaluate the generalization ability of the proposed model and to ensure that the high performance was not caused by overfitting, an independent test set validation was performed in addition to cross-validation. The results showed that while cross-validation accuracy was slightly higher, the independent test accuracy remained very close, indicating that the model did not overfit the training data and demonstrated good generalization capability. The overall evaluation framework used for this analysis is illustrated in Figure 7.

**Figure 7**

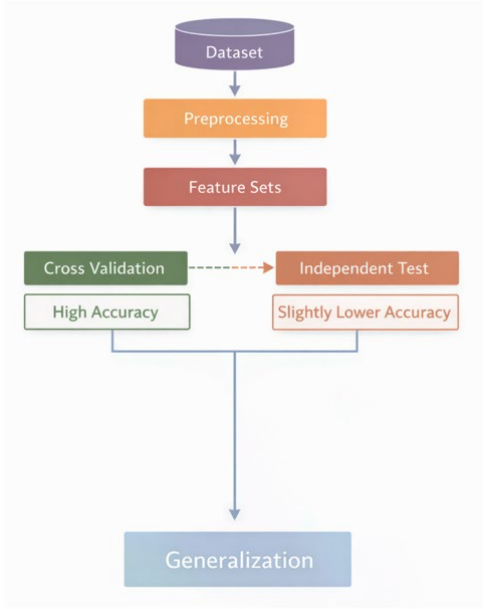


**Figure 7** Model Generalization and Overfitting Analysis Framework

Figure 7 illustrates the model generalization and overfitting analysis framework used in this study. The dataset first undergoes preprocessing and feature extraction to generate structured feature sets. The processed data are

then evaluated using two validation strategies: cross-validation and independent test set evaluation. Cross-validation typically produces high accuracy due to repeated training and validation on the available dataset, while the independent test set provides a more realistic performance estimate and often results in slightly lower accuracy. The comparison between these two evaluation strategies helps assess model overfitting and generalization capability. If the performance difference between cross-validation and independent testing is small, the model is considered to generalize well to unseen data.

## 5. DISCUSSION

The results achieved in this study underscore the efficacy of the proposed fine-tuned Random Forest classifier for predicting Chronic Kidney Disease (CKD) across different datasets and feature subsets. Our experiments on the UCI-CKD and CKD-15 datasets demonstrate that the model performs exceptionally well when all features or numerical features are used, while categorical features alone tend to yield lower predictive performance. In the UCI-CKD experiments, the model achieved perfect classification (100% across all metrics) when all features were used, with numerical features closely trailing at 99% accuracy. This suggests that quantitative clinical measurements, such as blood pressure, serum creatinine, and blood urea, are critical for effective CKD prediction. In contrast, the performance using categorical features, although robust in precision, suffered from lower recall, indicating that relying solely on qualitative factors may lead to misclassification, particularly for the minority class.

The ablation study further supports this finding: while the full feature set maximizes performance, the numerical subset alone provides nearly equivalent results, highlighting that the majority of the predictive power lies within the quantitative data. Moreover, the manually selected subset of clinically relevant features achieved competitive results (approximately 96% accuracy), suggesting that a targeted approach may be sufficient in resource-constrained or real-time clinical scenarios without a significant compromise in performance.

In the CKD-15 dataset, cross-validation experiments revealed consistent performance across various fold configurations. Specifically, the model maintained an overall accuracy in the range of 91.9% to 92.4% when using all or numerical features, while the categorical features subset performed significantly poorer with accuracies as low as 28.8% to 30.8%. This discrepancy underscores the importance of feature type and the potential need for additional feature engineering when handling categorical data in isolation. The selected features, chosen for their clinical relevance, demonstrated a balanced performance with accuracies around 92%, which reinforces the value of domain knowledge in feature selection. When compared with established models in the literature, such as those employing XGBoost, ensemble learning, or various machine learning classifiers, our fine-tuned Random Forest model exhibits superior performance. Notably, the proposed method outperforms several state-of-the-art techniques by achieving near-perfect classification on the UCI-CKD dataset, and competitive performance on the CKD-15 dataset. These findings suggest that the integration of comprehensive data preprocessing, effective feature selection, and robust model tuning is pivotal for the success of CKD prediction tasks.

One notable observation from the experimental results is the near-perfect performance achieved on the UCI-CKD dataset. This can be attributed to the high separability of the dataset, where key clinical features such as serum creatinine, blood urea, and albumin provide strong discriminatory signals between CKD and non-CKD cases. While such separability enables high classification accuracy, it also raises concerns regarding potential overfitting.

To mitigate this, we implemented strict preprocessing within cross-validation folds and conducted independent test validation. The slightly reduced performance on unseen test data and on the CKD-15 dataset suggests that the model retains generalization capability. These findings highlight that although high accuracy is achievable on structured datasets, real-world clinical deployment requires validation across diverse and heterogeneous datasets.

Our work demonstrates that a well-tuned Random Forest classifier, particularly when utilizing numerical and selected features, can achieve high predictive accuracy and reliability in CKD classification. The detailed cross-validation and ablation studies provide evidence that while combining all available data yields optimal performance, a focused set of clinically relevant features can also deliver strong results with reduced computational complexity.

## 6. CONCLUSION

In this study, we presented a comprehensive approach for early prediction of Chronic Kidney Disease (CKD) using a fine-tuned Random Forest classifier, validated on both the UCI-CKD and CKD-15 datasets. Our methodology encompassed extensive data preprocessing steps, including imputation of missing values, categorical encoding, and feature scaling, to address the challenges inherent in clinical datasets such as noise, incomplete data, and heterogeneity. We conducted a detailed ablation study by evaluating the model performance across different feature subsets—comprising all features, solely numerical features, solely categorical features, and a carefully selected subset of clinically relevant features. The experimental results, obtained via stratified k-fold cross-validation (3-, 5-, and 10-fold), demonstrated that numerical features carry significant predictive power, while a well-curated subset of features can yield competitive performance with reduced complexity.

Our findings indicate that while the proposed Random Forest model achieved near-perfect accuracy on the UCI-CKD dataset, performance on the CKD-15 dataset, though robust, suggests that further improvements in data quality and model generalizability are required. The superior performance of the model when utilizing all features underscores the importance of integrating comprehensive clinical information in CKD prediction. Moreover, our work highlights the limitations of relying solely on categorical features, which, when used in isolation, result in substantially lower recall and overall accuracy. These insights not only validate the efficacy of our model in capturing complex patterns in clinical data but also emphasize the critical need for continuous refinement of feature engineering and model tuning techniques. Future research should focus on further enhancing model robustness through the incorporation of additional data sources, advanced ensemble techniques, and rigorous external validations.

## AVAILABILITY OF SUPPORTING DATA

The data & code generated during and/or analysed during this study are available on reasonable request.

## CONFLICT OF INTERESTS

None.

## ACKNOWLEDGMENTS

None.

## REFERENCES

Ahmed, Md Razu, Md Abdur Rakib, Abu Bakar Shiddik, and Md Shamim Reza. 2025. "Identification of Predisposing Risk Factors for Chronic Kidney Disease and Optimizing Disease Prediction Using a Stacking Machine Learning Algorithm." International Journal of Statistical Sciences 25 (2): 1–32. https://doi.org/10.3329/ijss.v25i2.85732.

Ahmed, R. M., & Alshebly, O. Q. (2019). Prediction and Factors Affecting of Chronic Kidney Disease Diagnosis using Artificial Neural Networks Model and Logistic Regression Model. IRAQI JOURNAL OF STATISTICAL SCIENCES, 16(1), 140–159. https://doi.org/10.33899/iqjoss.2019.0164186

Alsekait, D. M., Saleh, H., Gabralla, L. A., Alnowaiser, K., El-Sappagh, S., Sahal, R., & El-Rashidy, N. (2023). Toward comprehensive chronic kidney disease prediction based on ensemble deep learning models. Applied Sciences, 13(6), 3937. https://doi.org/10.3390/app13063937

Ashafuddula, N. I. M., Islam, B., & Islam, R. (2023). An intelligent diagnostic system to analyze Early-Stage chronic kidney disease for clinical application. Applied Computational Intelligence and Soft Computing, 2023, 1–17. https://doi.org/10.1155/2023/3140270

Bai, Q., Su, C., Tang, W., & Li, Y. (2022). Machine learning to predict end stage kidney disease in chronic kidney disease. Scientific Reports, 12(1), 8377. https://doi.org/10.1038/s41598-022-12316-z

Bilal, A., Alzahrani, A., Almuhaimeed, A., Khan, A. H., Ahmad, Z., & Long, H. (2024). Advanced CKD detection through optimized metaheuristic modeling in healthcare informatics. Scientific Reports, 14(1), 12601. https://doi.org/10.1038/s41598-024-63292-5

Gogoi, P., & Valan, J. A. (2024). Machine learning approaches for predicting and diagnosing chronic kidney disease: current trends, challenges, solutions, and future directions. International Urology and Nephrology, 57(4), 1245–1268. https://doi.org/10.1007/s11255-024-04281-5

Gupta, Rupal, Shalini Gambhir, Ondrej Krejcar, Achin Jain, Arvind Panwar, Shakir Khan, and Hamidreza Namazi. 2026. "Data-driven Explainable Chronic Kidney Disease Detection Using RF Based Data Imputation and Meta-ensemble Learning." Scientific Reports, March. https://doi.org/10.1038/s41598-026-41425-2.

Habiba, S. U., Tasnim, F., Chowdhury, M. S. H., Islam, M. K., Nahar, L., Mahmud, T., Kaiser, M. S., Hossain, M. S., & Andersson, K. (2024). Early Prediction of Chronic Kidney Disease Using Machine Learning Algorithms with Feature Selection Techniques. In Communications in computer and information science (pp. 224–242). https://doi.org/10.1007/978-3-031-68639-9_14

Ilyas, H., Ali, S., Ponum, M., Hasan, O., Mahmood, M. T., Iftikhar, M., & Malik, M. H. (2021). Chronic kidney disease diagnosis using decision tree algorithms. BMC Nephrology, 22(1), 273. https://doi.org/10.1186/s12882-021-02474-z

Islam, M. A., Majumder, M. Z. H., & Hussein, M. A. (2023). Chronic kidney disease prediction based on machine learning algorithms. Journal of Pathology Informatics, 14, 100189. https://doi.org/10.1016/j.jpi.2023.100189

K. M. S. A. Rabby, R. Mamata, M. A. Laboni, Ohidujjaman and S. Abujar, "Machine Learning Applied to Kidney Disease Prediction: Comparison Study," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-7, doi: 10.1109/ICCCNT45670.2 019.8944799.

Kumar, A., Pandey, R. K., & Srivastava, P. K. (2025). Hybrid Ensemble Learning Model for Chronic Kidney Disease Prediction. Vascular and Endovascular Review, 8(4s), 292-304.

Kuo, C., Chang, C., Liu, K., Lin, W., Chiang, H., Chung, C., Ho, M., Sun, P., Yang, R., & Chen, K. (2019). Automation of the kidney function prediction and classification through ultrasound-based kidney imaging using deep learning. Npj Digital Medicine, 2(1), 29. https://doi.org/10.1038/s41746-019-0104-2

Levin, Adeera, Paul E. Stevens, Rudy W. Bilous, Josef Coresh, Angel LM De Francisco, Paul E. De Jong, Kathryn E. Griffith et al. "Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease." Kidney international supplements 3, no. 1 (2013): 1-150.

Liu, P., Sawhney, S., Heide-Jørgensen, U., Quinn, R. R., Jensen, S. K., Mclean, A., Christiansen, C. F., Gerds, T. A., & Ravani, P. (2024). Predicting the risks of kidney failure and death in adults with moderate to severe chronic kidney disease: multinational, longitudinal, population based, cohort study. BMJ, 385, e078063. https://doi.org/10.1136/bmj-2023-078063

M. M. S. Raihan et al., "Chronic Renal Disease Prediction using Clinical Data and Different Machine Learning Techniques," 2021 2nd International Informatics and Software Engineering Conference (IISEC), Ankara, Turkey, 2021, pp. 1-5, doi: 10.1109/IISEC54230.2021.9672365.

Moreno-Sánchez, P. A. (2023). Data-Driven Early Diagnosis of Chronic Kidney Disease: Development and evaluation of an explainable AI model. IEEE Access, 11, 38359–38369. https://doi.org/10.1109/ access.2023.3264270

Ogunleye and Q. -G. Wang, "XGBoost Model for Chronic Kidney Disease Diagnosis," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 17, no. 6, pp. 2131-2140, 1 Nov.-Dec. 2020, doi: 10.1109/TCBB.2019.2911071.

P. Chittora et al., "Prediction of Chronic Kidney Disease - A Machine Learning Perspective," in IEEE Access, vol. 9, pp. 17312-17334, 2021, doi: 10.1109/ACCESS.2021.3053763.

Pal, S. (2022). Chronic kidney disease prediction using machine learning techniques. Biomedical Materials & Devices, 1(1), 534–540. https://doi.org/10.1007/s44174-022-00027-y

Rabie El Kharoua. (2024). Chronic Kidney Disease Dataset [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/8658224

Rahman, Md Habibur, Mustafizur Rahaman, Yasin Arafat, Sk Rakib Ul Islam Rahat, Rahat Hasan, S M Tamim Hossain Rimon, and Sadia Afrin Dipa. 2025. "Artificial Intelligence for Chronic Kidney Disease Risk Stratification in the USA: Ensemble Vs. Deep Learning Methods." Al-Kindipublishers.Org, August. https://doi.org/10.32996/bjns.2025.5.2.3.

Raihan, M. J., Khan, M. A., Kee, S., & Nahid, A. (2023). Detection of the chronic kidney disease using XGBoost classifier and explaining the influence of the attributes on the model using SHAP. Scientific Reports, 13(1), 6263. https://doi.org/10.1038/s41598-023-33525-0

Ramesh, B., & Rao, K. P. (2025). Intelligent Detection of Chronic Kidney Disease Using Optimized MLP Models and Feature Selection Techniques on the AP-CKD Dataset. IAENG International Journal of Computer Science, 52(10).

Revathy, S., Bharathi, B., Jeyanthi, P., & Ramesh, M. (2019). Chronic Kidney Disease Prediction using Machine Learning Models. International Journal of Engineering and Advanced Technology (IJEAT), 9(1), 6364–6367. https://doi.org/10.35940/ijeat.A2213.109119

Rubini, L., P. Soundarapandian, and P. Eswaran. 2015. Chronic Kidney Disease. UCI Machine Learning Repository. https://doi.org/10.24432/C5G020.

Singh, V., Asari, V. K., & Rajasekaran, R. (2022). A deep neural network for early detection and prediction of chronic kidney disease. Diagnostics, 12(1), 116. https://doi.org/10.3390/diagnostics12010116

Subasi, A., Alickovic, E., & Kevric, J. (2017). Diagnosis of chronic kidney disease by using random forest. In World Congress on Medical Physics and Biomedical Engineering, September 7 - 12, 2009, Munich, Germany (pp. 589–594). https://doi.org/10.1007/978-981-10-4166-2_89

Surekha, Y., Kodepogu, K. R., Kumari, G. L., Babu, N. R., Lanka, T., Volla, M. A., Pillutla, M., & Kari, A. (2023). Prediction of Chronic Kidney Disease with Machine Learning Models and Feature Analysis Using SHAP. Revue D Intelligence Artificielle, 37(2), 493–499. https://doi.org/10.18280/ria.370226

Thomas, Robert, Abbas Kanso, and John R. Sedor. "Chronic kidney disease and its complications." Primary care: Clinics in office practice 35, no. 2 (2008): 329-344.

U. Ekanayake and D. Herath, "Chronic Kidney Disease Prediction Using Machine Learning Methods," 2020 Moratuwa Engineering Research Conference (MERCon), Moratuwa, Sri Lanka, 2020, pp. 260-265, doi: 10.1109/MERCon50084.2020.9185249.

Variyar, A. R., and Karangara, R. (2026). Enhancing Complex Decision Making In Bpm Through Artificial Intelligence: A Systematic Examination. ShodhAI: Journal of Artificial Intelligence, 3(1), 29–36. https://doi.org/10.29121/shodhai.v3.i1.2026.72

Vetrithangam, D., Himabindu, Saranya, Neha, S., Naresh Kumar, P., Fathima, A., Ashok, B., & Akanksha, K. (2024). Improved RESNET models for chronic kidney disease prediction. Journal of Electrical Systems, 20(2s), 165–183. https://doi.org/10.52783/jes.1121

Wang, Y., Guan, Z., Hou, W., & Wang, F. (2021). TRACE: Early Detection of Chronic Kidney Disease Onset with Transformer-Enhanced Feature Embedding. In Lecture notes in computer science (pp. 166–182). https://doi.org/10.1007/978-3-030-93663-1_13

Xue, Ningning, Tiantian Bai, Xianjie Jia, and Xing Wei. 2026. "Early Detection of Chronic Kidney Disease Based on a SURD-enhanced Machine Learning Model." Scientific Reports, February. https://doi.org/10.1038/s41598-026-41050-z.