PARTICLE SWARM OPTIMIZATION-BASED FEATURE SELECTION FOR CHRONIC KIDNEY DISEASE PREDICTION

Adarsh Maurya [1], Arpit Pathak [2], and Bhanu Priya [3]

[1]Department of Botany, Banaras Hindu University, Varanasi, UP, India.
[2]Computer Science and Engineering, Indian Institute of Information Technology Bhagalpur, Bihar, India.
[3]Electronics and Communication Engineering, Indian Institute of Information Technology Bhagalpur, Bihar, India.

ABSTRACT

Chronic Kidney Disease (CKD) is a progressive medical condition that requires early detection to prevent severe health complications. In this work, we present a machine learning framework for CKD prediction that integrates feature selection and classification models. The UCI CKD dataset was utilized, and preprocessing steps included handling missing values, encoding categorical variables, and normalization. To reduce dimensionality and select the most relevant features, Particle Swarm Optimization (PSO) was applied, which significantly improved model efficiency. Multiple machine learning algorithms, including Random Forest (RF), Gradient Boosting (GB), Support Vector Machine (SVM), Logistic Regression (LR), and Artificial Neural Networks (ANN), were evaluated using 5-fold, 10-fold, and 15-fold cross-validation. The experimental results demonstrate that ensemble methods achieved the best performance after feature selection. Random Forest and Gradient Boosting consistently outperformed other models, achieving accuracies above 99% with near-perfect precision, recall, and F1-scores, while maintaining balanced classification between CKD and non-CKD cases. These results highlight the effectiveness of feature selection in improving diagnostic accuracy and confirm the superiority of ensemble learning for CKD prediction. The proposed approach provides a reliable and efficient tool that can assist healthcare professionals in early CKD detection and decision-making, ultimately contributing to better patient outcomes.

KEYWORDS

Chronic Kidney Disease · Particle Swarm Optimization, ·Machine Learning , Explainable AI.

1. INTRODUCTION

Chronic Kidney Disease (CKD) is a progressive condition characterized by the gradual loss of kidney function, often without overt symptoms until the damage becomes substantial. It poses a significant global health challenge, largely because many individuals remain unaware of its onset, and by the time symptoms appear, treatment becomes more complex and costly. Globally, CKD affects hundreds of millions of people. The Global Burden of Disease (GBD) [1] studies estimate that nearly 700 million individuals live with CKD. When combined with Acute Kidney Injury (AKI) and kidney failure, the total burden rises to around 850 million people worldwide. Between 1990 and 2017, the prevalence of CKD rose by about 33%, indicating a trend of increasing risk. Furthermore, CKD is disproportionately concentrated in low- and middle-income countries, where detection and treatment resources are often limited. In the United States [2], about 14% of adults—approximately 35.5 million people—have CKD at stages 1–4, as shown in Figure 2.

Bibhu Dash et al: ICFCSAI-2025
pp. 219-230, 2026. IJCI – 2026 DOI:10.5121/ijci.2026.150123

Alarmingly, nearly 9 in 10 affected individuals are unaware of their condition. The median global prevalence [3] is estimated at 9.5% (interquartile range 5.9–11.7%), with wide regional variations. Because CKD often remains undiagnosed until advanced stages, early detection is crucial to slow progression toward kidney failure, reduce associated comorbidities—especially cardiovascular disease—and alleviate healthcare costs. This urgency has motivated the increasing use of computational techniques, particularly machine learning (ML), to predict CKD from clinical and laboratory data. Title Suppressed Due to Excessive Length

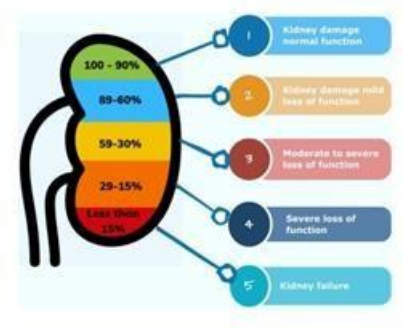

Figure.1. Different Stages of CKD

Several research efforts have recently applied machine learning algorithms to CKD detection and prediction. Rezk et al. [4] proposed an explainable AI framework combining GAN-based data imputation with few-shot learning and tested multiple classifiers such as SVM, Logistic Regression, Decision Trees, and ensemble methods. Iliyas et al. [5] (2025) surveyed CKD studies conducted between 2021 and 2024, highlighting limitations in feature selection, interpretability, and generalizability. Halder et al. [6] (2024) developed ML-CKDP, focusing on robust preprocessing and algorithm comparison to enhance prediction accuracy. Kovesdy et al. [7] updated global prevalence estimates, showing that CKD stages 3–5 affect 4.7% of men and 5.8% of women aged 20 and above. Jadoul et al. [8] (2024) reported median prevalence figures of around 9.5%, emphasizing treatment gaps even in regions where healthcare resources are available. Tangri et al. [9] (2024) introduced risk models predicting CKD progression using routinely collected laboratory data, demonstrating strong predictive performance. Despite these advances, several gaps remain. Many studies rely on all available features or use simple feature selection techniques, with few leveraging meta-heuristic optimization methods such as Particle Swarm Optimization (PSO). The balance between predictive accuracy and interpretability is often underexplored, leaving clinicians with models that are either highly accurate but opaque or transparent but less reliable. Furthermore, generalizability is limited, as many models are built on small or region-specific datasets without rigorous handling of missing or noisy data. Comprehensive pipelines that integrate feature optimization, algorithm comparison, and interpretability remain relatively scarce. This study addresses these gaps through a threefold contribution. First, PSO is employed to select the most relevant subset of features, reducing redundancy and improving both performance and efficiency. Second, multiple machine learning algorithms are evaluated to identify the most effective predictive models, enhancing generalizability across different settings. Third, SHAP (SHapley Additive exPlanations) is applied to provide both global and local interpretability, enabling clinicians to understand not only which features are most important overall but also why a specific prediction was made. This integrated pipeline—PSO-based feature selection, multi-model comparison, and SHAP-driven interpretability—has the potential to improve early CKD detection, reduce diagnostic costs, and foster clinical trust in AI-based decision support systems.

2. RELATED WORK

Several researchers have investigated the use of machine learning methods for the early detection of Chronic Kidney Disease (CKD). For instance, Islam et al. [10] proposed a machine learning approach that utilized both numerical (24) and categorical (13) attributes. To identify the most informative variables, they applied Principal Component Analysis (PCA), followed by classification using XGBoost. Their model achieved an impressive accuracy of 99.16%, highlighting the effectiveness of dimensionality reduction in CKD prediction. Similarly, Sawhney et al. [11] explored artificial intelligence–based prediction by combining feature extraction and feature selection strategies with a multilayer perceptron. The Artificial Neural Network (ANN) they developed demonstrated superior predictive performance compared to traditional classifiers. Alsekait et al. [12] presented an ensemble deep learning architecture for CKD prediction using the UCI dataset with 400 patient cases. Their framework incorporated preprocessing steps such as label encoding and outlier removal, followed by feature selection techniques including mutual information and recursive feature elimination. They combined RNN, LSTM, and GRU networks in a stacked design, achieving strong predictive results with precision, recall, F1-score, and accuracy all approaching 99.69%. In another study, Arif et al. [13] integrated Boruta feature selection, advanced preprocessing techniques (e.g., Min-Max scaling, robust scaling, and iterative imputation), and hyperparameter tuning into a CKD prediction pipeline. Using the K-Nearest Neighbor (KNN) classifier with grid search optimization, they reported accuracy levels as high as 100% on the UCI dataset. Poonia et al. [14] developed a feature learning framework in which multiple models were evaluated on the UCI dataset (400 cases, 24 features). Their work emphasized the role of feature selection through Chi-Square testing, achieving 98.75% accuracy when combined with Logistic Regression. Similarly, Pal [15] employed an ensemble learning approach using bagging techniques over multiple machine learning classifiers. Among them, the Decision Tree model performed best, with accuracy close to 97.23%. Almustafa et al. [16] also examined CKD classification using the UCI dataset. Their evaluation across several machine learning algorithms showed that Decision Trees provided the strongest performance (99% accuracy). They further demonstrated that incorporating feature selection with Naïve Bayes and KNN classifiers enhanced predictive outcomes. Other contributions further highlight the role of machine learning in CKD detection. Ilyas et al. [17] built a Decision Tree–based diagnostic system and applied Random Forest to classify patients into five CKD stages. Their method was particularly effective in the early stages, achieving performance levels up to 98%. Halder et al. [6] developed a web application for CKD screening, incorporating preprocessing steps such as missing value imputation, normalization, and feature selection. Their system improved upon existing classifiers in terms of accuracy, precision, and recall. Similarly, Uddin et al. [18] conducted a comparative analysis of various machine learning models on the UCI dataset. By applying preprocessing techniques and selecting relevant features, they found that ensemble methods such as Random Forest and AdaBoost offered the best predictive accuracy.

Figure.2. Work flow diagram of proposed methodology.

3. DATA SET DESCRIPTION

In this study, we employed the Chronic Kidney Disease (CKD) dataset from the UCI Machine Learning Repository, which consists of 400 patient records (250 CKD cases and 150 non-CKD cases) described by 24 clinical and physiological attributes. These features include both numerical measures, such as age, blood pressure, hemoglobin, blood urea, and serum creatinine, as well as categorical indicators such as the presence of anemia, diabetes mellitus, hypertension, and pedal edema. Since the dataset reflects real clinical records, it contains missing values that require preprocessing through imputation and encoding to ensure data consistency and reliability. The UCI CKD dataset is widely used in machine learning research because it provides a balanced combination of medical and demographic attributes, making it a reliable benchmark for developing and evaluating predictive models aimed at early CKD detection.

3.1 Data Preprocessing

Before feature selection, the CKD dataset required several preprocessing steps to ensure data quality and consistency. Since the dataset contained missing values, different strategies were applied depending on the type of feature. For categorical attributes, missing entries were replaced using random value imputation in order to preserve the original distribution of categories. For numerical attributes, missing values were handled through iterative imputation, which estimates each missing value by considering its relationships with other features in the dataset. Once the missing values were addressed, all categorical variables were converted into numerical form through encoding to make them suitable for machine learning algorithms. Finally, normalization was applied to scale the numerical features into a common range, preventing attributes with larger values from dominating the training process. These preprocessing steps collectively prepared the dataset for reliable feature selection and accurate prediction.

3.2 Feature selection

To identify the most relevant subset of features for CKD prediction, we employed Particle Swarm Optimization (PSO), a population-based metaheuristic inspired by the social behavior of birds flocking or fish schooling. In PSO, each particle represents a candidate solution, that is, a subset of features, and moves through the search space by updating its velocity and position based on its own experience as well as the experience of neighboring particles. The objective of the optimization process is to maximize classification accuracy while simultaneously minimizing the number of selected features, thereby achieving an optimal balance between predictive performance and model complexity.

Each particle updates its velocity and position according to the following equations:

$$v_i^{t+1} = w \cdot v_i^t + c_1 \cdot r_1 \cdot (pbest_i - x_i^t) + c_2 \cdot r_2 \cdot (gbest - x_i^t) \quad (1)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (2)$$

the inertia weight www controls the influence of the particle's previous velocity on its current movement. The parameters c1c_1c1 and c2c_2c2 are acceleration coefficients that regulate the impact of the particle's personal best position and the global best position, respectively. The variables r1r_1r1 and r2r_2r2 are random numbers uniformly distributed within the interval [0,1][0,1][0,1]. The term pbestipbest_ipbesti denotes the best solution found by particle iii, while

g best g best g best represents the best global solution identified by the swarm thus far.

$$\text{Fitness} = \alpha(1 - \text{Accuracy}) + \beta\frac{|S|}{|F|} \quad (3)$$

where Accuracy is the classification accuracy of the model with selected features, |S| is the number of selected features, |F|is the total number of features, and α, βare weigh ting factors balancing accuracy and feature subset size. [!htb] Particle Swarm Optimization for Feature Selection Input :Population size N, number of iterations T, dataset D with n features Output: Optimal binary vector Xg best[1] Initialize a swarm of particles with random binary vectors of length n Initialize velocity vi randomly for each particle each p article I Evaluate fitness using a classification model Set p best i ←xi Set g best as the best pbestiin the swarm each iteration t= 1to Teach particle I each feature index j Generater and om number sr 1,r2∈[0,1] Update velocity:

$$v_{i,j}^{t+1} = w \cdot v_{i,j}^{t} + c_1 \cdot r_1 \cdot (pbest_{i,j} - x_{i,j}^{t}) + c_2 \cdot r_2 \cdot (gbest_j - x_{i,j}^{t})$$

Apply thres holding : ifxt+1>0.5, setxi,j=1;otherwise,setxi,j=0 Recalculate fitness for updated particle i fitness (xi) better than fitness (pbesti) Update pbesti← xi Update gbestas the best pbesti among all particles Return the binary solution vector of the best - performing particle best.

3.3 Machine learning models

Once the most informative features were selected, different machine learning models were trained and evaluated for CKD prediction. These models were chosen because they represent a mix of ensemble methods, statistical techniques, and neural network approaches, thereby providing a comprehensive view of how different algorithmic families perform on structured medical data. Each model offers distinct advantages when applied to datasets such as the CKD dataset.

Random Forest (RF)

Random Forest (RF) [19] is an ensemble learning method that constructs multiple decision trees and combines their predictions. By aggregating the outputs of many trees, it reduces the risk of over fitting and enhances overall robustness. This characteristic makes Random Forest particularly well-suited for medical datasets containing both numerical and categorical features, such as the CKD dataset.

Gradient Boosting (GB)

Gradient Boosting (GB) [20] is another tree-based ensemble technique. Unlike Random Forest, which builds trees independently, Gradient Boosting constructs trees sequentially. Each new tree is trained to correct the errors made by the previous ones. This iterative refinement process enables Gradient Boosting to achieve high predictive accuracy, especially in structured classification problems. Its built-in regularization mechanisms also help control over fitting, making it a strong candidate for CKD prediction.

Support Vector Machine (SVM)

Support Vector Machine (SVM) [21] aims to identify the optimal decision boundary that maximizes the margin between different classes. Through the use of kernel functions, SVM can model non-linear relationships, making it effective when the relationship between patient attributes and CKD status is complex and not strictly linear.

Logistic Regression (LR)

Logistic Regression (LR) [22] is a statistical classification model that estimates the probability of a binary outcome using a logistic function. Although it is less complex than ensemble or neural network models, its primary advantage lies in its interpretability. Clinicians and researchers can clearly understand the contribution of individual features to the final prediction, which is highly valuable in healthcare applications.

Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) [23] are inspired by the structure of the human brain and consist of interconnected layers of neurons. They are capable of capturing complex, non-linear patterns in data. In CKD classification, ANNs can adapt to variations in patient data and achieve competitive performance compared to tree-based models. However, they typically require greater computational resources and careful hyper parameter tuning. In summary, these models were selected to balance predictive power and interpretability. Ensemble methods such as Random Forest and Gradient Boosting offer high accuracy and robustness. Support Vector Machine performs well in complex feature spaces, while Logistic Regression provides simplicity and transparency. Artificial Neural Networks offer flexibility in capturing non-linear dependencies. Together, these models provide a comprehensive and well-rounded evaluation of CKD classification performance.

3.4 Performance evaluation

To evaluate the effectiveness of the classification models ,several standard performance metrics were used [24]. These include *accuracy*, *precision*, *recall*, and the *F1-score*.

4. RESULTS

Before applying feature selection, multiple machine learning models were evaluated using 5-fold, 10-fold, and 15-fold cross-validation. The results indicate that Random Forest and Gradient Boosting consistently outperformed the other models across all validation settings. Both methods achieved very high accuracy, precision, recall, and F1-scores, demonstrating their strong capability to capture complex patterns within the dataset. For instance, Random Forest achieved approximately 98.8% accuracy in 10-fold cross-validation, while Gradient Boosting produced comparable results, making them the most reliable models at this stage.

The Support Vector Machine (SVM) model also delivered strong performance, with accuracy and F1-scores in the range of 97–98%. Although competitive, its results were slightly lower than those of the ensemble methods. The Artificial Neural Network (ANN) achieved reasonably good performance as well, surpassing Logistic Regression but still not matching the accuracy of Random Forest and Gradient Boosting. This outcome may be attributed to the relatively small dataset size, as ANN models typically require larger datasets to reach optimal performance.

In contrast, Logistic Regression demonstrated the weakest performance among the evaluated models, with accuracy ranging between 94–95% and consistently lower precision, recall, and F1-scores. These findings suggest that simpler linear models are less effective for CKD prediction compared to more advanced ensemble and non-linear approaches.

Overall, the results before feature selection highlight that ensemble methods—particularly Random Forest and Gradient Boosting—are the most effective models, while SVM and ANN provide competitive but slightly weaker alternatives. Logistic Regression, despite its interpretability, struggles to achieve comparable predictive performance.

Table1. Performance analysis of individual models on different Cross-Validation before feature selection

Cross-Validation	Model	Accuracy	Precision	Recall	F1-Score
5-Fold	RandomForest	0.9829	0.9834	0.9829	0.9830
	GradientBoosting	0.9803	0.9808	0.9803	0.9804
	SVM	0.9779	0.9783	0.9783	0.9783
	LogisticRegression	0.9443	0.9470	0.9443	0.9445
	ANN	0.9736	0.9740	0.9736	0.9737
10-Fold	RandomForest	0.9879	0.9882	0.9879	0.9880
	GradientBoosting	0.9854	0.9859	0.9854	0.9855
	SVM	0.9779	0.9783	0.9779	0.9780
	LogisticRegression	0.9417	0.9442	0.9417	0.9419
	ANN	0.9729	0.9733	0.9729	0.9730
15-Fold	RandomForest	0.9829	0.9834	0.9829	0.9830
	GradientBoosting	0.9854	0.9859	0.9854	0.9855
	SVM	0.9779	0.9783	0.9779	0.9780
	LogisticRegression	0.9443	0.9470	0.9443	0.9445
	ANN	0.9685	0.9689	0.9685	0.9686

Random Forest in 10-fold and 15-fold cross-validation, while Random Forest demonstrated equally strong stability in the 5-fold validation setting. The SVM model also exhibited significant improvement after feature selection, achieving performance levels close to those of the ensemble models, with accuracy and F1-scores around 99.5%. This indicates that the reduced feature set enabled SVM to handle the dataset more efficiently. The ANN model maintained strong predictive capability as well, recording scores between 98.5% and 98.7% across different validation folds. These results reflect its ability to generalize effectively even after dimensionality reduction.

On the other hand, Logistic Regression continued to lag behind the more complex models. Although it benefited from feature selection, reaching approximately 96% accuracy and similar F1-scores across all folds, its overall performance remained comparatively weaker. This suggests that simpler linear models may not fully capture the complex feature interactions inherent in CKD data.

Overall, the findings demonstrate that feature selection not only reduces dimensionality but also enhances model efficiency and predictive performance. Ensemble methods such as Random Forest and Gradient Boosting remained the most effective approaches, while SVM significantly narrowed the performance gap. These results indicate that with optimized features, even non-ensemble methods can achieve near state-of-the-art performance.

After applying feature selection, confusion matrices were generated for the Random Forest (RF) and Gradient Boosting (GB) models using 5-fold, 10-fold, and 15-fold cross-validation. The results (Figure 5) show that both ensemble models achieved highly accurate predictions, with

only a minimal number of misclassifications across CKD and non-CKD classes.

In the 5-fold cross-validation setup, Random Forest correctly classified 398 out of 400 samples, while Gradient Boosting misclassified only three samples. In the 10-fold cross-validation setting, both models demonstrated even stronger generalization performance: Random Forest misclassified only one instance, and Gradient

Table 2. Performance analysis of individual models on different Cross-Validation after feature selection.

Cross-Validation	Model	Accuracy	Precision	Recall	F1-Score
5-Fold	Random Forest	0.9950	0.9950	0.9950	0.9950
	Gradient Boosting	0.9924	0.9925	0.9924	0.9924
	SVM	0.9950	0.9950	0.9950	0.9950
	Logistic Regression	0.9571	0.9588	0.9571	0.9573
	ANN	0.9874	0.9877	0.9874	0.9874
10-Fold	Random Forest	0.9975	0.9975	0.9975	0.9975
	Gradient Boosting	0.9975	0.9975	0.9975	0.9975
	SVM	0.9950	0.9950	0.9950	0.9950
	Logistic Regression	0.9571	0.9588	0.9571	0.9573
	ANN	0.9874	0.9877	0.9874	0.9874
15-Fold	Random Forest	0.9950	0.9950	0.9950	0.9950
	Gradient Boosting	0.9975	0.9975	0.9975	0.9975
	SVM	0.9950	0.9950	0.9950	0.9950
	Logistic Regression	0.9571	0.9588	0.9571	0.9573
	ANN	0.9849	0.9853	0.9849	0.9849

15-fold cross-validation setting, both models maintained excellent performance, with at most two misclassifications across the entire dataset. These confusion matrices clearly demonstrate the robustness of ensemble learning methods for CKD prediction. Random Forest and Gradient Boosting not only achieved high accuracy, precision, recall, and F1-scores, but also provided well-balanced predictions for both CKD and non-CKD patients. This indicates that the models are not biased toward the majority CKD class and are equally effective in identifying minority non-CKD cases, which is critically important in medical diagnosis.

Overall, the confusion matrix results confirm that, after feature selection, ensemble models deliver near-perfect classification performance with minimal misclassification errors.

4.1 Model Interpret Ability Usings Hap

In addition to performance evaluation, SHAP (SHapley Additive exPlanations) was employed to interpret the predictions of the best-performing models. Figure 4 presents the SHAP summary plot, which highlights the most influential features contributing to CKD prediction. Key attributes

such as serum creatinine, hemoglobin, blood pressure, and albumin were identified as the strongest predictors. These findings are consistent with established medical knowledge, confirming that the model is not only highly accurate but also clinically meaningful. Figure ?? illustrates the SHAP decision plot, which demonstrates how individual features contribute to the cumulative prediction score for a specific patient sample. The plot provides a step-by-step explanation of how the model aggregates feature contributions to reach the final classification decision. This level of transparency is particularly important in medical applications, as it enables healthcare practitioners to trace and validate the reasoning behind each prediction. Such interpretability is essential for fostering trust in AI-driven clinical decision support systems. By allowing predictions to be understood and verified by medical professionals, SHAP enhances both the credibility and practical applicability of the proposed model. Together, these SHAP visualizations confirm not only the reliability of the model but also its suitability for real-world clinical deployment.

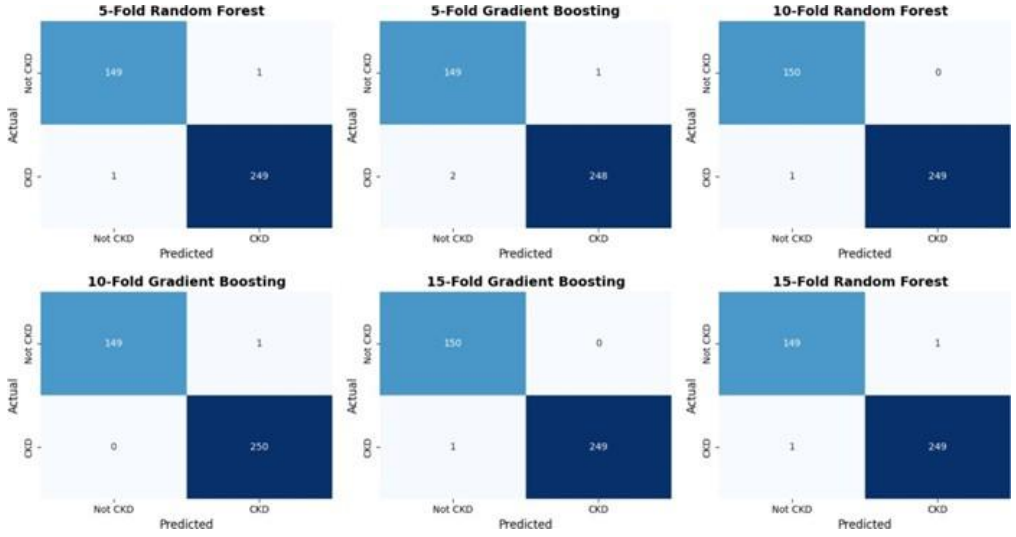

Figure.3. Confusion matrices of Random Forest and Gradient Boosting models at different cross-validations after feature selection

Also provide action able insights, making the frame work suitable for real-world clinical decision support.

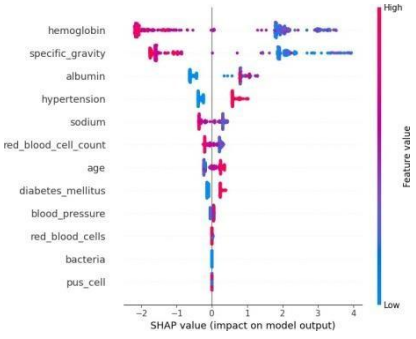

Figure. 4. SHAP summary plot showing the most influential features.

4.2 Comparative study

The comparative study presented in Table3 highlights the performance of our proposed approach against several existing state-of-the-art(SOTA)methods for CKD prediction. Previous works employed a variety of classifiers such as XG Boost, Artificial Neural Networks (ANN),Support Vector Machines (SVM),K-Nearest Neighbors(KNN),Logistic Regression(LR),and decision tree-based models like J-48.Thereportedaccu-

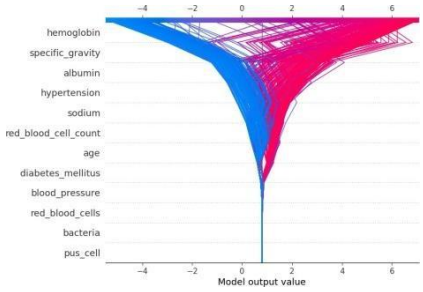

Figure.5. SHAP decision plot explaining feature contributions for a sample

racies of these studies range from 85.5% to 100%. For instance, the study by [17] us- ingJ-48with15-foldcross-validationachieved85.5 accuracy, which is significantly lower compared to ensemble earning methods. On the other hand, ANN [11]and KNN [13] achieved perfect classification (100%), showing the effectiveness of these mod- els on certain splits. Our proposed method, which leverages Random Forest (RF) and Gradient Boosting (GB) under K-Fold cross-validation (10and15),achieved an accuracy of 99.75%. This performance is comparable to the best existing methods while providing consistent and reliable results across different validation folds. The results demons trade that our approach is competitive with e existing SOTA methods, particularly ensemble-based classifiers, and offer sa more stable and generalizabl solution for CKD prediction.Byensuringhighpredictiveperformanceacrossmultiplefolds,ourmethod reduces the like LaHood of over fitting ,which is an important advantage over models that report perfect accuracy under a single split ratio.

Table3.Comparative analysis with existing state-of- the -art (SOTA) techniques

Methodology	SplitRatio	Classifiers	Accuracy
[10]	70:30	XgB	98.3%
[11]	70:30	ANN	100%
[12]	80:20	SVM	99.69%
[13]	80:20	KNN	100%
[14]	80:20	LR	98.75%
[16]	K-Fold=10	J-48	99.75%
[17]	K-Fold=15	J-48	85.5%
[18]	70:30	RF,AdaB	99%
OurMethod	K-Fold=10,15	RF,GB	**99.75%**

5. CONCLUSION

In this study, we developed and evaluated multiple machine learning models for the prediction of Chronic Kidney Disease (CKD). Feature selection using Particle Swarm Optimization (PSO) significantly improved the performance of the classifiers by reducing the number of input features while retaining essential information. Among all the models, ensemble approaches such as Random Forest and Gradient Boosting consistently achieved the highest accuracy, precision, recall, and F1-scores across different cross-validation strategies.

The results, further supported by confusion matrix analysis, demonstrate that these models can effectively differentiate between CKD and non-CKD cases with minimal misclassifications. This highlights their strong potential for practical implementation in clinical decision support systems.

For future work, this study can be extended in several directions. First, incorporating a larger and more diverse dataset would improve the generalizability of the models across different populations and clinical environments. Additionally, advanced deep learning methods and hybrid optimization techniques could be explored to further enhance predictive performance. Finally, integrating explainable AI approaches will be essential to provide clinicians with clear and transparent insights into model decisions, thereby improving trust, interpretability, and adoption in real-world healthcare applications.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to their supervisors and mentors for their continuous guidance, encouragement, and valuable feedback throughout this research work. The authors also extend their heartfelt thanks to the creators and maintainers of the CKD dataset, without which this study would not have been possible.

REFERENCES

[1] A. Francis et al., "Chronic kidney disease and the global public health agenda: An international consensus," Nature Reviews Nephrology, vol. 20, no. 7, pp. 473–485, 2024.

[2] Centers for Disease Control and Prevention (CDC), Chronic Kidney Disease in the United States, 2019. Atlanta, GA, USA: U.S. Department of Health and Human Services, 2019.

[3] K. Bello et al., "An update on the global disparities in kidney disease burden and care across world countries and regions," The Lancet Global Health, vol. 12, no. 3, pp. e382–e395, 2024.

[4] N. G. Rezk, S. Alshathri, A. Sayed, and E. E.-D. Hemdan, "Explainable AI for chronic kidney disease prediction in medical IoT: Integrating GANs and few-shot learning," Bioengineering, vol. 12, no. 4, p. 356, 2025.

[5] I. Iliyas, S. Boukari, and A. Y. Gital, "Recent trends in prediction of chronic kidney disease using different learning approaches: A systematic literature review," Journal of Medical Artificial Intelligence, vol. 8, 2025.

[6] R. K. Halder et al., "ML-CKDP: Machine learning-based chronic kidney disease prediction with smart web application," Journal of Pathology Informatics, vol. 15, p. 100371, 2024.

[7] C. P. Kovesdy, "Epidemiology of chronic kidney disease: An update 2022," Kidney International Supplements, vol. 12, no. 1, pp. 7–11, 2022.

[8] M. Jadoul, M. Aoun, and M. M. Imani, "The major global burden of chronic kidney disease," The Lancet Global Health, vol. 12, no. 3, pp. e342–e343, 2024.

[9] N. Tangri et al., "Machine learning for prediction of chronic kidney disease progression: Validation of the Klinrisk model in the CANVAS Program and CREDENCE trial," Diabetes, Obesity and Metabolism, vol. 26, no. 8, pp. 3371–3380, 2024.

[10] M. A. Islam, M. Z. H. Majumder, and M. A. Hussein, "Chronic kidney disease prediction based on machine learning algorithms," Journal of Pathology Informatics, vol. 14, p. 100189, 2023.

[11] R. Sawhney et al., "A comparative assessment of artificial intelligence models used for early

prediction and evaluation of chronic kidney disease," Decision Analytics Journal, vol. 6, p. 100169, 2023.
[12] D. M. Alsekait et al., "Toward comprehensive chronic kidney disease prediction based on ensemble deep learning models," Applied Sciences, vol. 13, no. 6, p. 3937, 2023.
[13] M. S. Arif, A. Mukheimer, and D. Asif, "Enhancing the early detection of chronic kidney disease: A robust machine learning model," Big Data and Cognitive Computing, vol. 7, no. 3, p. 144, 2023.
[14] R. C. Poonia et al., "Intelligent diagnostic prediction and classification models for detection of kidney disease," in Healthcare, vol. 10, no. 2. MDPI, 2022, p. 371.
[15] S. Pal, "Chronic kidney disease prediction using machine learning techniques," Biomedical Materials & Devices, vol. 1, no. 1, pp. 534–540, 2023.
[16] K. M. Almustafa, "Prediction of chronic kidney disease using different classification algorithms," Informatics in Medicine Unlocked, vol. 24, p. 100631, 2021.
[17] H. Ilyas et al., "Chronic kidney disease diagnosis using decision tree algorithms," BMC Nephrology, vol. 22, no. 1, p. 273, 2021.
[18] K. M. M. Uddin et al., "Machine learning-based chronic kidney cancer prediction application: A predictive analytics approach," Biomedical Materials & Devices, vol. 2, no. 2, pp. 1028–1048, 2024.
[19] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[20] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Annals of Statistics, pp. 1189–1232, 2001.
[21] V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, pp. 273–297, 1995.
[22] R. A. Fisher, "The use of multiple measurements in taxonomic problems," Annals of Eugenics, vol. 7, no. 2, pp. 179–188, 1936.
[23] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," Nature, vol. 323, no. 6088, pp. 533–536, 1986.
[24] O. P. Pathak, O. P. Singh, and U. Biswas, "Predicting chronic kidney disease using machine learning methodologies," in 2025 8th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech). IEEE, 2025, pp. 1–6.

AUTHOR BIOGRAPHY

Adarsh Maurya is a second-year undergraduate student in the Department of Botany at Banaras Hindu University, Varanasi. His academic interests include botany, the medical and healthcare applications of plant sciences, and the use of machine learning techniques for biological and medical data analysis. He is particularly interested in interdisciplinary research that integrates computational methods with botany for applications in medical research and healthcare.

Arpit Pathak is a third-year undergraduate student pursuing a Bachelor of Technology in Computer Science and Engineering at the Indian Institute of Information Technology, Bhagalpur. His research interests include biomedical signal processing, quantum machine learning, and computer vision.He has worked as a Research Intern at IIT (BHU), Varanasi, and UIET, Panjab University, where his work focused on biomedical signal processing for healthcare and prosthetic applications. He has co-authored multiple papers published in international conferences.His current interests involve applying advanced AI techniques to medical imaging and brain–computer interface systems.

Bhanu Priya received the Ph.D. degree from the Department of Electronics and Electrical Engineering (EEE) at the Indian Institute of Technology Guwahati, Guwahati, India, in 2018 She is currently working as an Assistant Professor (Temporary) in the Department of Electronics and Communication Engineering at the Indian Institute of Information Technology Bhagalpur, Bhagalpur, Bihar, India. She was previously associated as an Assistant Professor with the Bundelkhand Institute of Engineering and Technology, Jhansi, Uttar Pradesh, India. She has published several research articles in internationally reputed journals and conference proceedings. Her research interests include speech signal processing, biomedical signal processing, and brain–computer interfacing (BCI).