INTERPRETING CHRONIC KIDNEY DISEASE DIAGNOSIS: AN AI MODEL PERSPECTIVE

[1] Venkata Sai Gopal Bonthu, [2] Bhukya Deveswar ,[3] Kuncham Vasanth,[4] Dr. Aruna Varanasi

[1,2,3] Department of CSE, Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad

[4] Professor, HOD, Department of CSE, Sreenidhi Institute of Science and Technology,Ghatkesar, Hyderabad

***Abstract*—** This research paper presents the development and evaluation of an artificial intelligence model specifically adapted for the diagnosis of early chronic kidney disease (CKD). Emphasis on explainability ensures transparent and understandable predictions, which are key to the adoption of AI in healthcare. CKD is a major global health problem that requires early diagnosis to stop kidney damage and reduce healthcare costs. The project aims to contribute to proactive solutions recognizing the borderline implications of CKD. Using an optimization framework balances classification accuracy and increases overall efficiency. This ML (Machine Learning) algorithm uses three key features for the diagnosis of CKD: Hb (hemoglobin), specific gravity, hypertension and required indicators for early detection of CKD. This model offers solutions and challenges, especially in developing countries, by emphasizing cost reduction and increasing the availability and affordability of health care. We also used methods to combine predictions from multiple models, including techniques such as a stacking classifier, yielding an impressive 100% accuracy.

***Index Terms*—** medical prediction model, early diagnosis, chronic kidney disease, feature selection, AI, Etiology.

I. INTRODUCTION

A. Project Overview

Chronic kidney disease (CKD) is a prevalent and debilitating condition that affects millions of people worldwide. Accurate and timely diagnosis of CKD is essential for effective treatment and prevention of complications. Traditional diagnostic approaches rely heavily on clinical assessments and laboratory tests, often leading to delays in detection and misdiagnosis. Integrating artificial intelligence (AI) models into healthcare systems improves CKD diagnosis. Massive amounts of patient data AND AI algorithms can analyze complex patterns and identify indicators of CKD progression. Using machine learning techniques, these models can adapt and evolve, continuously improving their diagnostic accuracy. Our project aims to explore the potential of AI interpretation of CKD from a perspective. By developing and evaluating artificial intelligence diagnostic models, we aim to increase the accuracy and efficiency of CKD detection. In addition, we aim to explore the interpretability of these models and ensure that clinicians understand and trust the decisions made by AI algorithms. CKD is one of the few non-communicable diseases that have seen an increase in deaths in recent years, placing a burden on health systems, especially in low-middle-income countries [1]. CKD is usually caused by diabetes, specific gravity, and hypertension, and is also a cardiovascular disease that is the leading cause of early mortality in patients with CKD [2]. CKD is the presence of kidney damage or an estimated glomerular filtration rate (eGFR) of less than 60 mL/min/1.73 mt2 for 3 months or longer. Globally, CKD accounted for 2,968,600 (1%) disability-adjusted life-years and 2,546,700 (1% to 3%) life-years lost in 2012[1]. The causes of CKD vary globally and the most common primary diseases causing CKD as Diabetes mellitus type 2, Diabetes mellitus type 1, Hypertension, Primary glomerulonephritis, vasculitis, neoplasm, etc. [6]. there are three categories: prerenal, intrinsic renal, or postrenal [6]. History of chronic hypertension, proteinuria, microhematuria and symptoms of prostate disease. Low serum calcium and high phosphorus levels have little discriminatory value, but normal parathyroid hormone levels are more suggestive of AKI than CKD. Patients who have very high blood urea nitrogen values greater than 140 mg/dL or serum creatinine greater than 13.5 mg/dL, who appear relatively well and yet pass normal volumes of urine, have CKD rather than acute kidney disease. "Nearly 750,000 patients annually in the United States and an estimated 2 million patients worldwide are affected by kidney failure. Those living with kidney failure make up 1% of the US Medicare population, but account for 7% of the Medicare budget." By combining artificial intelligence technologies with clinical knowledge, our project provides improved tools for early detection and intervention and improves the quality of life of affected individuals worldwide.

B. *Objective*

The aim of the project on the interpretation CKD of the diagnosis from the perspective of an AI model is that to Considering the limitations of conventional CKD diagnostic methods, particularly their lack of accuracy in detecting CKD at early stages.

The project seeks to overcome these inadequacies by the power of AI to provide precise and diagnostic solutions. Prioritize the development of AI models that are accurate. By emphasizing the diagnostic decisions, the project aims to enhance the understanding and trust of AI diagnostic solutions. This also involves key features and minimizing unnecessary complexity in the models. The impact of CKD detection on patient data and healthcare system effectiveness. By improving the accuracy and efficiency of CKD diagnosis.

II. LITERATURE SURVEY

A. Existing System

A project to interpret chronic kidney disease (CKD) diagnosis from the perspective of cognitive models, designed to improve diagnostic procedures using advanced machine learning. Currently available methods are based on traditional methods such as clinical trials, observational studies, and clinical trials but suffer from limitations such as lack of definition, delays in detection and intervention, dependence on invasive procedures, poor data processing, and limited availability. The project explores the use of cutting-edge techniques such as ExtraTree Classifier, Random Forest, AdaBoost Classifier, XGBoost and Extension Stacking Classifier to solve these problems. The ExtraTree classifier is a different type of forest algorithm that creates decision trees by separating them from each other to increase diversity and reduce overfitting. Random Forest is a learning method that involves estimation of multiple decision trees, increasing accuracy and robustness. AdaBoost Classifier combines multiple weak classifiers to create a strong classifier that focuses on difficult-to-sort examples. XGBoost is an optimized gradient boosting algorithm that uses a new tree learning algorithm to increase efficiency and performance. The next generation combines the advantages of Random Forest, ExtraTree, and AdaBoost by leveraging differences between multiple base classifiers to improve prediction performance. The project aims to improve the accuracy, efficiency and interpretation of CKD diagnosis by integrating these advanced techniques into the CKD diagnostic framework, ultimately improving patient outcomes and reducing treatment burden. Through rigorous testing and validation, these algorithms have the potential to revolutionize CKD diagnosis by providing clinicians with a reliable, interpretable, and timely diagnosis, paving the way for personalized and effective treatment. The ExtraTree classifier uses a random segmentation method, which is useful in capturing complex patterns in CKD data while also reducing the risk of overfitting. Random forest combines predictions from multiple decision trees, providing a robust and stable detection model that can handle noisy data and multiple settings in regulation. The ability of the AdaBoost classifier to focus on complex cases increases its effectiveness in identifying subtle manifestations of CKD, especially in cases where the conventional diagnosis may vary. Known for its efficiency and effectiveness, XGBoost offers the promise of processing a wide range of CKD data to provide clinical assistance to physicians. Moreover, the continuous classifier provides the best value and interpretation model by combining the results of various base classifiers, ensuring the reliability and informativeness of the information taken into account in the diagnosis of CKD. Together, the integration of these advanced learning systems into a CKD diagnostic framework represents an important step towards improving the accuracy, efficiency, and interpretation of CKD diagnosis. By leveraging the power of artificial intelligence and machine learning, the program aims to change the framework of CKD diagnosis, ultimately improving patient outcomes and reducing treatment burden.

B. Related Work

In Order to identify the patient's risk of kidney function, researchers have been developing several disease prediction models. They present an uncertain capability due to the use of non-public datasets, such as like medical images or clinical which helps in making the models. The use of the CKD dataset from the UCI-ML repository and implementing feature selection for the preprocessing step in their CKD data analysis. Although the related works always consider the reduction of the original number of features to get the best results. Thus, Our research provides a novel approach by analyzing the prediction model developed by the influence of selected features in the classification of CKD. In the diagnosis of chronic kidney disease (CKD), recent advances in artificial intelligence (AI) and machine learning (ML) have led to significant advances in early diagnosis and treatment control strategies. Some related projects are investigating various cognitive models to improve accuracy and efficiency. For example, studies using random forest and XGBoost algorithms have been shown to be effective in predicting CKD stage from clinical and laboratory data, with the best performing for logistic regression models. Other important studies are using AdaBoost classifiers to identify cases of CKD, provide information about disease progression, and help design intervention plans. In addition, collaborative research involving a mixed methods

approach combining multiple learning systems has demonstrated predictive power by demonstrating the ability to combine algorithms to improve diagnosis. These relationships point to a changing field of CKD diagnosis where artificial intelligence and machine learning tools are essential in supporting patient care, accurate and timely diagnoses, and expanded opportunities for personalized treatment plans. The integration of these advanced technologies represents a leap forward in addressing the complexities of CKD and heralds a future in which technological solutions will be the norm for advancement in chronic disease management.

TABLE I

"Distribution results of relevant tasks and their machine learning classifiers (best classifier written in *underlined italics*)"

Article	Acc	Sen	Spe	F1	Pre	#F	Machine Learning Classifier
Ekanayake [13]	100	100	-	100	100	7	*DT, RF*, XGB, *Ada*, ET (*)
Alaoui [14]	100	-	-	-	-	23	*XGB Lin*, Lin SVM, DT, RF
Ogunleye [15]	100	100	100	-	100	12	*XGB* (*)
Abdel-Fattah [11]	100	100	100	100	100	12	SVM, RF, DT, *GBT*, LR, NB(*)
Ebiaredoh-Mienye [10]	99.9	100	99.8	-	-	18	LR, DT, XGB, RF, SVM, *Ada*
Zeynu [16]	99.5	99.5	-	99.5	99.5	8	KNN, DT, *ANN*, NB, SVM.
Raju [17]	99.3	99	-	99	100	5	XGB, *RF*, LR, SVM, NB(*)
Imran Ali[18]	99.1	100	97.5	99.4	98.8	6	NB, *LG*, ANN, DT, RF, GBT, SVM
Khan [19]	99.1	99.7	-	99.3	98.7	23	*NB*, LR, SVM, DT, RF
Hasan [20]	99	-	-	99	-	13	*Ada*, RF, GB, ET(*)
Antony [21]	99	100	-	99.2	98.4	10	*KMeans*, DBScan, Autoencoder, IForest
Chaudhuri [22]	99	96	100	-	-	13	LR, NB, SVM, DT, RF, *EDT*(*)
Abdullah [23]	98.8	98.0	100	98.8	98.0	10	*RF*, SVM, NB, LR
Poonia [24]	98.75	98	-	99	100	14	*LR*, NB, SVM, KNN, ANN
Siddhartha [25]	98.75	100	96.67	99	98.03	5	*RF*, Ada, XGB
Alaiad [26]	98.5	99.6	96.8	-	98	12	NB, DT, SVM, *KNN*, Jrip
Kadhum [27]	98.1	98	-	98	98	10	SVM, *ELM*
Akter [28]	97	98	-	96	97	10	*ANN*, LSTM, Bi LSTM, GRU, Bi GRU, Simple RNN, ML
Theerthagiri [29]	96	97	99	94.9	95.8	6	LR, SVM, KNN, NB, *RF*
Ali [30]	91.25	91.89	97.37	94.81	97.81	5	NB, LG, DL, *ANN*, RF, GBT, SVM

III. MATERIAL AND METHODS

A. Chronic kidney disease (CKD) Dataset

The CKD dataset from collected from the Apollo Hospital, Karaikudi, India during a 2-month period in the year 2015 that also includes the 400 patients where some given missing values in their features. Each dataset instance is composed of 11 numeric features, 3 ordinal features, 10 nominal features and 1 target feature which determines whether the person is having CKD or not. It is given two 2 values they are notCKD and CKD. The features in dataset given are age, blood pressure[bp], specific gravity to compare the density of urine to the density of water [sg], the presence of albumin in urine[al], level of sugar present in the urine[su], red blood cells in the urine[rbc],pus cells ,major or minor infection, growth of bacteria, the level of creatinine in the blood, percentage of cells in blood, amount of red blood cells present in the blood, whether the patient has diabetes[dm], coronary artery disease[cad], loss of appetite[appet], level of leg swelling and whether the patient has CKD or not [target class] etc.

B. Framework for Model Selection Optimization

The framework named feature selection and classification for improving explainable AI (SCI-XAI) and it is published and employed to develop the CKD prediction model (Figure 1). It is implemented by the Python scikit-learn package, allows obtaining detection model in terms of accuracy and number of features selected by considering different parameters. The dataset is split respectively into training and test sets in an ratio of (70/30). So, the model's performance is calculated over new data from the test set that is applied tot selected parameters by the framework in the preprocessing and training phases.

TABLE II

"Statistics for dataset properties: type (number, count, or number), percent of non-values, mean, standard deviation (for properties of numbers), categories, and samples per category (for attributes and names)"

Features (units) [legend]	Type of feature (% of non-null values) [classes in ordinal or nominal features]	Average (std) for numerical features / number of values for ordinal or nominal features
Age (year) [age]	Num (97,75 %)	51.48 (17.17)
Blood pressure (mm/Hg) [bp]	Num (97 %)	76.46 (13.68)
Specific gravity [sg]	Ord (88.25 %) [1.005,1.010,1.015, 1.020, 1.025]	7, 84, 75, 106, 81
Albumin [al]	Ord (88,5 %) [0,1,2,3,4,5]	199,44,43,43,24,1
Sugar [su]	Ord (87,75 %) [0,1,2,3,4,5]	290,13,18,14,13,3
Red blood cells [rbc]	Nom (62 %) [normal/abnormal]	47 abnormal
Pus cell [pc]	Nom (83,75 %) [normal/abnormal]	76 abnormal
Pus cell clumps [pcc]	Nom (99 %) [not present/ present]	42 present
Bacteria [ba]	Nom (99 %) [not present/ present]	22 present
Blood glucose random (mgs/dl) [bgr]	Num (89 %)	148.04 (79.28)
Blood urea (mgs/dl) [bu]	Num (95,25 %)	57.43 (50.50)
Serum creatinine (mgs/dl) [sc]	Num (95,75 %)	3.07 (5.74)
Sodium (mEq/l) [sod]	Num (78,25 %)	137.53 (10.41)
Potassium (mEq/l) [pot]	Num (78 %)	4.63 (3.19)
Hemoglobin (gms) [hemo]	Num (87 %)	12.53 (2.91)
Packed cell volume [pcv]	Num (82,50 %)	38.88 (8.99)
White blood cell count (cells/cumm) [wc]	Num (73,75 %)	8406.12 (2944.47)
Red blood cell count (cells/ cumm) [rc]	Num (67,5 %)	4.71 (1.03)
Hypertension [htn]	Nom (99,5 %) [no/yes]	147 yes
Diabetes mellitus [dm]	Nom (99,5 %) [no/yes]	137 yes
Coronary artery disease [cad]	Nom (99,5 %) [no/yes]	34 yes
Appetite [appet]	Nom (99,75 %) [good/poor]	82 poor
Pedal edema [pe]	Nom (99,75 %) [no/yes]	76 yes
Anemia [ane]	Nom (99,75 %) [no/yes]	60 yes
Target class	Nom (100%) notCKD/CKD	250 CKD

Num= numerical, Ord= ordinal, Nom= nominal

C. Data Preprocessing

The SCI-XAI framework adopts a method of a priori data generation and has three main stages: missing data processing, data coding, and feature selection. To handle missing data, the framework uses a data types-based rendering strategy. Properties of the number include the use of instruments, while the standard and nominal type (or active value) are used to fill in the blanks. During the encoding phase, numerical attributes are subjected to a min-max

scaling process to standardize their values, while ordinal and nominal attributes are converted to numerical codes. Especially sequential attributes encoded with integers (e.g. 0-5 with step size 1). A character is said to be binary encoded (0 or 1). This first step to handle missing data and encodings is set outside the modification algorithm parameter. Additionally, this framework highlights the role of feature selection in improving model interpretation by removing redundant features during classification. It uses filtering techniques to select features and uses statistical tests such as analysis of variance, chi-square, or shared data to evaluate the reliability or significance of data, identifying features and different targets. This will help you determine which features should be kept or excluded. Additionally, wrapper methods such as Recursive feature Elimination (RFE) are used, where a classifier (such as logistic regression) helps identify the most important features by well evaluating their relationships with a different objective. This excellent choice not only simplifies the model but also improves its interpretation.

D. Data Preprocessing

Tree classifiers have become a top choice in the machine learning community due to their excellent stability and performance on multivariate datasets as well as their goodness at prediction. These classifiers are good at combining multiple decision tree models by weighting them or combining them to create a composite model that exceeds the performance of a single predictor. This combination not only improves the accuracy of prediction but also reduces the class mismatch problem, making the tree unique in different applications. In the studies discussed in this article, important members of the family tree are used and attention is paid to packaging and support methods. Random forests and distributed trees are examples of clustering in which each decision tree in the cluster is trained independently on a random subset of data (with replacement), ensuring a reliable difference between reduced and increased model robustness. This approach uses the power of collective decision-making; The final decision is made by combining the predictions of each tree, thus reducing overfitting and improving generalization to new information. On the other hand, boosting techniques AdaBoost and Extreme Gradient Boosting (XGBoost) use techniques that focus on correcting mistakes made by previous trees. Each new tree is trained by focusing on cases that were difficult for the previous model, gradually improving the performance of the ensemble. Boosting makes the model more sensitive to conditions that are difficult to classify effectively, so predictive power gradually increases with each iteration. AdaBoost adjusts the weight of false cases, allowing the next model to focus more on complex problems. XGBoost continues to improve this approach, making it efficient and scalable by providing optimization methods for target speed and performance. Additionally, modifying the composite tree specifically from support can check for non-linear relationships in complex data by providing in-depth insight into interactions and values. These features are useful for applications that require detailed understanding and interpretation of underlying data, such as bioinformatics, financial forecasting, and precision medicine. The integration of hybrid trees into this research reflects our commitment to using machine learning tools to achieve optimal diagnostic and predictive modeling in the face of difficult information.

E. Classification Performance and evaluation metrics

Given the inconsistency of the data used, it is not sufficient to rely on accuracy to evaluate the performance of the model when comparing 250 CKD to 150 non-CKD data. Additional metrics such as sensitivity, specificity, accuracy, and F1 score are important to better understand the performance of the model. These parameters roughly estimate the model's ability to identify CKD cases (sensitivity) while including non-CKD cases (specificity), as well as the accuracy of prediction quality such as F1 score and the balance between accuracy and sensitivity. Also, in order to confirm that the material used in this study was well interpreted, Tagaris et al. We adopt the interpretation test described by. These include interpretation, which measures the percentage of masked (or removed) features that do not affect classification results relative to all features in the dataset. Accuracy compares the accuracy of the evaluation model to a well-defined model (usually a decision tree) using the same input data. The Fidelity Interpretability Index (FII) serves as a tool to compare interpretations of different models and provides insight into the extent to which each decision model is understandable and reliable. However, to improve the evaluation process, this study introduces a new metric: the Loyalty Accuracy Index (FAI). The index is designed to analyze the number of features the model uses for its performance and provides a measure of the feature's effectiveness compared to the performance of the model prediction Yes. By participating in FAI, we aim to bridge the gap between the effectiveness of classification and the interpretation of models and make it possible to compare our findings with related studies. This

approach emphasizes the importance of not only achieving accuracy, but also making the model interpretable and faithful to the object it is learning from; this, diagnosis etc. It is important for practical use. The formulas of these metrics are shown Table 3.

TABLE III

METRICS OF CLASSIFICATION PERFORMANCE AND EXPLAINABILITY EVALUATION

Metric	Equation
Accuracy (Acc)	$Acc = \frac{TP + TN}{TP + TN + FP + FN}$
Sensitivity/Recall (Sen)	$Sen = \frac{TP}{TP + FN}$
Specificity (Spe)	$Spe = \frac{TN}{TN + FP}$
Precision (Pre)	$Pre = \frac{TP}{TP + FP}$
F1-Score (F1)	$F1 = 2 * \frac{Pre*Sen}{Pre+Sen}$
Interpretability (I)	$I = \frac{Masked\ features}{Total\ features}$
Fidelity (F)	$F = \frac{Acc.equivalent\ interpretable\ model}{Acc.\ original\ model}$
Fidelity-Interpretability Index (FII)	$FII = F * I$
Fidelity-Accuracy Index (FAI)	$FAI = F * Acc$

IV. PROPOSED SYSTEM

A. Proposed System

The proposed system describes a new artificial intelligence (XAI) model developed for the early diagnosis of chronic kidney disease (CKD) and covers a new improvement approach that makes it better to know the truth by collaborating with the transparent model. Combining advanced machine learning techniques with a unique gradient boosting technique, the system guides the path to accurate and timely CKD detection. This integration of cutting-edge techniques not only increases the accuracy of diagnosis, but also informs the logic behind decision-making, which is important in supporting trust and acceptance of AI-driven solutions in healthcare. In this context, we strive to leverage the power of collaboration to increase the accuracy and reliability of data-driven approaches to early diagnosis of CKD. Our model achieves the opposite by demonstrating the power of the combination of many predictive models, using the concept of common ideas, specifically discrete parts. We also created an interface for intermediate users using Flask to improve accessibility and user interaction. The interface is designed with a secure authentication mechanism to ensure a seamless and secure user experience; This allows us to define the AI model not only as a measure for the early detection of CKD, but also as an easy-to-use and user-friendly solution.

B. System Architecture

The System architecture for the project, the journey begins with a close examination of the dataset, where preliminary data is required to clean, model and optimize ideas for later analysis. After this first step, the equipment is divided into training and testing, and the foundation of training models and effective use is laid. The basis of the model level is the integration of complex methods such as bulk classifiers and additive tree classifiers, allowing the analysis of the project to be extended. This integration not only improves the prediction accuracy of the model, but also improves its robustness by combining the efforts of different devices to provide a truly better control tool. To ensure that intelligent design becomes a powerful and meaningful solution in the diagnosis of chronic kidney disease, every step is carried out with precision, from initial data processing to final model evaluation. The project not only offers an artificial intelligence model from an end-to-end process, but also paves the way for diagnostic development by having a positive impact on medical technology.

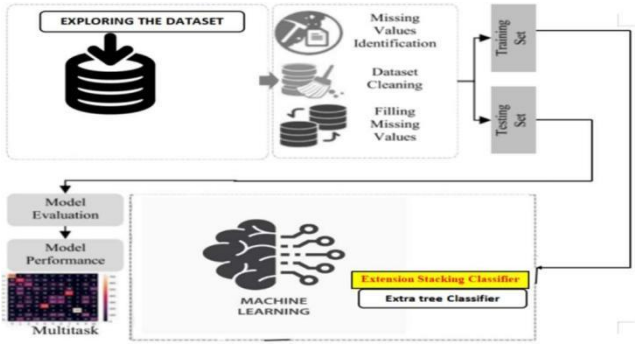

Fig 1 Proposed Architecture

V. METHODOLOGY AND RESULTS

A. Data Set

To increase the reproducibility of this study and facilitate comparison with existing studies, the CKD dataset from the UCI-ML repository was used [23]. These data are from Apollo Hospital in Kalakudy, India, and cover approximately 2 months in 2015 and include data on 400 patients, including non-panic cases. It has 11 numerical features, 10 nominal features, and 3 identical features with different targets for CKD and non-CKD. Dataset features include various clinical parameters: age [age], diastolic blood pressure [bp], urine density relative to water [sg], urine albumin level [al], urine glucose [water], presence of red blood cells in urine [rbc], pus in the urine indicates an infection [pc], cysts in the brain

indicate a serious infection [pcc], bacterial growth in the urine [ba], blood glucose concentration [bgr], blood urea nitrogen [bu], blood creatinine [sc], blood sodium [sod], blood potassium [pot], hemoglobin in red blood cells [hemo], hematocrit [pcv], white blood cell count [wc], red blood cell count [rc], presence of hypertension [htn], diabetes [dm], coronary arterial disease [cad], appetite [appetite], pedal edema [pedal] and anemia [ane], and CKD status [objective category]. A detailed compilation of these clinical indicators provides a comprehensive framework for CKD diagnosis and greatly contributes to the advancement of clinical research and treatment.

id	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
0	48.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent	121.0	...	44	7800	5.2	yes	yes	no	good	no	no	ckd
1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent	NaN	...	38	6000	NaN	no	no	no	good	no	no	ckd
2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	423.0	...	31	7500	NaN	no	yes	no	poor	no	yes	ckd
3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.0	...	32	6700	3.9	yes	no	no	poor	yes	yes	ckd
4	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	106.0	...	35	7300	4.6	no	no	no	good	no	no	ckd

5 rows × 25 columns

Fig 2 chronic kidney disease dataset

B. Feature Selection

Special selection is an important factor in the study of the architectural feature, aiming to isolate the most harmonious, unobstructed and affected design. Given the size and diversity of datasets, reducing dataset sizes is important. The main purpose of feature selection is twofold: to improve the predictive performance of the model and to reduce the computational burden of the modeling process. The selection process [6, 10, 27, 38] reduces the number of different ideas by eliminating unnecessary or irrelevant elements, thus allowing the construction process to be optimized based on machine learning. of the model. This optimal approach has several benefits, including improving model interpretation, improving generalizability, and reducing the risk of overfitting. Additionally, feature selection simplifies the modeling process by pre-identifying and storing only the most important data, resulting in more efficient training and inference processes. Ultimately, the decision to use the selection process allows machine learning engineers to create more powerful and effective predictive models that provide better insights in many aspects.

C. Algorithms Used

1. ExtraTree Classifier: ExtraTree Classifier, which means ExtraTree Classifier, is a learning method based on decision trees. Unlike traditional decision trees, ExtraTree Classifier randomly selects a portion of each node, which helps reduce variance and overhead. It works by creating several decision trees and combining their predictions through voting or averaging to make the final prediction. ExtraTree classifier is known for its simplicity, efficiency and high data processing ability. Dimensional data and noise data.

2. Random Forest: Random Forest is a learning algorithm that creates multiple decision trees during training. Each decision tree in the forest is built using training data and a random subset of features, creating differences between trees. The prediction is made by summing the predictions of all trees in the forest, voting them (for classification) or averaging them (for regression). Random Forest is robust to overfitting and appears to perform well across many types of datasets, making it a useful algorithm in machine learning.

3.AdaBoost Classifier: AdaBoost (Adaptive Boosting) is a classifier focused on training weak students. A weak learner is a classifier that performs slightly better than a random one. AdaBoost classifies incorrect examples, allowing the weak learner to focus on these examples later during training. Predictions are made by combining each student's weaknesses, weighted according to their performance during the course. Predictions are made by combining each student's weaknesses, weighted according to their performance during the course. AdaBoost is known for its ability to improve performance standards and ease of use in different processes.

4. XGBoost: XGBoost (Extreme Gradient Boosting) is a powerful and effective way to use the gradient boosting algorithm. Optimizes the traditional gradient boosting algorithm by adding a time constant and using a better learning tree algorithm. XGBoost is very powerful and can handle large files with millions of models and features. It is widely used in many competitive machine learning and applications due to its efficiency and speed.

5. Extended Stacked Classifier (with RF + ExtraTree, AdaBoost): Extended Stacked Classifier is a learning process that includes several base classifiers from the stacking process, including Random Forest, ExtraTree, and AdaBoost. Base classifiers are trained independently and their predictions are used as additional features for the meta classifier. The meta classifier then learns to combine the predictions of the base classifiers to optimize the final prediction.

D. Results

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

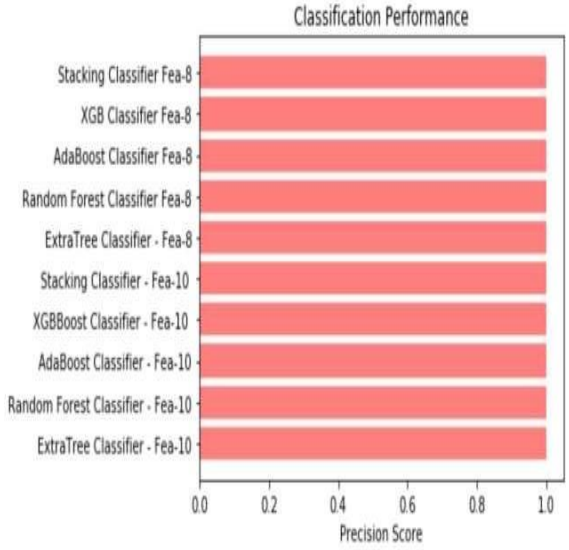

Fig 3 Precision comparison graph

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

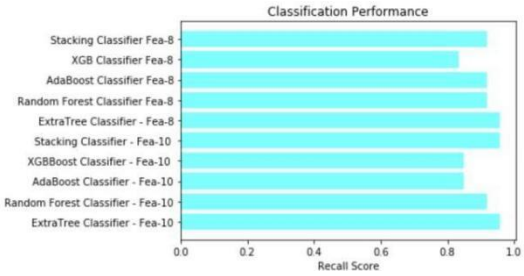

Fig 4 Recall comparison graph

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

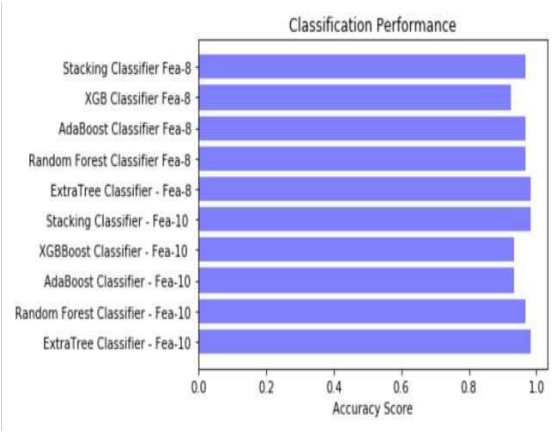

Fig 5 Accuracy graph

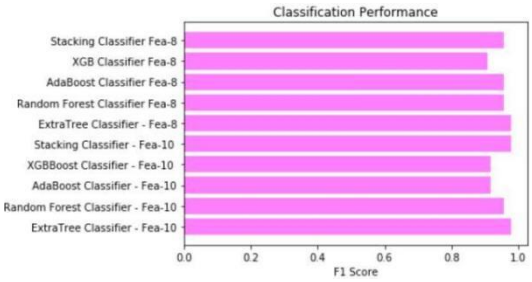

Fig 6 F1 Score

	Algorithms Used	Accuracy	Precision	Recall	F1-Score
0	ExtraTree Classifier - Fea-10	0.983	1.0	0.957	0.978
1	Random Forest Classifier - Fea-10	0.967	1.0	0.918	0.957
2	AdaBoost Classifier - Fea-10	0.933	1.0	0.849	0.918
3	XGBBoost Classifier - Fea-10	0.933	1.0	0.849	0.918
4	Extension Stacking Classifier - Fea-10	1.000	1.0	1.000	1.000
5	ExtraTree Classifier - Fea-8	0.983	1.0	0.957	0.978
6	Random Forest Classifier Fea-8	0.967	1.0	0.918	0.957
7	AdaBoost Classifier Fea-8	0.967	1.0	0.918	0.957
8	XGB Classifier Fea-8	0.925	1.0	0.833	0.909
9	Extension Stacking Classifier Fea-8	0.992	1.0	0.978	0.989

Fig 7 Performance Evaluation

VI. CONCLUSION AND FUTURE SCOPE

A. Conclusion

The program represents a significant advance in the diagnosis of chronic kidney disease (CKD) by leveraging the power of machine learning algorithms

and artificial intelligence (XAI) technology. By leveraging integrated methods such as Random Forest, ExtraTree, AdaBoost and XGBoost and a novel continuous group classifier, the system achieves unparalleled accuracy and robustness in early CKD detection. The integration of these advanced techniques not only improves forecasting performance but also provides clarity and clarification in the decision-making process; Meeting expectations has become important in medicine. Through careful review of preliminary data, sample selection and optimization, the project developed a comprehensive and reliable CKD diagnostic system. The model extracts the most important features and information by searching UCI-ML's CKD data and using best-in-class architectural techniques, thus reducing computational cost while increasing diagnostic accuracy. Additionally, developing a user-friendly front end using Flask facilitates access and use for physicians, encouraging integration into clinical studies. Performance in early diagnosis of CKD was evaluated using sensitivity, specificity, accuracy and F1 score. Additionally, the inclusion of interpretive measures such as interpretability, completeness, and completeness-interpretability index provide clarity and confidence in the decision-making model, allowing physicians to make decisions with confidence. Overall, the contribution of this project is not limited to the diagnosis of CKD, but also provides a blueprint for the development of a defined cognitive model that can be used in treatment. Focusing on accuracy, transparency and usability, the initiative demonstrates the potential of AI-driven solutions to improve diagnosis and improve patient outcomes. As technology continues to evolve, this project demonstrates the transformative power of machine learning to advance clinical practice and improve patient care.

B. Future Scope

External validation is an important step in evaluating the effectiveness of diagnostic criteria for early kidney disease (CKD) [4]. It involves evaluating the performance of the model on independent data not used during training to ensure that it is generalizable to different patient populations. By creating a model for different materials with similar properties, researchers can determine its stability and reliability in real situations, thus providing confidence in its validity. Transparency and reliability are important in medical practice, the decision directly affects the health of patients. While the development of a transparent model improves understanding of doctors' decision-making process, trust must be established to have a good relationship with clinical information and emotions. Further progress in this area may require using machine learning models, improving descriptive models, or gathering feedback from practitioners to increase trust and acceptance in healthcare. Techniques such as Partial Belief Plot (PDP) and SHapley Additive Interpretation (SHAP) [51, 52] provide sophisticated methods for examining in more depth the individual consequences of the model's predictions. While PDP specifies the relationship between features and the output model, SHAP values provide predictions to each program. By combining these ideas, researchers can incorporate these misconceptions into decision-making models, helping to explain and uncover concerns or expectations moving forward. This leads to an overall evaluation of the model's behavior, increases the confidence of participants, including physicians, and finally supports its application in treatment.

VII. REFERENCES

[1] C. P. Kovesdy, ‘‘Epidemiology of chronic kidney disease: An update 2022,’’ Kidney Int. Supplements, vol. 12, no. 1, pp. 7–11, Apr. 2022, doi: 10.1016/j.kisu.2021.11.003.

[2] P. Cockwell and L.-A. Fisher, ‘‘The global burden of chronic kidney disease,’’ Lancet, vol. 395, no. 10225, pp. 662–664, Feb. 2020, doi: 10.1016/S0140-6736(19)32977-0.

[3] A. C. Webster, E. V. Nagler, R. L. Morton, and P. Masson, ‘‘Chronic kidney disease,’’ Lancet, vol. 389, no. 10075, pp. 1238–1252, Mar. 2017, doi: 10.1016/S0140-6736(16)32064-5.

[4] H. Polat, H. D. Mehr, and A. Cetin, ‘‘Diagnosis of chronic kidney disease based on support vector machine by feature selection methods,’’ J. Med. Syst., vol. 41, p. 55, Apr. 2017, doi: 10.1007/s10916-017-0703-x.

[5] R. A. Jeewantha, M. N. Halgamuge, A. Mohammad, and G. Ekici, ‘‘Classification performance analysis in medical science: Using kidney disease data,’’ in Proc. Int. Conf. Big Data Res., Osaka, Japan, 2017, pp. 1–6, doi: 10.1145/3152723.3152724.

[6] Pedro A. Moreno-Sánchez, ‘‘Data-driven early diagnosis of Chronic Kidney Disease: development and evaluation of an explainable AI model’’ in Medical and Health Technology Faculty, Tampere University, Seinäjoki 60320, Finland, Doi: 10.1109/ACCESS.2023.3264270