Detection of Chronic Kidney Disease via comparing Multiple Algorithms

1st Ram Jee Dixit
Department of MCA
ABES Institute of Technology, Ghaziabad
AKTU, Lucknow, India
dikshitaram@gmail.com

2nd Vivek Kumar Pandey
Department of CSE
United College of Engineering and Research, Prayagraj
AKTU, Lucknow, India
Vivek5confidencial@gmail.com

3rd Avdhesh Kumar Tiwari
Department of CSE,
ABES Institute of Technology, Ghaziabad
AKTU, Lucknow, India
Eravdhesh77@gmail.com

***Abstract*— Millions of people throughout the world suffer from CKD (Chronic Kidney Disease), which significantly affects the rates of mortality and morbidity.To slow progression of the CKD while preventing respective complications, prompt identification and therapy are essential. Acknowledgement to the improved measures of variables that can define via predictive modeling, we are able to create a set of predictive models with the use of machine learning and prediction analytics. In fact, Classifiers based on machine learning algorithms have been studied thoroughly showing that they are thebest performance indicators which reveal that innovative logics and better machine learning models produce superior outcomes. The most crucial variables for the prediction model were chosen after pre-processing the information and selecting the features. The models' performance was assessed using their accuracy, specificity, precision, and sensitivity and F1-score metrics. The findings suggest that Decision Tree algorithms produced more precise and effective detection outcomes. Specific gravity, hypertension, hemoglobin, diabetes mellitus, albumin, appetite, and red blood cell count were among the most crucial factors for predicting CKD. In summary, decision tree algorithms, in particular Random Forest can be useful for predicting CKD and can help with early detection and management of the condition. According to the methodology applied in the study, it can be said that finding newer results that can be useful in accessing the mandatory predictions in the area of renal disease and beyond is an exciting way to take advantage of current advancements in machine learning and predictive modeling**.

***Keywords*— Chronic Kidney Disease (CKD), Decision tree, Deep Learning, Machine Learning**

I. INTRODUCTION

Data mining is the analysis or discovery of useful knowledge used to create a meaningful collection of data from a huge amount of information reservoir. The majority of health care record data is beautiful and comes from a variety of sources around the world, including sensor equipment, photographs, and text stored in electronic records. In this method, data were gathered and depicted using hints from various trials, which were then handled collectively and the original data were analyzed. A major global health crisis is CKD. The World Health Organization acknowledged that there were 58% more deaths and 35% more people with chronic renal disease. Numerous surveys have been conducted over the globe, and many logical predictions have been made. It is now estimated that 850 million individuals would have renal disease by2050. With estimated 850 million individuals projected to have renal disease from a variety of reasons, CKD is the sixth fastest growing cause of disaster and death, accounting for

The blood is filtered through the kidneys as part of its function. It gets rid of extra blood and controls fluid and electrolyte stability. It filters blood, and the kidney's two bean-shaped structures produce urine. Each kidney is surrounded by one million nephron-like structures. The main cause of CKD is diabetes, but other risk factors include hypertension, smoking, being overweight or obese, having a heart condition, having a history of heart disease in the family, drinking alcohol, and being older. We may experience a variety of symptoms, including rash, back pain, stomach pain, nosebleeds, diarrhea, vomiting and fever, if one or both of human kidneys are not functioning properly. The list of warning indicators includes things like changes in urinary functionality, bulges and pain, plasma in the urine, weakness, and extreme fatigue.

CKD (Chronic Kidney Disease), is also known as Chronic Kidney Failure or Chronic Renal Disease, It is a potentially fatal condition caused by the kidneys' failure to carry out their normal functions. It is a recognized health problem that causes Glomerular Filtration Rate (GFR) to continuously degrade for three months or longer. Different Levels of CKD:

i. Acute Prerenal Kidney Failure
ii. Acute Intrinsic Kidney Failure
iii. Chronic Prerenal Kidney Failure
iv. Chronic Intrinsic Kidney Failure

Usually, in the early stages victim does not en counters any symptoms related to CKD. This is because a significant decline in kidney function can usually be tolerated by the human body. Unless a normal examination for another condition, such as a blood or urine test, uncovers a potential concern, kidney disease is frequently not recognized until this stage. If it is found early on, prescription treatment and periodic screenings may help stop it from advancing to a more advanced form.

It's in its later phases. There may be a lot of indicators if the kidney disease is not detected early or keeps growing worse despite the treatment. CKD's last stage is kidney failure. It is also known as established renal failure or end- stage renal disease. At some point, it’s likely that a transplant of kidneys or kidney dialysis process will be required.

II. LITERATURE REVIEW

Several research papers are available in order to carry research activities on chronic kidney disease. Different researcher proposed their own work for understanding the kidney disease. In Table 1, we are going to present research work and study of several authors to get insights of CKD.

ISBN: 978-93-343-1044-3

TABLE 1: Literature Review

Table :Literature Survey Table				
Year	*Author*	*Title*	*Approach*	*Result*
2024	Imran, Muhammad, et al.	Predictive modeling of chronic kidney disease using extra tree classifier: A comparative analysis with traditional methods.	A number of machine learning algorithms are used, and their performances are contrasted.	Some machine learning algorithms are more accurate than other.
2024	Arora, Aditi, Chirag Sehgal, and Neha Agarwal	An Analysis of Machine Learning Algorithms for Chronic Kidney Disease Prediction	Development of prediction model using various machine learning algorithms	Random Forest algoritjm provides more accuracy
2023	Arif, Muhammad Shoaib, Aiman Mukheimer, and Daniyal Asif.	Enhancing the early detection of chronic kidney disease: a robust machine learning model	UCI-CKD dataset ML algorithms like Naïve Bayes, KNN.	Robust ML model for CKD detection with 100% accuracy
2023	Khalil, Nariman, Mohamed Elkholy, and Mohamed Eassa.	A comparative analysis of machine learning models for prediction of chronic kidney disease	Multiple ML models are trained	LR and KNN have accuracy of 99%
2023	El Sherbiny, Moataz Mohamed, et al	Classification of chronic kidney disease based on machine learning techniques	Dataset from Irvin ML, 9 ML models	AdaBoost have 99.17% accuracy.
2022	Silveira, Andressa CM da, et al.	Exploring early prediction of chronic kidney disease using machine learning algorithms for small and imbalanced datasets.	Medical record of adults age>=18. Decision Tree Model .	Decision Tree model provides accuracy of 98.99%.
2021	Ilyas, Hamida, et al	Chronic kidney disease diagnosis using decision tree algorithms	Used machine learnig algorithm like decision tree and random forest.	Radom Forest provides accuracy of 85%.
2020	Yashfi, Shanila Yunus, et al.	Risk prediction of chronic kidney disease using machine learning algorithms.	Prediction UCI ML dataset	Random Forst provides accuracy of 97 %.
2020	Lateef, Z.	A Comprehensive Guide to Random Forest in R	R code for Random Forest	Model implemented in R.
2017	Gupta, Bhumika, et al.	Analysis of various decision tree algorithms for classification in data mining	Various Decision tree algorithms	Understanding of decision tree algorithms.
2017	Webster, Angela C., et al.	Chronic kidney disease	Description of chronic kidney desease	Understad about CKD

III. PROPOSED METHODOLOGY

Numerous researchers have expressed interest in examining CKD datasets to create potentially effective models that could aid healthcare practitioners in keeping track of probable CKD patients. That is the main contribution of this work, which we hope adds value and supports the healthcare given to identify specific CKD cases at early stages based on the presented feature in our feature selection section of this paper. To the best of our knowledge, no work done in the literature has described the classification and feature section and/or sensitivity analysis of a CKD dataset.

The initial goal is to use the Chronic Kidney Disease dataset to make an early diagnosis of CKD patients with risk level. In light of the fact that this disease affects a large number of patients worldwide, this objective is important for current research. The evaluations will be carried out by using Jupyter Notebook web tool utilizing the Python 3.8 programming language. The Python- based Scikit-learning packages, a free platform for machine learning process were used in a number of the studies related to the kidneys. This analysis takes into consideration the various factors like

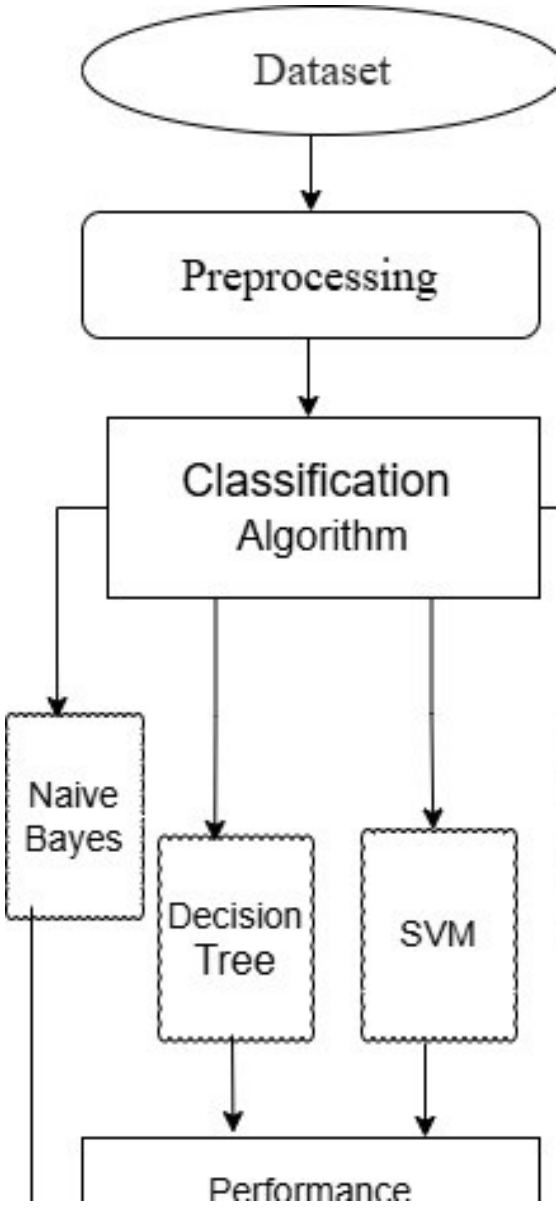

Fig-1: Flow of Proposed Methodology

ISBN: 978-93-343-1044-3

Table : Variable Descriptor Table			
	Attribute Symbol and Description	*Type*	*Class*
1	age (Age)	Numerical	Predictor
2	bp (Blood pressure)	Numerical	Predictor
3	sg (Specific gravity)	Nominal	Predictor
4	al (Albumin)	Nominal	Predictor
5	su (Sugar)	Nominal	Predictor
6	rbc (Red Blood Cells)	Nominal	Predictor
7	pc (pus cells)	Nominal	Predictor
8	pcc (Pus cells Clumps)	Nominal	Predictor
9	rc (Race)	Nominal	Predictor
10	bgr (Blood Glucose Random)	Numerical	Predictor
11	bu (Blood urea)	Numerical	Predictor
12	sc (Serum creatinine)	Numerical	Predictor
13	sod (Sodium)	Numerical	Predictor
14	pot (Potassium)	Numerical	Predictor
15	hemo (Hemoglobin)	Numerical	Predictor
16	pcv (Packed Cell Volume)	Numerical	Predictor
17	sex (Sex)	Nominal	Predictor
18	rc (Red Blood Cells count)	Numerical	Predictor
19	htn (Hypertension)	Nominal	Predictor
20	dm (Diabetes Mellitus)	Nominal	Predictor
21	cad (Coronary artery disease)	Nominal	Predictor
22	appet (Appetite)	Nominal	Predictor
23	pe (Pedal Edema)	Nominal	Predictor
24	ane (Anemia)	Nominal	Predictor
25	class (Class)	Nominal	Target

Table 1: variable Descriptor table

Specificity, AUC (area under the curve), sensitivity, and accuracy as determined by the F1-measurement which are the evaluation criteria. In various studies, the Random Forest Algorithm, AdaBoost Classifier, Gradient Boosting, Logistic Regression, Naive Byes, KNN, and Kernel SVM were used, among other machine learning techniques, in order to analyze the CKD dataset. The missing numerical values were computed at the preprocessing step using the mean method, and the mode approach was used to compute the missing nominal values. In this work, CKD was identified using a variety of classifiers, including SVM, KNN, AdaBoost Classifier, and others.

A. DATASET

The dataset for the respective system has been opted out from an open source data set named as kaggle. The dataset is represented in variable description table**.** All attributes derived from the data are shown in Table's attribute symbols and descriptions, and the type column displays the datatype of attributes, whereas class as the third column of the table really categorizes the dataset's attributes into two categories, namely predictor and target. Target will be predicted using respective attributes. The class/stage of chronic renal disease will be projected using all predictor variables shown in Fig-2.

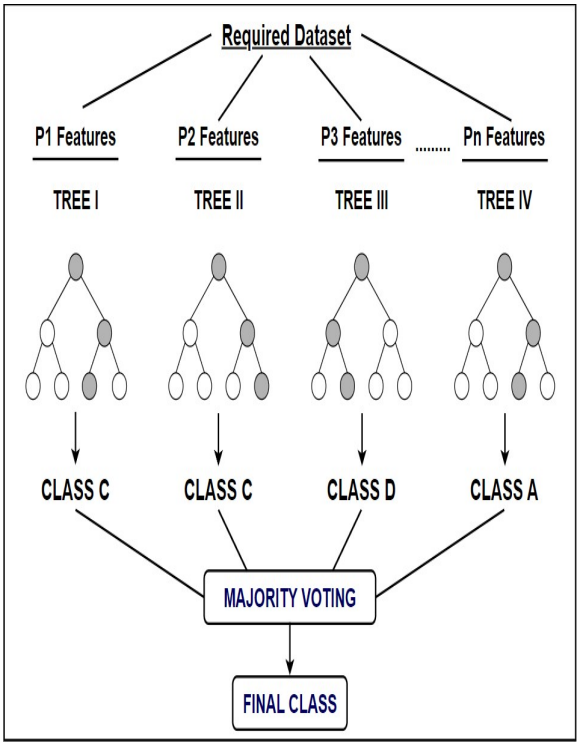

Fig-2; Required Data

stage indicates normal kidney function; II stage represents mildly impaired kidney function, and so forth. The suggested method further separates the CKD class into these various stages. 3A stage of CKD indicates comparatively reduced kidney function, 3B stage indicates significantly reduced kidney function, stage indicates severely reduced kidney function, and V stage indicates end-stage renal failure based on the computed GFR values.

B. PREPROCESSING

The gathering of a patient dataset with CKD marks the start of this phase. Numerous mathematical formulas are used in processing the estimation of GFR; however, in the respective investigation the CKD-EPI Equation [9] is used. A supposed to the Modification of Dietin Renal Disease (MDRD) equation which only considers gender, ethnicity and age as well as serum creatinine and is also known to be accurate only when GFR is greater than 60,which is the case for later stages of CKD so this equation can be considered reliable for the computation of all stages of CKD.

The most important phase in data mining techniques is data preparation, which comprises cleaning, extracting, and converting data into a format that can be used by machines. Raw data is a nightmare for machine learning prediction because it has inaccurate information, missing information, and unsuitable formats. The variables that are not present in the database must be filled in for the entire CKD dataset. The procedures can be synchronized in some cases to create discrete characteristics when continuous features are present. Each one of them contains values that are erratic and some missing values. In order to improve the behavior of medical data, the real data is preprocessed. Data preprocessing can be defined as the method of changing raw data to make it suitable for insertion in a machine learning model.

ISBN: 978-93-343-1044-3

The absent values in the CKD dataset were processed and infused after the categorical variables were encoded. This phase is considered completed once the category variables had been encoded. The median value of the pertinent variable which is considered for samples is used to fill in the gaps when numerical variables have missing values. On the other hand the category from the completed samples that appears most frequently in the variable is chosen to fill in the missing values for the category variables. This is done to make sure that every sample is fairly represented.

C. CCOMPUTATION

A substance's rate of the clearance from plasma is estimated to determine the GFR which is defined as the volume of plasma filtered by glomeruli per unit time. It is regarded as one of the finest characteristics to gauge renal health and gauge the severity of chronic kidney disease [3]. Filtration markers, a kidney-excreted material, are used to compute the GFR value. The GFR is then calculated using a formula that uses the clearance of the filtration marker. There are other mathematical formulae that can be used to estimate GFR, but the following equations are the most frequently used ones:

Equation1: Modification of Dietin Renal Disease (MDRD)

Equation 2: Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI).

Modification of Diet in Renal Disease (MDRD) method is considered to be much more accurate than the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) for predicting the glomerular filtration rate (eGFR). As a result, in the proposed work, we will calculate GFR using the respective CKD-EPI equation.

To determine the matching person's GFR, the equation (CKD- EPI) requires four inputs: sex, race, serum creatinine, and age. Execution time and classification accuracy are used as performance indicators when comparing classification algorithms. The classification's performance evaluation is then completed.

D. CLASSIFICATION ALGORITHMS

Random Forest Algorithm

An ensemble machine learning approach commonly known as the Random Forest algorithm integrates various decision trees to upgrade the robustness and accuracy of the respective model. Each decision tree in a random forest is built independently using a randomly picked subset of the characteristic features and a subset of the information used in training. Bootstrapped sample of the training data used by the Random Forest technique is used to build a huge number of decision trees, each of which is trained on a random portion of the features. The algorithm chooses the best feature among a randomly chosen subset of features at every split throughout the tree construction process rather than taking into account all the features [1].

AdaBoost Algorithm

The Boosting phenomenon called as AdaBoost algorithm, sometimes referred to as Adaptive Boosting, is used as an Ensemble Method in machine learning techniques. The weights are redistributed to each instance, with much higher weights being given to instances that were mistakenly categorized, hence is referred to as "adaptive boosting". AdaBoost can be used to upgrade the performance of any machine learning algorithm. While teaching or training slow learners, it works the best. These are models whose have categorization accuracy is slightly better than random chances. The method that works best with AdaBoost and is frequently processed is known as single level decision tree [2].

Gradient Boosting

For regression and classification type of functions, gradient boosting is a machine learning technique that specifies a machine learning model in the form of weak predictive models. Through the gradual construction of a model, this boosting method generalizes the whole model by enabling the optimization of any differentiable loss function. Gradient boosting essentially iteratively combines weak learners into a single powerful learner. A new model is fitted for each additional weak learner to give a more accurate estimate of the response variable. By combining a number of relatively weak prediction models, gradient boosting seeks to create a stronger prediction model [4].

Naïve Bayes Algorithm

The probabilistic machine learning algorithm is known as the Naïve Bayes algorithm is utilized for classification problems. It is founded on the Bayes theorem, according to which the likelihood of a hypothesis given certain observable evidence (features) is inversely proportional to the likelihood of the evidence given the hypothesis, multiplied by the prior likelihood of the hypothesis.

The algorithm learns the probability distribution of the features for each class label from a training dataset before applying it to classification. The likelihood of a new instance belonging to each class label is then determined given a new instance with an unknown class label using the learnt probability distributions. The instance is then given the class label with the highest probability [3].

K-nearest Neighbor (KNN) Algorithm

KNN (K-nearest Neighbors) algorithm is stated to be a machine learning algorithm used for functioning of classification and regression tasks. It is required to predict the class label of a new instance in the classification task using the k- nearest examples from the training set, where k is a user-defined parameter. The distances between the new instance and every other instance in the training set are first calculated by the KNN algorithm. Depending on the chosen distance metric, such as Manhattan distance or Euclidean distance, it then chooses the k instances that are closest to the new instance.

In the classification task, the algorithm labels the new instance with the class that is most prevalent among the k closest neighbors. The approach calculates the value of the new instance in the regression job as the mean of the k closest

ISBN: 978-93-343-1044-3

Logistic Regression

When attempting to predict a binary outcome, such as whether or not a consumer would buy a product, machine learning algorithms like logistic regression are used. The technique uses a logistic function to represent the relationship between the binary target variable and the predictor variables, commonly known as features.

Any real-valued input is converted by the logistic function, also referred to as the sigmoid function that transforms any real-valued input into a value between 0 and 1, which can be stated as a probability [6].

Kernel SVM

Kernel Support Vector Machine (SVM) can handle non-linearly separable data by transferring it to a more advanced feature space using a kernel function. Kernel Support Vector Machine (SVM) is a machine learning approach used for binary classification and regression applications. The SVM algorithm seeks to locate the hyper plane in the modified feature space that maximizes the margin between the two classes.

Without explicitly calculating the coordinates of the modified data, the kernel function is a similarity measure that computes the dot product between two points in the feature space. The techniques can effectively handle high-dimensional data and non-linear decision boundaries thanks to the kernel function. SVM performs relatively well when there is a clear line of separation between classes [4].

Decision Tree

A decision tree is considered to be a machine learning approach that creates a model which nearly resembles a tree to make judgments based on a number of queries or requirements regarding the input data. The tree has nodes and branches, where each node denotes a choice or question based on a characteristic of the input data and each branch denotes a potential result of that decision.

The first step of the Decision Tree algorithm is to choose the feature that divides the data into distinct subgroups while maximizing information gain or minimizing impurity. The method recursively divides the respective data into subsets according to the chosen characteristics and their values up until a stopping requirement, such as a maximum depth or a minimum number of samples per leaf are satisfied [5].

IV. Performance Evaluation

Following are mathematical connections that are used to calculate the sensitivity, F-Measure, specificity, accuracy and confusion matrix in order to assess the performance of classification.

Accuracy

It can be considered as the ratio between the correctly classified samples to the sum of total number of samples.

$$Accuracy = \frac{(tn + tp)}{(tn + tp + fn + fp)}$$

Specificity

Inverse recall or True Negative Rate (TNR) are other names for the following method. The ratioof accurately classified Negative instances to all negative instances are measured.

$$Specificity = \frac{(tn)}{(fp + tn)}$$

Sensitivity

It is also known as hit rate, recall, and TPR (true positive rate). This displays proportion of the correctly identified Positive instances to sum of all the positive instances.

$$Sensitivity = \frac{(tp)}{(fp + tn)}$$

F-Measure

The weighted average of sensitivity and precision values is used to generate F-Measure. For the estimation of classification performance, F- Measure turns to the subject of information retrieval.

$$F - measure = \frac{(2 * Sensitivity * Precision)}{(Sensitivity + Precision)}$$

Precision

Precision is defined as what proportion of positive identifications was actually correct.

$$Precision = \frac{(tp)}{(fp + tp)}$$

Confusion matrix

A model's predictions are tabulated and shown in the confusion matrix. Numerous both accurate and inaccurate forecasts are displayed. The k-test data is compared with the classification findings of the respective table which is used to determine the decisive forecast. The matrix is represented as an X-by-X matrix, where X can be stated as the number of classes in the dataset. Confusion matrix is a influential tool for determining a classifier's accuracy.

According to the stated formulas above, the letters F.P, F.N, T.P, and T.N stand for False Positive, False Negative, True Positive, and True negative values respectively.

V. Experimental Results

We have evaluated the accuracy of every algorithm using k-fold cross-validation, and the results are as presented in Table 3. The results can be also represented using chart in Fig-3.

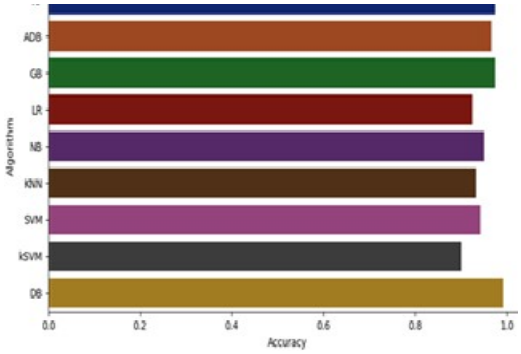

Fig-3: Chart for algorithm comparison

ISBN: 978-93-343-1044-3

TABLE 3: Accuracy Comparison Table

Table : Variable Descriptor Table		
S. No.	*Algorithm*	*Accuracy (%)*
1	*Decision Tree*	*99.17*
2	*Random Forest*	*97.61*
3	*KNN (K-Nearest Neighbors)*	*93.33*
4	*AdaBoost Classifier*	*96.76*
5	*Kernal SVM*	*90.11*
6	*Naïve Bayes*	*95.00*
7	*Logistic regression*	*92.50*
8	*SVM(Support Vector Machine)*	*94.17*
9	*Gradient Boosting*	*97.50*

Based on the results, we chose Random Forest as the final algorithm for CKD prediction due to its high accuracy of 97.6%. The Random Forest algorithm provides a reliable prognosis of the CKD, allowing for early detection and treatment of the disease.

VI. CONCLUSION

F-Measure is produced using the weighted average of the sensitivity and precision measurements. F-Measure goes to the topic of information retrieval for the measurement of classification performance. It allows a model to be evaluated taking both the precision and recall into account using a single score, which is helpful in describing the performance of the model.

The decision tree algorithm has a simple visualization and less accurate forecasts than the random forest, which has a complex visualization. The benefits of Random Forest are its ability to forecast outcomes more accurately and prevention of overfitting.

As a result, Random Forest Algorithm can be stated to be accurate and efficient when compared to other algorithms, demonstrating that it produces results that are more accurate. Multiple decision trees are constructed by Random Forest Algorithm, which then combines them to create a reliable prediction model. The algorithm is exceptionally good at making predictions in real time because of this method. Accordingly, we suggest employing Random Forest to assist doctors in creating an automated decision support system for detecting CKD based on our findings.

VII. FUTURE WORK

It is necessary to have a larger dataset in order to accommodate enough samples because adding more data will increase accuracy and efficiency. With enough samples, the prediction can be expanded to include and pinpoint geographic locations that are more susceptible to CKD.

Additionally, by using information on genetics, water consumption habits, and food kinds in the research, better understanding of CKD can be acquired.

In future we will apply deep learning techniques to detect chronic kidney diseases. Also we use large datasets for detection of CKD. In future we will enhance accuracy of disease detection

REFERENCES

[1] Imran, Muhammad, et al. "Predictive modeling of chronic kidney disease using extra tree classifier: A comparative analysis with traditional methods." *Journal of Computing & Biomedical Informatics* 6.02 (2024): 261-271.

[2] Arora, Aditi, Chirag Sehgal, and Neha Agarwal. "An Analysis of Machine Learning Algorithms for Chronic Kidney Disease Prediction." *2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. IEEE, 2024.

[3] Arif, Muhammad Shoaib, Aiman Mukheimer, and Daniyal Asif. "Enhancing the early detection of chronic kidney disease: a robust machine learning model." *Big Data and Cognitive Computing* 7.3 (2023): 144.

[4] Khalil, Nariman, Mohamed Elkholy, and Mohamed Eassa. "A comparative analysis of machine learning models for prediction of chronic kidney disease." *Sustainable Machine Intelligence Journal* 5 (2023): 3-1.

[5] El Sherbiny, Moataz Mohamed, et al. "Classification of chronic kidney disease based on machine learning techniques." *Indones. J. Electr. Eng. Comput. Sci* 32.2 (2023): 945-955.

[6] Silveira, Andressa CM da, et al. "Exploring early prediction of chronic kidney disease using machine learning algorithms for small and imbalanced datasets." *Applied Sciences* 12.7 (2022): 3673.

[7] Ilyas, Hamida, et al. "Chronic kidney disease diagnosis using decision tree algorithms." *BMC nephrology* 22.1 (2021): 273.

[8] Yashfi, Shanila Yunus, et al. "Risk prediction of chronic kidney disease using machine learning algorithms." *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. IEEE, 2020. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed.,vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[9] Lateef, Z. "A Comprehensive Guide to Random Forest in R." *Edureka. co* (2020).

[10] Gupta, Bhumika, et al. "Analysis of various decision tree algorithms for classification in data mining." *International Journal of Computer Applications* 163.8 (2017): 15-19.

[11] Webster, Angela C., et al. "Chronic kidney disease." *The lancet* 389.10075 (2017): 1238-1252.

[12] Subasi, Abdulhamit, Emina Alickovic, and Jasmin Kevric. "Diagnosis of chronic kidney disease by using random forest." *CMBEBIH 2017: Proceedings of the International Conference on Medical and Biological Engineering 2017*. Springer Singapore, 2017.

[13] George, Cindy, et al. "Chronic kidney disease in low-income to middle-income countries: the case for increased screening." *BMJ global health* 2.2 (2017).

[14] Salekin, Asif, and John Stankovic. "Detection of chronic kidney disease and selecting important predictive attributes." *2016 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, 2016.

[15] Vijayarani, S., and S. Dhayanand. "Kidney disease prediction using SVM and ANN algorithms International Journal of Computing and Business Research (IJCBR)." (2015).

ISBN: 978-93-343-1044-3