ADROID Conference Series: Engineering and Technology
Volume 2, Issue 3, pp.132-140, 2026

Predictive Analysis of Chronic Kidney Disease using Explainable Ensemble Learning Models

Siddharthan L, Mohammad Ishtiyaq Qureshi

Department of Computational Intelligence, SRM Institute of Science and Technology, Chennai, India

manchu.e18260@cumail.in, 22bai71383@cuchd.in, 22bai71389@cuchd.in, 22bai71358@cuchd.in
22bai71408@cuchd.in

Abstract. Chronic kidney disease (CKD) progresses and is potentially fatal. Early detection of CKD is essential to prevent severe effects and kidney failure. This work provides an intelligent and comprehensible method of predicting CKD using ML and ensemble methods. It works with the CKD dataset from the UCI ML Repository, comprising 400 patient records from Apollo Hospital in India and containing 25 numeric and nominal attributes. We apply statistical and model-based imputation to replace missing values, and SMOTE to address class imbalance (CKD vs. non-CKD). We identified six clinically significant items, including haemoglobin, serum creatinine, albumin, high blood pressure, age, and diabetes mellitus. A combination of preprocessing, feature engineering, and feature selection techniques, such as Mutual Information, Variance Thresholding, RFE, and Sequential Feature Selection, was used to identify these. Various models are evaluated using Stratified K-Fold cross-validation. These are LR, Gaussian NB, SVM, DT, RF and AdaBoost. A mixed Voting Classifier composed of RF and AdaBoost achieves an F1-score of 99.3% and 99.4% accuracy, precision, and recall. The model is easier to comprehend because individual forecasts and feature contributions are explained using LIME and SHAP. The web application, developed using Flask, allows users to register and log in, enter data securely, make real-time predictions, and receive interpretable results. It provides outcomes such as “CKD detected” and “no CKD detected” that help doctors make informed decisions.

Keywords: CKD, Machine Learning Ensemble models, Explainable AI, Interpretability.

1. INTRODUCTION

Clinically, CKD is an indication that the kidneys are impaired or their ability has been reduced irreversibly. It is typically evidenced by a Glomerular Filtration Rate (GFR) that is less than 60 mL/min/1.73 m 2 over a period of over three months [1]. When it becomes severe, it tends to destroy the kidneys which cannot be replaced resulting in AKI or ESRD, which require costly treatment methods such as dialysis or transplants [2]. The disease has many other conditions associated with it such as Diabetes Mellitus and High Blood Pressure that predispose the possibility of the development of CKD significantly. The clinical diagnosis of CKD is detected by the standard lab tests that include blood and urine tests. The most significant signs include albuminuria and elevated blood creatinine level as they indicate that the kidneys are not filtering as effectively as they are supposed to [3]. In addition, it has albumin in the urine that is usually an initial indication of renal destruction among diabetics. The reason why this is so is that broken glomerular screens allow the leakage of proteins into the urine [4].

Since the kidneys are not able to regulate the amount of erythropoietin and vitamin D secretions, it may result in anemia, bone frailty, edema, hypertension, and increased risk of heart diseases [5]. Majority of these individuals are unaware of its existence in them. More than 360 individuals commence dialysis per hour due to the failure of the kidneys to perform their functions. Almost 75% of new cases of ESRD are brought about by diabetes and high blood pressure [3]. The financial cost is equally frightening as well: ESRD treatment costs over 32 billion a year in America, and the poor in other countries such as India are unable to receive kidney replacement treatment due to health insurance or lack thereof [6].

Being an engineer, CKD analysis is not easy to conduct since there are not a lot of medical datasets available and varied. Even in the situation when available, healthcare data is frequently imperfect in terms of cognitive and sampling biases [1]. This complicates the use of ML apps that are trustworthy. The Apollo Hospital data of the UCI ML Repository is the most complete publicly available source on CKD prediction, although it lacks some clinical variables and a more diverse demographic sample [7]. Recent advancements in XAI have provided us with a clearer interpretation of the way to analyze models, and this has allowed us to discover significant biomarkers that influence the process of CKD development. XAI is better when used with ensemble learning models as it enhances prediction accuracy and transparency of clinical data. This allows physicians to make evidence-based decisions regarding the way to identify and treat CKD at an early stage [8].

2. RELATED WORK

The article by Bayram et al. [9]developed a DL-based system to locate and predict kidney diseases. They have included XAI to make the framework easier to follow and gain confidence in the professional community. The study emphasized that though the DL models were more precise compared to conventional methods, they were not easy to utilize in the clinical setting because of their black box character. The model also indicated the most important clinical biomarkers to predict CKD using XAI techniques such as LIME and SHAP. The authors concluded that the use of XAI increased the accuracy of prediction and provided the users with a greater sense of confidence in the system. This is why it is a promising alternative to use in healthcare applications when the ability to explain things is highly valued.

Silveira et al.[10] researchers investigated the possibility of early detection of CKD by ML techniques that are best applied when data is small and uneven. They explored several models, including RF, SVM, and Gradient Boosting algorithm to identify early signs of CKD. The issue of unevenly distributed data was resolved by authors by applying SMOTE which increased the sensitivity of the classifier to minor classes. RF performed the most accurate tests and best memory rates indicating that it would be suitable in the clinical prediction tasks.

Mehrabi et al. [11] conducted a massive study that investigated ML bias and fairness. It demonstrated significant ethical and technical issues of healthcare data analysis. Cognitive, sampling, and labelling biases can be detrimental to predictive models, particularly in medicine such as predicting CKD where there is not much demographic and socioeconomic diversity. The authors proposed learning methods and debiasing algorithms that recognize fairness to ensure that the model performs equally across all groups of populations. This ensures that healthcare prediction models remain transparent, ethical and practical in the real world.

Ventrella et al. [12] measured the progress made by CKD by using supervised ML techniques. This was more of an attempt to determine how the diseases would worsen rather than the initial diagnosis of the disease. The data on continuous patients and the techniques of LR, DT, and ensemble learning were utilized by the researchers to classify CKD stages into groups. We discovered that some of the most significant clinical manifestations in monitoring the worsening of the disease are serum creatinine and blood urea nitrogen, as well as haemoglobin using the feature importance analysis. The paper demonstrated the role that time-series data and longitudinal models play in the understanding of the changes CKD experiences over time and will result in improved predictions of patient outcomes over time.

Nishat et al. in the paper [13]considered the problem of implementing ML techniques to locate CKD by examining such models as KNN, NB, DT, and RF. Based on a correlation based feature selection, the authors identified haemoglobin, albumin, blood pressure, serum creatinine as the most significant factors in the prediction of CKD. According to their findings, RF was the most suitable clinical decision support since it was easy to comprehend and fix. The study also indicated that, ensemble techniques came up with a number of models, which could further be combined to provide predictions which are much accurate.

Chittora et al. provided a comprehensive ML insight into CKD prediction in their paper [14]. They developed an optimized pipeline that involved selection of features, preprocessing and tuning the model. They identified the optimal parameter selection of predicting CKD using algorithms such as SVM, LR, DT, and AdaBoost. The study revealed that, as compared to individual classifiers, the ensemble learning techniques were far more effective with medical characteristics that lack linear relationships, among themselves. According to the research, AdaBoost was the most accurate and recalls the most, which made it among the best classifiers of CKD finding.

Dave et al. [15] examined the possibility of utilizing XAI in healthcare by using a heart disease dataset, which provides us with the information that can be used to enhance CKD prediction models. The SHAP and LIME interpretability methods were used to explain the decisions taken by ML models. The findings indicated that both explainable AI and ensemble methods are more effective with regard to trust and diagnostic accuracy. This allows the same to be used in schemes to predict CKD.

Ghosh [16] made the process of predicting CKD more accurate through ML techniques and emphasis on fine-tuning parameters, and model evaluation. The paper relied on the UCI CKD dataset and examined the functionality of various algorithms such as SVM, RF, and KNN. According to the feature importance metrics, haemoglobin, serum creatinine and albumin were discovered as the most significant predictors. The author made the conclusion that RF was superior to other algorithms, in fact both in accuracy and in ease of use.

As stated in Aljaaf et al. [17], they developed a predictive analytics and ML-based system to detect CKD at an early stage. The technique involved preprocessing, normalization, and ensemble data models in the effort to enhance the precision of diagnostics. Some of the predictors that were compared in this study were DT, LR and NB. They were evaluated based on several factors, including their accuracy, sensitivity and specificity. Their model was a great success, demonstrating that predictive analytics may be applied to assist doctors in making real-life decisions. The authors further noted the combination of these models with healthcare information systems could be useful in early intervention and proactive tracking of patients.

Webster et al. [18] provided a complex clinical account of CKD including its general causes, mechanism, and impacts on health in the global context. The research mentioned the key risk factors, including diabetes, hypertension, and family backgrounds. It also emphasized the importance of early detection of such conditions with the help of such biomarkers as GFR, albuminuria, and serum creatinine levels. Their findings provided the clinical foundation on the use of AI and data-driven approaches to diagnose and treat CKD.

3. MATERIALS AND METHODS

The proposed CKD prediction model is a ML framework designed to offer a high-quality, convenient, and scalable clinical decision support. The UCI CKD data is the base that contains 400 patients of Apollo Hospital in India with 25 clinical and biological features. During the data preparation stage, missing values are imputed both through statistical and model-based imputation, categorical variables are encoded, numerical features are normalized, and the SMOTE corrects the imbalance between classes [19]. Mutual Information, Variance Thresholding, RFE and Sequential Feature Selection are used to select six key attributes that include hemoglobin, serum creatinine, albumin, high blood pressure, age and diabetes mellitus. Some of the learned supervised classifiers include LR, Gaussian NB, SVM, DT, RF and AdaBoost. The data is simplified using LIME and SHAP. Flask structure is applied to create and start a hybrid ensemble Voting Classifier that utilizes both RF and AdaBoost to apply to web-based applications of clinical use that are accessible.

There are steps to be followed to determine CKD as illustrated in Figure 1. The initial one is Data Pre-processing and Feature Selection RFE FS. The second one is Data Balancing SMOTE. These ML models are LR, NB, SVM, AdaBoost, DT, and RF. We then train, test and rate a Voting Classifier with the help of metrics like Accuracy and F1-Score. The most promising model is stored, implemented in a Flask web application to be utilized, and its forecasts are explained with the help of XAI (LIME and SHAP).

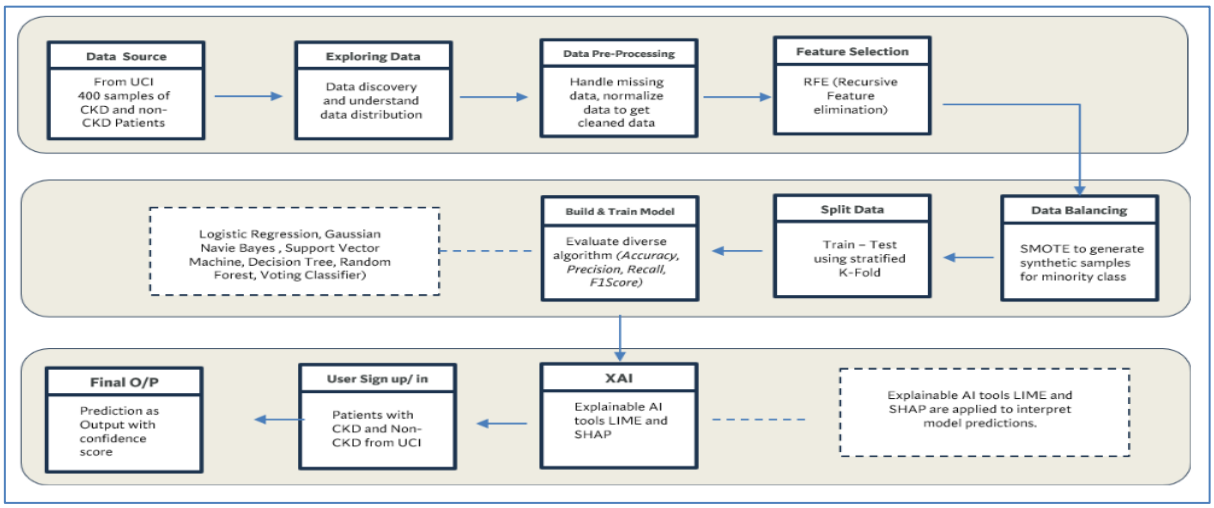

Fig. 1. System Architecture

A. Dataset Collection

The dataset used in the study was a reliable source of medical data and consisted of 25 characteristics in ARFF format of 400 records. It includes such things as age, high blood pressure, diabetes and biochemicals such as albumin, serum creatinine, hemoglobin and hypertension.[20]. The traits provide a complete picture of the health of a patient that can be utilized to locate CKD. The data was imported to a pandas Data Frame, and this provided it with a structured format to further analyze, develop a model towards its validation with an aim of predicting CKD with accuracy and early enough.

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wbcc	rbcc	htn	dm	cad	appet	pe	ane	class
0	48	80	1.020	1	0	<NA>	normal	notpresent	notpresent	121	...	44	7800	5.2	yes	yes	no	good	no	no	ckd
1	7	50	1.020	4	0	<NA>	normal	notpresent	notpresent	<NA>	...	38	6000	<NA>	no	no	no	good	no	no	ckd
2	62	80	1.010	2	3	normal	normal	notpresent	notpresent	423	...	31	7500	<NA>	no	yes	no	poor	no	yes	ckd
3	48	70	1.005	4	0	normal	abnormal	present	notpresent	117	...	32	6700	3.9	yes	no	no	poor	yes	yes	ckd
4	51	80	1.010	2	0	normal	normal	notpresent	notpresent	106	...	35	7300	4.6	no	no	no	good	no	no	ckd

5 rows × 25 columns

Fig. 2. CKD Dataset

B. Pre-Processing

Pre-processing is used to deal with the missing values, reduce noise, identify significant features, and equalize the distribution of classes to ensure that the data is correct and reliable. These measures make model learning quicker, and predictions to be more precise, and the system more stable in order that CKD classification could be reliable.

Data Preprocessing: Not all values are present as not all clinical tests were done or reported differently. Statistical imputation of numerical features and mode based, or model driven estimation of categorical variables can make this possible. This approach maintains the integrity of data and prevents the loss of information that may adversely affect the accuracy of predictions. Normalization and encoding techniques ensure that features are represented in a consistent manner as well as that different methods can be compatible with one another. Handling missing data appropriately reduces bias, increases stability, and allows the learning algorithms to discover meaningful clinical trends to help CKD be predicted accurately.

Feature Selection: The essential element of feature selection is to find the most significant clinical features and make models more helpful and comprehensible. RFE works by removing features that are less significant individually depending on the performance of a model and the ranking of their importance. RFE ensures that the traits that are most crucial contribute to classification through multiple trainings of the model and dropping of the weakest features at every stage [21]. It reduces the number of dimensions, prevents overfitting as much as possible and enhances generalization. The selected characteristics are very much related to medicine and thus prediction and explanation are more valid. Computations are also easier to make with RFE, and faster training and more accurate estimates. This is quite significant to real-time clinical decision support systems.

Data Balancing: A lack of an equal number of cases in the CKD data can cause the learning methods to favor the majority of the cases and this complicates the possibility of locating cases in the minority. To overcome this issue, SMOTE is implemented in the sampling of the data. Rather than generating copy records, SMOTE generates false samples of the minority category by interpolating between the actual data points. This makes the spreading out of classes even as the fundamental data patterns remain unchanged. The balancing of input makes the classifier more sensitive, remembers better and more accurate in general, particularly in finding CKD. SMOTE enhances model learning, reduces bias, and facilitates the correct and consistent disease prediction outcomes by ensuring that the CKD and non-CKD cases are proportionally represented.

C. Training and Testing:

The processed data set was divided into training and testing sets with a 10-fold cross-validation technique which was stacked. Several ML methods were trained on the selected features. The models were evaluated by metrics such as accuracy, precision, memory and F1-score. Ensure that the system would be functional, capable of accurate prediction and reliable to ensure that it could detect CKD at the early stage through systematic training and testing was ensured.

D. Algorithms:

Logistic Regression: Employs the logistic function to simulate the correlation of important clinical features and forecast the probability of CKD. The model provides quantifiable probabilities, allowing physicians to determine the likelihood of CKD occurrence and take early preventive measures through data-guided diagnostic information.

$$\hat{y}_i = \sigma(w^T x + b) = \frac{1}{1 + e^{-(w^T x + b)}} \qquad (1)$$

Gaussian Naïve Bayes: Fits probabilistic models using Gaussian assumptions to classify cases as having or not having CKD [22]. It performs rapid and accurate predictions with limited quantities of information, enabling physicians to estimate the probability of CKD and support early diagnosis.

Support Vector Machine (SVM): Assigns cases to CKD and non-CKD categories by determining the optimal hyperplane that maximizes class separation. Non-linear relationships are modeled using kernel mapping to achieve high accuracy and reliability in clinical data analysis and early disease identification.

$$\min \frac{1}{2} \| W \|^2 + C \sum_{i=1}^{n} \xi_i \qquad (2)$$

Decision Tree (DT): Constructs a classification tree for CKD prediction based on clinical features [23]. The model provides well-defined decision paths, effectively handles missing values, and is easily interpretable by clinicians, thereby improving diagnostic transparency and early disease detection.

$$I(i) = 1 - \sum_{i=1}^{k} p_i^2 \qquad (3)$$

Random Forest (RF): Combines multiple Decision Trees to obtain a more stable classification process while reducing overfitting. It identifies important CKD indicators, improves prediction accuracy, and provides clinicians with feature importance information to support patient assessment and treatment planning.

$$\text{Gini} = 1 - \sum_{i=1}^{C} (P_i)^2 \qquad (4)$$

AdaBoost: Sequentially trains weak learners by focusing on previously misclassified CKD cases to improve prediction accuracy. This approach enhances model performance in identifying complex disease patterns and supports accurate, data-driven clinical assessment.

$$H(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t \, h_t(x) \right) \qquad (5)$$

Voting Classifier: Combines the predictions of multiple algorithms, including Random Forest and AdaBoost, using majority voting to produce more accurate and less biased CKD detection results. This ensemble approach ensures consistent predictions for supporting clinical decision-making.

$$\hat{y} = \arg\max_{c} \left(\sum_{i=1}^{n} I\left(\hat{y}_i = c\right) \right) \qquad (6)$$

E. Integration of XAI and Flask Framework

By integrating Explainable Artificial Intelligence (XAI) with the Flask framework, machine learning models can generate predictions that are transparent, interpretable, and accessible through an interactive web interface. The implementation utilizes **LIME** and **SHAP** to explain CKD prediction outcomes.

LIME provides local explanations by analyzing the contribution of individual features for a single patient prediction. Important clinical variables such as age, blood pressure, and laboratory test results are highlighted to explain their influence on the predicted outcome.

SHAP complements this approach by providing global explanations across multiple patient records. It visualizes feature importance using summary plots and other visualization techniques, enabling clinicians to understand the overall factors influencing model predictions.

By integrating these XAI techniques into a Flask application, clinicians can directly input patient information, obtain CKD prediction probabilities, and interpret the contribution of each clinical feature to the prediction. This integration enhances model transparency, supports clinical decision-making, and increases physician trust in AI-assisted diagnosis.

4. EXPERIMENTAL RESULTS

Accuracy: The ability of a model to distinguish between CKD and non-CKD patients is measured using accuracy. It is defined as the proportion of correctly classified instances.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \qquad (7)$$

Precision: Precision is the ratio of correctly predicted positive cases to all predicted positive cases.

$$\text{Precision} = \frac{TP}{TP + FP} \qquad (8)$$

Recall: Recall measures the ability of the model to identify all actual positive CKD cases.

$$\text{Recall} = \frac{TP}{TP + FN} \qquad (9)$$

F1-Score: The F1-score combines Precision and Recall into a single performance metric by computing their harmonic mean.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100 \qquad (10)$$

TABLE I
PERFORMANCE EVALUATION

Model	Accuracy (%)	Precision (Weighted %)	Recall (Weighted %)	F1 Score (Weighted %)
Voting	99.4%	99.43%	99.4%	99.39%
AdaBoost	99.2%	99.23%	99.2%	99.19%
RF	98.2%	98.28%	98.2%	98.19%
DT	96.8%	97.01%	96.8%	96.79%
Log Reg	96.8%	97.12%	96.8%	96.78%
SVM	96.2%	96.59%	96.2%	96.18%
GNB	95.2%	95.50%	95.2%	95.18%

Table 1. indicates that the mixed Voting Classifier is better than the other models in terms of CKD prediction models with accuracy, weighted precision, recall, and F1-score.

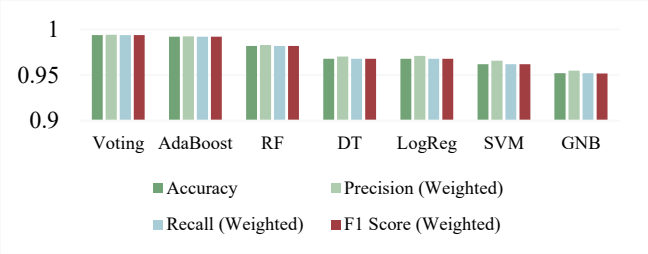

Fig. 3. Comparison Graph

The Figure 3 is the comparison graph of CKD prediction models based on accuracy, weighted precision, recall, and F1-score. The metrics are represented by the colors of the colors: accuracy is blue, precision is green, recall is orange and F1-score is red. This indicates that the hybrid Voting Classifier had higher performance.

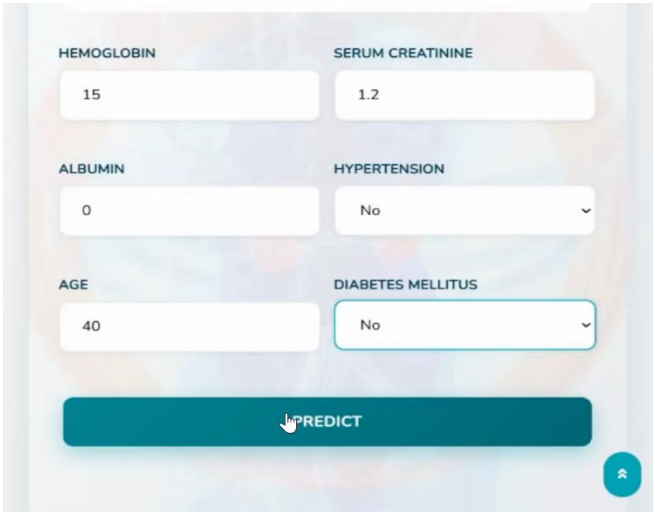

Fig. 4. Upload the Data

Fig. 4 The input form provides individuals with an opportunity to add data concerning patients with CKD, and this will be used to make predictions within the ML-based classification system.

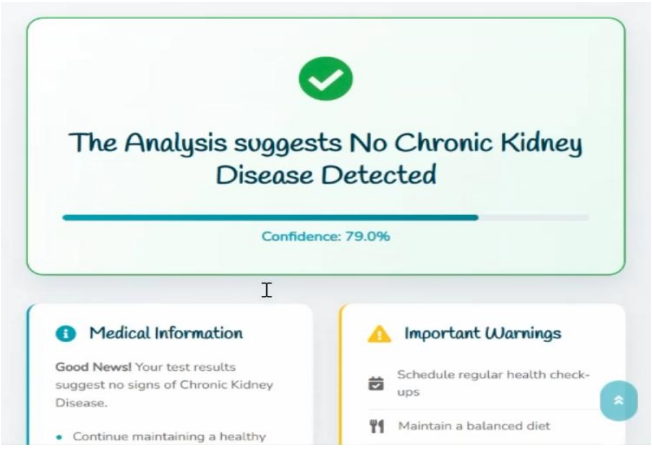

Fig. 5. Predicted Results

Fig 5 The forecast outcome reveals that It is The Analysis says No CKD Detected with a confidence score of 79.0 and this implies the model is certain.

Fig. 6. Upload the Data

Fig.6 input interface allows individuals to provide CKD-specific clinical information concerning a patient which allows the system to predict the results with confidence and interpretability scores.

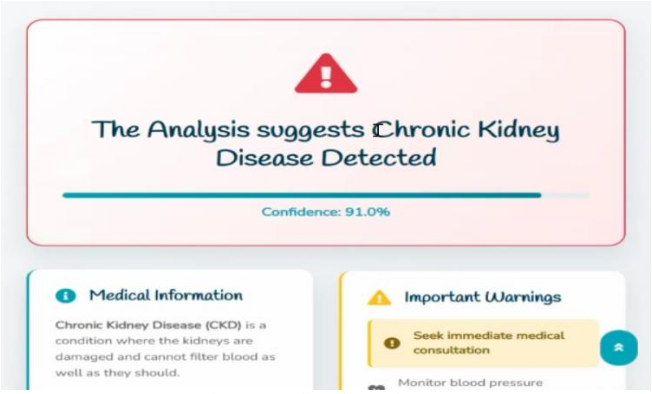

Fig. 7. Predicted Results

The forecasts with the highest confidence score of 91.0 is shown in Fig.7 and is labeled as The Analysis suggests CKD Detected.

5. CONCLUSION

This piece of work demonstrates how it is possible to create a highly accurate and interpretable CKD prediction system by integrating cutting-edge ML practices with a convenient web-based implementation. It was then possible to discover clinically meaningful characteristics that significantly influence the diagnosis of CKD due to strong feature selection, good preprocessing, and balanced learning using SMOTE. The Voting Classifier, which was a combination of RF and AdaBoost, achieved the best performance among all the tested models. Its

accuracy, precision, recall and F1-score were 99.4, 99.4, 99.4 and 99.3 respectively indicating that it was very reliable and consistent in distinguishing between CKD and non-CKD cases. The results were more comprehensible when using XAI techniques, mostly LIME and SHAP, displaying the contribution of various features. This simplified the predictions and earned confidence in the predictions. The entire system is implemented with the help of a Flask-based web application, which allows users to safely register and log in and allows them to input structured patient information, perform real-time predictions, and visualize the expected outcomes. The system provides articulate diagnostic responses such as CKD detected, or no CKD detected that assists the doctors to make intelligent decisions and initiate the treatment at the earliest. On the whole, the findings indicate that the mixed ensemble Voting model, together with interpretable insights and an interactive Flask interface is a powerful, precise, and convenient means of assessing the CKD risk in the real-life context of healthcare.

Wearable devices and electronic health records also can be linked with real-time management systems to provide predictions on demand and customized notifications. To top this, the predictions based on the sophisticated DL models can be made more precise, and telemedicine systems can be linked in such a way that patients can receive remote assistance. Improving the methods of explainability will also assist clinicians to have more trust in the system and it can be easy to make it convenient so both healthcare providers and patients can use it to pursue proactive kidney health management on a mobile-friendly interface.

Conflict of Interest

The authors declare no conflicts of interest regarding the current research.

References

[1] P. A. Moreno-Sanchez, "Data-Driven Early Diagnosis of Chronic Kidney Disease: Development and Evaluation of an Explainable AI Model," *IEEE Access*, vol. 11, pp. 38359–38369, 2023, doi: 10.1109/ACCESS.2023.3264270.

[2] K. M. T. Jawad, A. Verma, F. Amsaad, and L. Ashraf, "AI-Driven Predictive Analytics Approach for Early Prognosis of Chronic Kidney Disease Using Ensemble Learning and Explainable AI," Jan. 2025. [Online]. Available: http://arxiv.org/abs/2406.06728

[3] S. K. Ghosh and A. H. Khandoker, "Investigation on explainable machine learning models to predict chronic kidney diseases," *Scientific Reports*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-54375-4.

[4] K. Jhumka, M. M. Auzine, M. Heenaye-Mamode Khan, S. M. Casseem, S. A. Fedally, and Z. Mungloo-Dilmohamud, "Explainable Chronic Kidney Disease (CKD) Prediction using Deep Learning and Shapley Additive Explanations (SHAP)," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Oct. 2023, pp. 29–33, doi: 10.1145/3633598.3633604.

[5] S. K. Ghosh, N. Widatalla, and A. H. Khandoker, "Machine Learning Framework for Early Detection of Chronic Kidney Disease Stages Using Optimized Estimated Glomerular Filtration Rate," *IEEE Access*, vol. 13, pp. 78057–78072, 2025, doi: 10.1109/ACCESS.2025.3565549.

[6] M. V. N., A. N., and K. S. K., "Decision Tree-Based Explainable AI for Diagnosis of Chronic Kidney Disease," in *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2023, pp. 947–952, doi: 10.1109/ICIRCA57980.2023.10220774.

[7] F. Sanmarchi, C. Fanconi, D. Golinelli, D. Gori, T. Hernandez-Boussard, and A. Capodici, "Predict, diagnose, and treat chronic kidney disease with machine learning: a systematic literature review," May 01, 2023, Springer Science and Business Media Deutschland GmbH, doi: 10.1007/s40620-023-01573-4.

[8] A. Holzinger, C. Biemann, C. S. Pattichis, and D. B. Kell, "What do we need to build explainable AI systems for the medical domain?," Dec. 2017. [Online]. Available: http://arxiv.org/abs/1712.09923

[9] A. BAYRAM, C. Gurkan, A. BUDAK, and H. KARATAŞ, "A Detection and Prediction Model Based on Deep Learning Assisted by Explainable Artificial Intelligence for Kidney Diseases," *European Journal of Science and Technology*, Mar. 2022, doi: 10.31590/ejosat.1171777.

[10] A. C. M. da Silveira, Á. Sobrinho, L. D. da Silva, E. de B. Costa, M. E. Pinheiro, and A. Perkusich, "Exploring Early Prediction of Chronic Kidney Disease Using Machine Learning Algorithms for Small and Imbalanced Datasets," *Applied Sciences (Switzerland)*, vol. 12, no. 7, Apr. 2022, doi: 10.3390/app12073673.

[11] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A Survey on Bias and Fairness in Machine Learning," Jan. 2022. [Online]. Available: http://arxiv.org/abs/1908.09635

[12] P. Ventrella, G. Delgrossi, G. Ferrario, M. Righetti, and M. Masseroli, "Supervised machine learning for the assessment of Chronic Kidney Disease advancement," *Computer Methods and Programs in Biomedicine*, vol. 209, p. 106329, 2021, doi: 10.1016/j.cmpb.2021.106329.

[13] M. Nishat *et al.*, "A Comprehensive Analysis on Detecting Chronic Kidney Disease by Employing Machine Learning Algorithms," *EAI Endorsed Transactions on Pervasive Health and Technology*, pp. 1–12, Mar. 2021, doi: 10.4108/eai.13-8-2021.170671.

[14] P. Chittora *et al.*, "Prediction of Chronic Kidney Disease - A Machine Learning Perspective," *IEEE Access*, vol. 9, pp. 17312–17334, 2021, doi: 10.1109/ACCESS.2021.3053763.
[15] D. Dave, H. Naik, S. Singhal, and P. Patel, "Explainable AI meets Healthcare: A Study on Heart Disease Dataset," Nov. 2020. [Online]. Available: http://arxiv.org/abs/2011.03195
[16] P. Ghosh, F. M. Javed Mehedi Shamrat, S. Shultana, S. Afrin, A. A. Anjum, and A. A. Khan, "Optimization of Prediction Method of Chronic Kidney Disease Using Machine Learning Algorithm," in *2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)*, 2020, pp. 1–6, doi: 10.1109/iSAI-NLP51646.2020.9376787.
[17] A. J. Aljaaf *et al.*, "Early Prediction of Chronic Kidney Disease Using Machine Learning Supported by Predictive Analytics," in *2018 IEEE Congress on Evolutionary Computation (CEC)*, 2018, pp. 1–9, doi: 10.1109/CEC.2018.8477876.
[18] A. C. Webster, E. V. Nagler, R. L. Morton, and P. Masson, "Chronic Kidney Disease," *The Lancet*, vol. 389, no. 10075, pp. 1238–1252, Mar. 2017, doi: 10.1016/S0140-6736(16)32064-5.
[19] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease Prediction by Machine Learning Over Big Data From Healthcare Communities," *IEEE Access*, vol. 5, pp. 8869–8879, 2017, doi: 10.1109/ACCESS.2017.2694446.
[20] Elslvler, I. D. Penman, S. H. Ralston, M. W. J. Strachan, and R. P. Hobson, *"24th Edition."*
[21] T. J. Hoerger *et al.*, "The future burden of CKD in the United States: A simulation model for the CDC CKD initiative," *American Journal of Kidney Diseases*, vol. 65, no. 3, pp. 403–411, Mar. 2015, doi: 10.1053/j.ajkd.2014.09.023.
[22] H.-Y. Kim, "Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test," *Restorative Dentistry & Endodontics*, vol. 42, no. 2, p. 152, 2017, doi: 10.5395/rde.2017.42.2.152.
[23] E. Seker, J. R. Talburt, and M. L. Greer, "Preprocessing to Address Bias in Healthcare Data," in *Studies in Health Technology and Informatics*, IOS Press BV, May 2022, pp. 327–331, doi: 10.3233/SHTI220468.