2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3) | 979-8-3503-3800-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/IC2E357697.2023.10262425

A LIME-Based Explainable Machine Learning Technique for the Risk Prediction of Chronic Kidney Disease

1st Ankit Vijayvargiya
Department of Electrical Engineering
Swami Keshvanand Institute of Technology, Management and
Gramothan, Jaipur, India
ankitvijayvargiya29@gmail.com

2nd Aarsh Raghav
Department of Electrical Engineering
Swami Keshvanand Institute of Technology, Management and
Gramothan, Jaipur, India
singhraghavsunny@gmail.com

3rd Anchal Bhardwaj
Department of Electrical Engineering
Swami Keshvanand Institute of Technology, Management and
Gramothan, Jaipur, India
bhardwajanchal02@gmail.com

4th Naveen Gehlot
Department of Electrical Engineering
Malaviya National Institute of Technology
Jaipur, India
gehlotnaveen27@gmail.com

5th Rajesh Kumar
Department of Electrical Engineering
Malaviya National Institute of Technology
Jaipur, India
rkumar.ee@gmail.com

***Abstract*—Chronic Kidney Disease (CKD) is on the rise in the current research scenario. The main function of the kidney is to remove and purify the waste and blood in the human body. Diabetes is the most prevalent cause of kidney disease. The key to preventing or curing CKD is identifying it at an early stage. If early detection is avoided, there is a greater chance of kidney failure as well as heart disease, bone disease, or an imbalance in potassium and calcium levels. Prediction at an early stage for a long and healthy life is made feasible with the assistance of a machine intelligence classifier. In this study, the prediction of CKD and Non-CKD patients is done by applying five machine learning classifiers. The results show that the Random Forest classifier has the greatest accuracy of all classifiers. Explainable artificial intelligence (XAI) is introduced for a trustworthy explanation of the result. XAI investigates how the Random Forest model gives high accuracy with input features imported into the classifier.**

***Index Terms*—Chronic Kidney Disease (CKD); Explainable Artificial Intelligence (XAI); Machine Learning (ML); Random Forest (RF).**

I. Introduction

The fundamental functions of the kidneys within a human being are waste elimination and blood purification. A kidney removes acid from the body and aids in the proper balance of salts, water, and minerals in the blood, such as calcium, sodium, and potassium. Muscles and nerves in our body cannot function effectively without this equilibrium [1]. The imbalance in the equilibrium leads to chronic kidney disease (CKD), also known as chronic renal failure, which is caused by kidney impairment, and waste accumulates in the body [2]. CKD is an irreversible decrease in renal function. Diabetes and high blood pressure are the most prevalent diseases.

CKD is considered chronic because it causes gradual damage. Anemia, fatigue, edema (swollen ankles, hands, and feet), bloody urine, and loss of appetite are some of the most common symptoms of CKD [3]. CKD affects approximately 850 million people globally, of whom 2 million have received a kidney transplant [4]. After COVID-19, 10% of patients with kidney disease and CKD patients doubled. Almost one-third of CKD patients who tested positive for COVID-19 died within 90 days [5].

In order to identify kidney illness, physicians conduct blood tests to measure the levels of waste products like creatinine that are present in the patient's blood [6]. In addition, urine tests aid medical professionals in diagnosing the abnormalities in urine samples that are the root cause of CKD. The early detection of kidney disease is made easier by artificial intelligence, which analyses the specifics of each patient's condition in order to provide precise disease forecasts. This is necessary since the number of people suffering from kidney disease is so high. Also, the number of people who have kidney problems that can be treated by professionals has increased because of Artificial Intelligence (AI) [7].

Raju et al. [8] have applied six different classifiers, Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), Neural Network (NN), XGBoost (XGB), and Naive Bayes (NB) classifiers. After comparing the results of these classifiers, the researchers found that the XGB and RF classifiers had the highest accuracy, at 99.29%. The machine learning methods are known as Gradient Boosting (GB), RF, AdaBoost (ADB), and Bagging are utilised by Nikhila and the other researchers for classification between CKD and Non-CKD subjects [9]. There are a total of seven performance measures that are computed for each method. These metrics include the F1-score, Accuracy, Sensitivity, and Specificity, as well as the Mathew Correlation Coefficient. The results of

TABLE I: Features available in dataset

Attribute	Description	Attribute	Description
ane	Anemia	**cad**	Coronary artery disease
htn	Hypertension	**pe**	Pedal edema
dm	Diabetes	**sg**	Urine specific gravity
age	Age	**class**	Target variables
appet	Appetite	**rbcc**	Red blood cell count
pc	Pus cell	**sc**	Serum creatinine
hemo	Hemoglobin	**bgr**	Blood glucose random
su	Sugar	**bp**	Blood pressure
al	Albumin	**rbc**	Red blood cells
ba	Bacteria	**pcc**	Pus cell clumps
bu	Blood urea	**wbcc**	White blood cell count
sod	Sodium	**pcv**	Packed cell volume
pot	Potassium		

the comparison suggest that ADB and RF provide the highest accuracy.

Islam et al. [10] used six machine learning methods to classify kidney disease: NB, RF, Simple Linear Regression, and Simple LR. Machine learning strategies are implemented following the collection of patient information from a medical college located in Bangladesh. RF is the algorithm with the best accuracy of 98.88% out of all the methods. Chittora and colleagues [11] used seven machine learning classifiers to predict CKD. Chi-square automated interaction detector, Artificial Neural Network, C5.0, SVM, LR, and RF are the methods employed. The dataset was obtained from the UCI repository, which consists of 400 samples with 24 attributes. The findings show that LSVM with penalty L2 has a maximum accuracy of 98.86%.

Devika et al. [12] investigate the performance of machine learning classifiers and evaluate the Accuracy, Precision, Recall, and F-measure of these classifiers in relation to the prediction of chronic kidney disease (CKD) using datasets that are freely accessible to the public from the UCI repository. Charleonnan and others [13] developed a predictive model for the prediction of CKD using machine learning techniques. For prediction, four machine learning methods are used: Decision Tree (DT), Linear Regression, K-Nearest Neighbor (KNN), and SVM. In this investigation, a 5-fold cross-validation test is performed on each method, and the results that are averaged are taken into account. The data set is composed of training data equal to 70% and test data equal to 30%. When the accuracy of machine learning algorithms is compared, the SVM has the best accuracy with 98.3% and DT, KNN, and LR have accuracy 94.8%, 98.1% and 96.55%.

Kunwar et al. [14] used a dataset of 400 patients from the UCI machine learning repository, which was reduced to 220 after data cleaning and missing value removal. Two machine learning algorithms: NB and Artificial Neural Networks (ANN) are applied to the dataset on the Rapidminer data mining tool for the prediction of CKD in the earlier stages. The accuracy of both algorithms is compared, and the result shows that NB has a higher accuracy of 100% when compared to ANN, which has an accuracy of 72.73%. Akter and others [15] applied an advanced machine learning approach to predict CKD.

In previous studies, researchers compared how well different classifiers worked, but it's hard to figure out how the classifiers predicted what people would do. Having a reliable explanation for the findings of these classifiers would be quite valuable. Explainable artificial intelligence (XAI) is presented in order to develop trustworthy predictive modelling results and to help comprehend how machine learning algorithms function [16]. Local interpretable model-agnostic explanation (LIME) is an Explainable AI approach that determines which features provide the most information and are most suited for operation [17].

In this study, a LIME based XAI approach is considered for finding the impacts of features on the machine learning classifier during the prediction of CKD. This research article is organised as follows: Section II provided a description of the methodology used. The outcomes of the work and a discussion are presented in Section III. The conclusion is provided in Part IV, along with suggestions for further research.

II. Methodology

The early identification of CKD may be accomplished using the following five classification algorithms: RF, SVM, DT, LR, and KNN. The CKD dataset is taken from a publicly available source. Data is preprocessed, which includes data cleansing and data mining. Based on their kind, data is then divided as nominal or numerical. Chronic illness information was acquired and applied to each algorithm, and the final performance evaluation of algorithms is based on accuracy. The proposed methodology is shown in Fig. 1.

A. Dataset

Using machine learning classifiers, this study attempts to determine if a patient has CKD or not. The CKD dataset that is used for analysis may be acquired from the repository for machine learning at UCI. The data for this dataset consist of 400 patients, which implies that there are 400 instances, each with 25 attributes. All of these 25 features have some connection to CKD, as indicated in Table I.

B. Data Preprocessing

The following phase, which comes after the collection of the dataset, is the preprocessing of the data. During this stage, the data that has question marks (?) or other symbols denoting missing values will have those symbols replaced with null values (NAN). Following that, the NAN values are replaced with the mean value of that attribute for numerical values, and the mode value is substituted for categorical values. This process is repeated until all the NAN values have been removed. In the dataset, there are fourteen numerical attributes and eleven category attributes. Following the application of this data encoding to the data set in order to convert string values to numerical values, it was necessary to assign the values '1' and '0' to string values in order for algorithms to be able to function properly. The machine learning model is only capable of operating on data that is numerical in

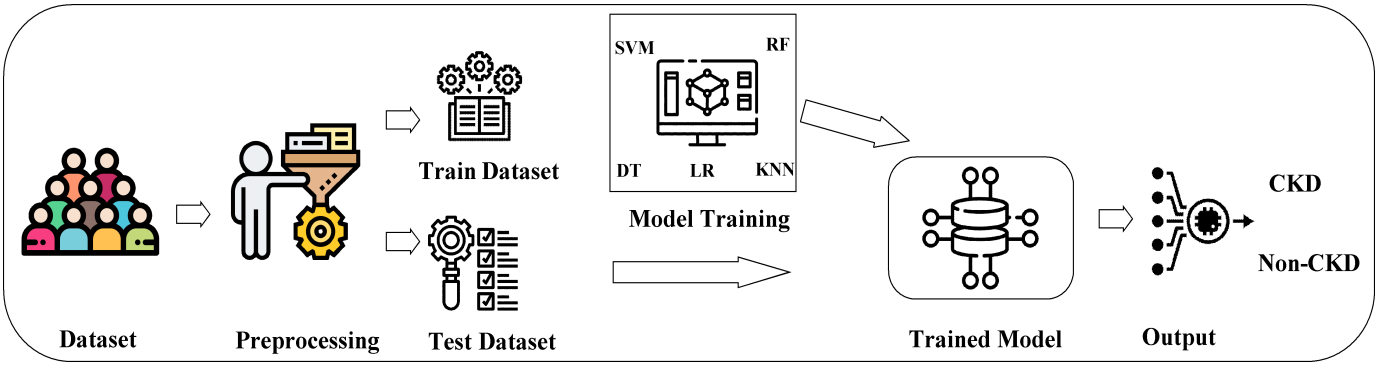


Fig. 1: The Block Diagram of Work Flow

nature. 'CKD' and 'Non-CKD' are both target labels that are included in the CKD dataset class.

C. Machine Learning Models

This section gives a quick summary of the machine learning models [18] that are used to separate people with and without CKD.

1) Decision Tree (DT): It is a kind of supervised machine learning algorithm with a tree-like structure. It consists of two types of nodes: decision nodes, which are used to make any decisions, and leaf nodes, which are the output of those decisions. This model is used for both classification and regression. The root node of the decision tree is selected on the basis of information gain. Information Gain (IG) is a measurement of how much information a feature contains. A feature having more information gain will be selected as a root node. Entropy is the randomness or impurity in a feature. The entropy describe in equation (1)

$$Entropy = -P_t log_2 P_t - P_f log_2 P_f \tag{1}$$

Where P_t is probability of a true outcome. P_f Probability of a false outcome.

The information gain describe in equation (2)

$$IG = E1 - [W * E2] \tag{2}$$

Where, 'IG' is the information gain, 'E1' is the entropy of root node, 'W' is the weighted average, 'E2' is the entropy of each feature.

2) Random Forest (RF): It is a machine learning models which is based on ensemble learning, in which multiple classifiers are combined to solve a problem. As we take multiple classifiers, the performance of the model improves. In this algorithm, a larger number of the decision tree are formed on various subset of the dataset. Finally, in this, take an average of the entire decision tree's output to improve the accuracy. The greater, number of decision trees taken, greater the accuracy. By randomly picking the attributes, N number of decision trees are formed.

3) Support Vector Machine (SVM): It is mainly used for both classification regression problems based on supervised machine learning. SVM creates a line in a manner to separate classes, data points on both sides of lines represent different categories, so there can be n numbers of lines from which it can choose. The best line is termed as a hyperplane. Datapoints closest to the hyperplane are called support vectors. The function of line is given in equation (3)

$$w^T x + b = 0 \tag{3}$$

Here, an n-dimensional weight vector is w, and a scalar quantity is b. The vector w points in a direction that is perpendicular to the hyperplane that separates the two sections. The margin may be made larger by increasing the value of the offset parameter b.

4) K-Nearest Neighbor (KNN): It is a type of machine learning method known as supervised learning, and it may be used to issues involving classification as well as regression. It has two properties: a lazy learning algorithm, Non-parametric learning algorithm. Because it does not have a specific training phase and employs all the data for training, it is referred to as a lazy learning algorithm. As KNN does not make any assumptions about the data in before, it is considered a non-parametric method.The selection of the K value is of the extreme significance since, whenever new data is entered, a comparison of K-neighbors is performed. Calculate the distance between each row of the test data and the corresponding row of the training data using the Euclidean distance algorithm. The formula for calculating Euclidean distance is shown in equation (4)

$$d(a,b) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2} \tag{4}$$

5) Logistic Regression (LR): It is one of the most popular kinds of supervised machine learning algorithms. It gives the probabilistic value of the categorical dependent output, which we want to predict by taking input as a set of independent variables. "S" shaped sigmoid logistic function is fitted to predict the values (0 or 1). It uses the concept of threshold to delineate the relationship between dependent output and independent input. When the value is greater than the threshold, it gives 1 output, and a value is less than the threshold, it gives 0. The logistic function or sigmoid function is defined in equation (5)

$$g(k) = \frac{1}{1 + e^{-value}} \tag{5}$$

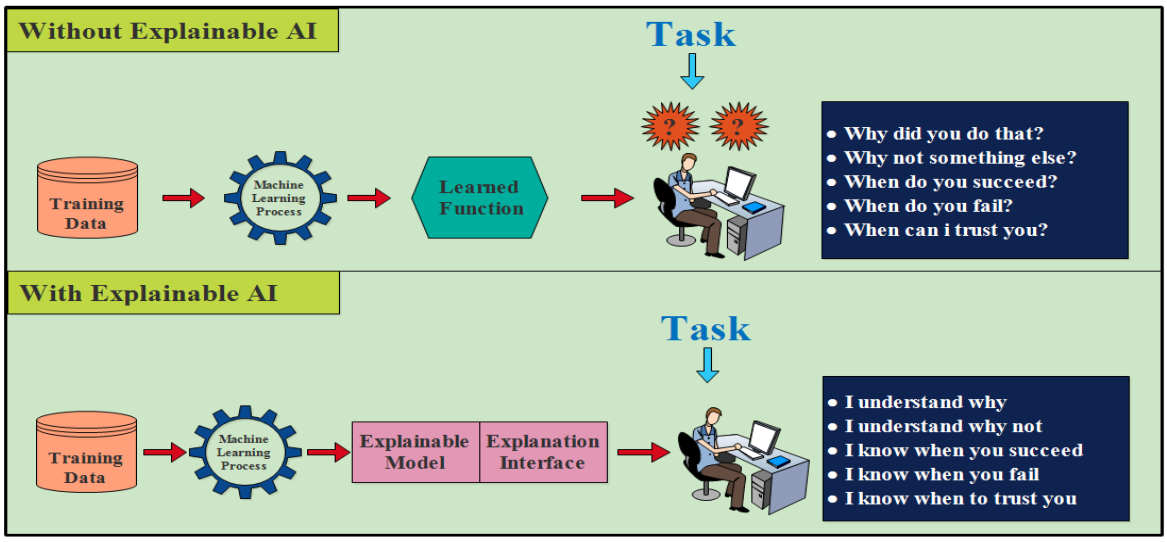


Fig. 2: Flow of Explanation Artificial Intelligence

Where 'value' is the real value that transforms.

D. Explainable AI (XAI)

Since a number of years ago, artificial intelligence (AI) had mostly theoretical study, with very few applications of practical consequence in the actual world. However, this has changed recently due to a combination of high processors, more powerful learning algorithmic models, and easier access to large amounts of data. In the past ten years, significant progress has been achieved in the field of machine learning (ML), leading to its widespread use in industry. A wide variety of real-time challenges across a wide range of industries are being handled as a result of the application of machine learning models.

However, higher model complexity is commonly utilized to achieve this enhanced predictive performance. The deep learning technique is an excellent illustration, since it enables machines to automatically recognize, understand, and extract the complex data representations necessary for detection and classification tasks.

Despite the fact that this hierarchy of increasing complexity and the use of enormous amounts of data to train and construct such advanced systems increases the system's predictive capability in the vast majority of cases, it significantly reduces their capacity to explain their workings and methods. As a result, it is harder to grasp the thinking behind their behaviours, making it more challenging to interpret their suggestions in the future.

It is difficult to trust in systems whose results are difficult to explain, particularly in domains such as healthcare or the development of autonomous vehicles, which unavoidably raise concerns of ethics and justice. The concept of XAI has seen a resurgence in recent years as a result of the creation of models for real-world applications that are efficient, fair, robust, and high-performing (XAI). It is presently gaining popularity as a result of its programmes that are simple, easy, and concise to explain.

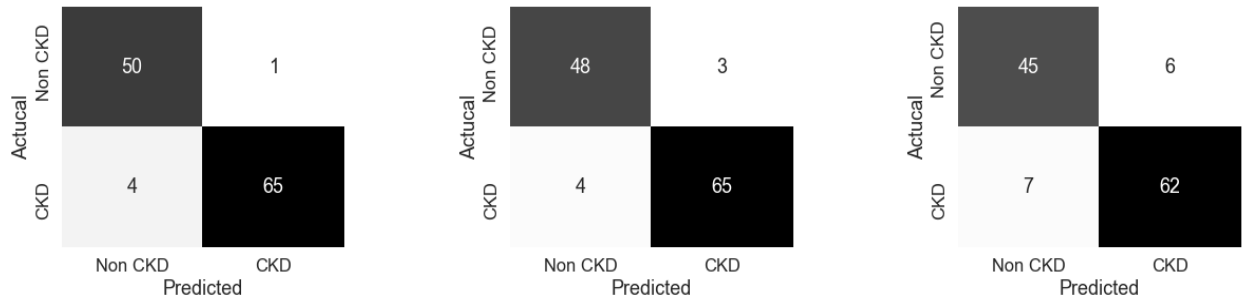
Explainable AI is a set of tools and frameworks for understanding and explaining machine learning model predictions, as well as analyzing and optimizing model performance and helping others understand the explainability of your model, as shown in Fig. 2. Here are three important factors for interpreting the model:

- Transparency
- The ease of understanding
- The ability to question

There are several ways to improve our models' interpretability, some of which are already known and used. Exploratory data analysis, visualizations, and model assessment metrics are examples of traditional methodologies. We can gain an understanding of the model's approach with their help. However, they have some limitations. In machine learning problems, normally build a model on a stationary database to acquire our objective function or loss function, which is then utilised to resolve the issue when the model meets particular criteria based on its performance. In practice, however, owing to unpredictability in collected data, extra limitations, and disturbance, a model's performance frequently drops and plateaus over time after deployment. This might include things like changes in the environment, changes in features, as well as added constraints. In order to ensure that model predictions are accurate on new data points, it is not sufficient to simply re-train a model on the same feature set. Instead, we must constantly assess how important particular features are in determining model predictions and how well models may perform on new data points. In order to overcome these constraints, several model interpretation approaches have been designed such as LIME, SHAP, ELI5, and SKATER etc.

Lime (Local Interpretable model-agnostic explanation) is an explainable AI technique that consists of interpretability and model-agnostic.It explains the behavior of an algorithm that used for explanation.

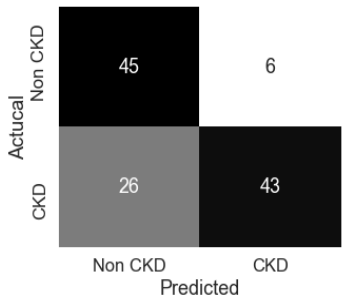
The following term will explain the lime process:



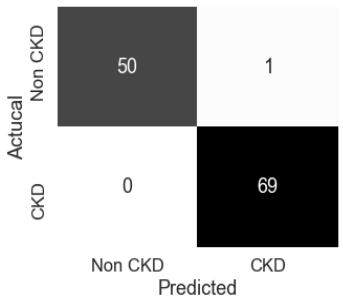
(a) SVM

(b) DT

(c) LR



(d) KNN



(e) RF

Fig. 3: Confusion Matrix of Classifier

- Local: It tells us about the behaviour of the classifier for predicting the disease or output.
- Interpretability: Lime can easily explain the classifier in its own words. Classifiers are complex to understand, but it is important to understand the classifier as it is used in the medical field.
- Model-agnostic: LIME can't directly go into the model. Inputs are perturbed around its nearer data points to check which feature contributes to the output prediction. Then perturb input is weighted by their nearness to the actual data points.

III. Results and Discussions

In this section, the aim is to explore a model which is better than any other model for the early prediction of CKD. For the early detection compared, five machine learning algorithms were based on the following performance factors: Precision, F-measure, Accuracy, and Recall [19]. Algorithms are applied to a dataset consisting of 400 instances and 25 attributes.

As shown in the Table II, RF, SVM, DT, and LR have a precision of 98.57%, 98.49%, 95.59%, 91.17% while KNN has the lowest precision of 87.76%. The recall of RF, SVM, DT, LR, KNN is 100%, 94.20%, 94.20%, 89.76%, 62.32 shows that among all five algorithms, the RF has the highest recall. The F-measure of RF, SVM, DT, LR, KNN is 99.28%, 96.29%, 94.89%, 90.51%, 72.88%, the F-measure of RF is highest.

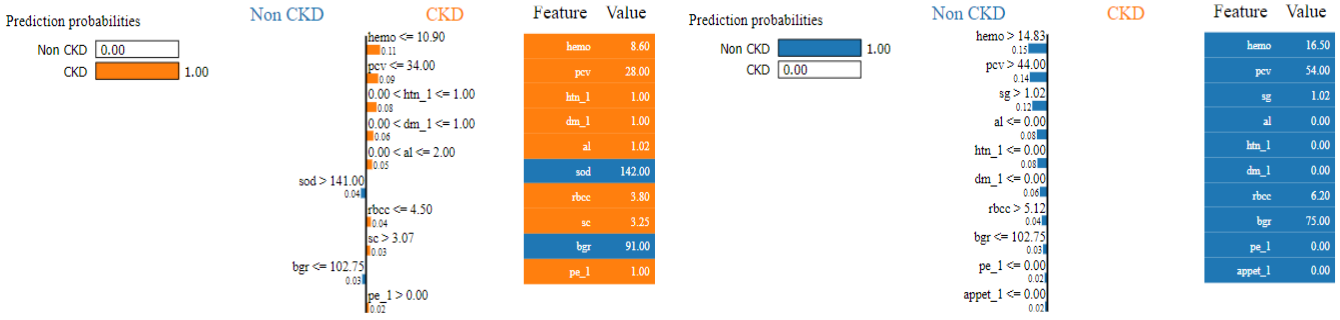
Now the accuracy of RF, SVM, DT, LR, KNN is 99.17%, 94.17%, 89.17%, 95.84% 73.34%. After comparing all the reliability measures like Precision, Recall, Accuracy, F-measure of all five machine learning algorithms, the result shows that RF is the top model among all other models for predicting CKD without XAI.

TABLE II: Performance analysis of classifier in (%)

Classifiers	Accuracy	Precision	Recall	F1-score
RF	**99.17**	**98.57**	**100**	**99.28**
SVM	95.84	98.49	94.20	96.29
DT	94.17	95.59	94.20	94.89
LR	89.17	91.17	89.86	90.51
KNN	73.34	87.76	62.32	72.88

Fig. 3 represents the confusion matrix of all five classifiers. Diagonal values represent the correctly predicted Non-CKD values and CKD values, and off-diagonal values represent the incorrectly predicted Non-CKD values and incorrectly predicted CKD values. Correctly predicted Non-CKD values for machine learning classifiers SVM, DT, LR, KNN, and RF are 50, 48, 45, 45, and 50. Correctly predicted CKD values for SVM, DT, LR, KNN, and RF are 65, 65, 62, 43, and 69. Incorrectly predicted CKD values for SVM, DT, LR, KNN, and RF are 1, 3, 6, 6, and 1. Incorrectly predicted Non-CKD values for SVM, DT, LR, KNN, and RF are 4, 4, 7, 26, and 0.

As shown in Table II and shown in Fig. 3(e), the RF got the best accuracy out of all five algorithms applied, as well seen in the matrix highly classified value among all other classifiers. So, now to get a better understanding, apply XAI on the RF algorithm as shown below in Fig. 4. With the help of XAI, found the features that help predict CKD more efficiently. As shown in Fig. 4, the features XAI found are hemoglobin, packed cell volume, diabetes mellitus, hypertension, albumin,



(a) For CKD

(b) For Non-CKD

Fig. 4: LIME Outcomes for RF Identification of CKD or Non-CKD

serum creatinine, pedal edema, blood glucose random, red blood cell count sodium, and pedal edema. The hemoglobin level is less than or equal to 10.9, the packed cell volume level is less than or equal to 34, the hypertension level is between 0 to 1, the diabetes mellitus level between 0 1, the albumin level is between 0 to 2, red blood cell count level less than or equal to 4.5, serum creatinine level greater than 3 and pedal edema level greater than 1 will lead to CKD. By using XAI get the range of values of features for which a model like an RF will make predictions.

IV. Conclusion

CKD is a chronic kidney disease that shows no symptoms at an early stage, which leads to kidney failure or even death. An earlier stage diagnosis of CKD, this work introduced five machine learning models to predict. In this paper, Machine Learning classifiers are RF, SVM, DT, LR, and KNN. Out of five, RF gave the best results having Accuracy (99.17%), Precision (98.57%), Recall (100%), and F-measure (99.28%). The advantage of this model is that the prediction takes lesser time and helps doctors to initiate treatment at the early stage. The used data set is smaller, so preferred a larger dataset in the future compared to the result of this dataset to learning model.

References

[1] S. D. Fraser and T. Blakeman, "Chronic kidney disease: identification and management in primary care," *Pragmatic and observational research*, pp. 21–32, 2016.

[2] F. P. Schena, V. W. Anelli, D. I. Abbrescia, and T. Di Noia, "Prediction of chronic kidney disease and its progression by artificial intelligence algorithms," *Journal of Nephrology*, vol. 35, no. 8, pp. 1953–1971, 2022.

[3] A. Yechake, *Knowledge Based System to Support Chronic Kidney Failure Diagnosis and Treatment*. PhD thesis, 2021.

[4] K. J. Jager, C. Kovesdy, R. Langham, M. Rosenberg, V. Jha, and C. Zoccali, "A single number for advocacy and communication—worldwide more than 850 million individuals have kidney diseases," 2019.

[5] A. Vijayvargiya, A. Panchal, A. Parashar, A. Gautam, J. Sharma, and R. Kumar, "Deep learning frameworks for covid-19 detection," in *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 1048–1053, IEEE, 2021.

[6] S. Gowda, P. B. Desai, S. S. Kulkarni, V. V. Hull, A. A. Math, and S. N. Vernekar, "Markers of renal function tests," *North American journal of medical sciences*, vol. 2, no. 4, p. 170, 2010.

[7] P. Liang, J. Yang, W. Wang, G. Yuan, M. Han, Q. Zhang, and Z. Li, "Deep learning identifies intelligible predictors of poor prognosis in chronic kidney disease," *IEEE Journal of Biomedical and Health Informatics*, 2023.

[8] N. G. Raju, K. P. Lakshmi, K. G. Praharshitha, and C. Likhitha, "Prediction of chronic kidney disease (ckd) using data science," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, pp. 642–647, IEEE, 2019.

[9] Nikhila, "Chronic kidney disease prediction using machine learning ensemble algorithm," in *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, pp. 1–6, IEEE, 2021.

[10] M. A. Islam, S. Akter, M. S. Hossen, S. A. Keya, S. A. Tisha, and S. Hossain, "Risk factor prediction of chronic kidney disease based on machine learning algorithms," in *2020 3rd international conference on intelligent sustainable systems (ICISS)*, pp. 952–957, IEEE, 2020.

[11] P. Chittora, S. Chaurasia, P. Chakrabarti, G. Kumawat, T. Chakrabarti, Z. Leonowicz, M. Jasiński, Ł. Jasiński, R. Gono, E. Jasińska, *et al.*, "Prediction of chronic kidney disease-a machine learning perspective," *IEEE Access*, vol. 9, pp. 17312–17334, 2021.

[12] R. Devika, S. V. Avilala, and V. Subramaniyaswamy, "Comparative study of classifier for chronic kidney disease prediction using naive bayes, knn and random forest," in *2019 3rd International conference on computing methodologies and communication (ICCMC)*, pp. 679–684, IEEE, 2019.

[13] A. Charleonnan, T. Fufaung, T. Niyomwong, W. Chokchueypattanakit, S. Suwannawach, and N. Ninchawee, "Predictive analytics for chronic kidney disease using machine learning techniques," in *2016 management and innovation technology international conference (MITicon)*, pp. MIT–80, IEEE, 2016.

[14] V. Kunwar, K. Chandel, A. S. Sabitha, and A. Bansal, "Chronic kidney disease analysis using data mining classification techniques," in *2016 6th International Conference-Cloud System and Big Data Engineering (Confluence)*, pp. 300–305, Ieee, 2016.

[15] S. Akter, A. Habib, M. A. Islam, M. S. Hossen, W. A. Fahim, P. R. Sarkar, and M. Ahmed, "Comprehensive performance assessment of deep learning models in early prediction and risk identification of chronic kidney disease," *IEEE Access*, vol. 9, pp. 165184–165206, 2021.

[16] P. A. Moreno-Sánchez, "Data-driven early diagnosis of chronic kidney disease: Development and evaluation of an explainable ai model," *IEEE Access*, vol. 11, pp. 38359–38369, 2023.

[17] A. Vijayvargiya, P. Singh, R. Kumar, and N. Dey, "Hardware implementation for lower limb surface emg measurement and analysis using explainable ai for activity recognition," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–9, 2022.

[18] B. Dhanka, A. Vijayvargiya, R. Kumar, and G. Singh, "A comparative assessment of machine learning techniques for epilepsy detection using eeg signal," in *2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, pp. 1–6, IEEE, 2020.

[19] A. Vijayvargiya, N. Kumari, P. Gupta, and R. Kumar, "Implementation of machine learning algorithms for human activity recognition," in *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*, pp. 440–444, IEEE, 2021.