Journal of Rare Cardiovascular Diseases
ISSN: 2299-3711 (Print) | e-ISSN: 2300-5505 (Online)
www.jrcd.eu

RESEARCH ARTICLE

# An improved methodology for diagnosis of Chronic Kidney Disease using Machine Learning Techniques

[1] Mahesh Babu, [2]Susmitha
[1, 2] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation,Vaddeswaram.,Guntur District, Andhra Pradesh 522302, India

*Corresponding Author

Article History
**Received:** 06.10.2025
**Revised:** 30.10.2025
**Accepted:** 12.11.2025
**Published:** 04.12.2025

**Abstract:** Chronic kidney disease is a widespread health problem because it has a high death and illness rate and can lead to other health problems. People who have chronic kidney disease often don't pay attention to their illness because it doesn't show any clear signs at first. When people with chronic kidney disease are diagnosed early, they can start treatment right away, which slows the illness's progression. Models based on ML might help doctors reach this goal more quickly and correctly, as they use quick and accurate markers. We have a ML model that uses six distinct strategies at the same time to identify persons with chronic renal disease. These include RF, logistic regression, naive bayes, KNN, SVM, and feed forward neural network. There were many missing values in UCI's ML-oriented library, which included this type of chronic kidney disease data. To fill in the missing numbers, we utilized KNN imputation. One way to do this is to pick a few complete specimens whose findings are the most like the missing information in each incomplete specimen. To figure out how to identify chronic kidney diseases, we need to be able to use six different machine learning methods at the same time to deal with the missing data set. The ratios for training and testing would then be 80/20. Finally, we use standard performance curves to get an idea of how well the six ML methods work and then check our results.

**Keywords:** Healthcare, Kidney, Chronic kidney disease, Machine Learning, and imputation".

## INTRODUCTION

One out of ten people in the world has CKD which is a severe health problem to the population [1, 2]. The chronic renal disease was widespread between 10 and 15 percent in America [4] as opposed to a total of more than 10 percent in China [3]. Other estimates show that the prevalence has been on the rise to over 14 percent among Mexicans [5]. The distinguishing feature of this illness is a progressive deterioration of renal functionality that comes to a final result in total renal failure. The chronic kidney disease does not have any external signs in its initial phases. Thus, the sickness may remain undiagnosed up until the kidney itself becomes impaired by about a quarter of regular capacity [6]. Also, chronic kidney disease is very high with morbidity and mortality, which influences the health of individuals worldwide [7]. It can predispose the risk of heart issues [8, 9]. Chronic kidney disease is in most instances, an incurable and persistent medical condition [10]. It is thus paramount to diagnose and detect CKD at early stages so that the patients can be treated in order to restrain the spread of the illness.

To derive the features of the concerned trend, machine learning is a computer method that approximates and identifies task-specific data [11]. This technique can be a feasible way of detecting chronic kidney disease as it is able to detect diseases effectively and cheaply. Information technology has evolved to become a new form of therapeutic resource [12] and the quick growth of digital medical records [13] has provided it with a diverse usage range.

ML has also been applied in medicine in the past to detect various diseases, asses disease-related variables and also the physical status of individuals [14, 15]. Various forms of structures that were created by ML-oriented algorithms have been applied to the detection of hyperglycemia and retinopathy [16,17], heart disease [18,19], cancer [20], severe kidney damage [21,22], and other diseases [23,24]. These types of structures involve the use of effective methods such as tree, regression, possibility, neural network and decision surface. Moreover, it is evident that the majority of these ML-based methods to recognize severe renal diseases have a restricted breadth of payments or lack precision due to the technique adopted by the task of inputting missing information.

### Motivation

ML algorithms can be used to analyze large sets of patient data to identify minimal tendencies that can be used in early risk prediction and action. This has the capacity of reducing the burden on healthcare departments, reduce the cost of healthcare and significantly improve patient outcomes. Moreover, ML will have an opportunity to design individual treatment programs, which will enhance therapy choices of a particular patient with chronic renal disease and, eventually, the quality of their lives.

### Problem Statement

As it is stated in the 2019 World Renal Day documentation, approximately 2.4 million individuals die each year because of kidney diseases [25]. Chronic kidney disease is now the sixth fastest growing cause of deaths across the world. Chronic renal disease has emerged as a challenging health issue in the world because of the rising prevalence [25].

One of the created methods used to predict and find different diseases, like breast cancer, heart disease, kidney problems, and a stroke [2628], is ML. Not only is this used in healthcare, but it's also used in other fields as energy-based business [2934]. Due to the recent breakthroughs in big data of digital health records, ML, the computational method of discovering correlations in large datasets with a number of complicated features, is making headway in the healthcare sector much faster [35]. Appropriate applications of machine learning forecasting algorithms can help identify various diseases and give patients high-quality and affordable medical assistance in time. Therefore, it can become an effective tool of chronic renal disease diagnosis [36].

The renal impairment and its subsequent effects are extremely challenging to manage in developing countries because of the unavailability of resources and specialists and because of the high cost of treatment [37, 38]. Early screening of chronic renal disease is therefore necessary to reduce the financial consequences and enhance the effectiveness of disease treatment [39]. Machine learning-based predictions can be used to find chronic kidney disease early on so that it can be treated more successfully and faster [40]. A lot of research has been done using well-known People with chronic renal disease can be identified using ML algorithms such as RF, SVM, and DT [25]. Even though the most well-known and current ML techniques were applied in these efforts, they were narrowed in their possibilities to evaluate single processes or minor groups of methods of how to diagnose and treat chronic renal disease. Moreover, computational optimization has been the focus of many studies in comparison to clinical application, and many have questioned the effectiveness of diagnosing and treating chronic renal disease overall. Consequently, there is a need to develop a more comprehensive and comparative science that can address any significant gaps in the research through the systematic evaluation and maneuvering of the myriad physiological indicators associated with chronic kidney disease.

## Objectives

Training a ML model on patient data to predict chronic kidney disease. Medical records that contain demographical information, test results, and patient history can be used to classify individuals as negative/positive of chronic renal disease. This project aims to make an accurate and useful prediction model that will help doctors find and treat chronic kidney disease early on, which will improve the health of patients.

Our goal is to find the significant aspects of the raw data, do the required preparation, and predict the serious kidney problems with the help of machine learning methods. This work will become the means of providing early and effective care of the risk factors that were discovered during any proper and safe identification of severe renal diseases. By simultaneously employing six separate methods, including RF, logistic regression, naive bayes, KNN, SVM, and feed forward NN, we develop a ML-based solution by virtue of the firm need to develop a more comprehensive and comparative methodological strategy. The methodology can seal a considerable gap in the literature because it will be possible to evaluate and navigate the complex physiological indicators of chronic kidney disease methodically.

## Literature Review

Murugesan et al. [41] employed fuzzy and adaptive neural fuzzy inference algorithms to find out if someone has kidney disease or chronic kidney disease. The primary aim was to improve the precision of medical diagnostics for disease identification. Setting up a neural fuzzy inference system that is fuzzy and adaptable requires a number of factors to be taken into account and they include the amount of glucose, blood pressure and height, weight, smoking, level of maturity, and workability of the nephron.

The study by Senan et al. [42] is novel in that it comes up with a diagnostic technique of diagnosing chronic renal diseases. The research will help practitioners investigate the possibilities of preventing or at minimum early diagnosing chronic kidney disease using ML methods. The major purpose of the research was to test a sample of 24 attributes that were gathered amongst 400 patients.

The study conducted by Dey et al. [43] was unique as it developed a diagnosis system that uses a hybrid feature selection method and several ML techniques to find people with chronic renal disease. The project used 400 medical records from people with CKD that were available in the ML repository.

Stages prediction through both binary and multiclassification has been carried out in the research conducted by Debal and Sitote [25]. RF, DT, and SVM are some of the forecast models that are used. Analysis of variance and cross-validation iterative feature removal were used to choose the features. Tenfold cross-validation was used to check how well the models worked.

The research undertaken by Farjana et al. [36] suggested nine machine learning models, viz. KNN, logistic regression, SVM, Naive Bayes, more tree classifiers, AdaBoost, Xgboost, and LightGBM. Through the construction of the prediction models based on a dataset on the disease with 400 records and 14 features, the most optimal of the classifiers that predict chronic kidney disease was determined.
Hassan et al. study [44] was distinguished by the fact that it was able to extract the best features of the data, to offer the best classification techniques to determine

people with chronic renal disease. Individuals with chronic renal disease were used to predict the condition using a number of well-known ML techniques using clinical records. K-means clustering was used after missing numbers were taken into account. After that, the XGBoost feature selection was used to pick out the features. After the dataset's features were found, various ML techniques were used to find the best classification models. These included a RF, a random tree, a neural network, a scoring tree model, and a support vector machine.

A study by Bai et al. [45] looked at how well the ML could predict the likelihood of EKSD in people who already had CKD. The results were based on longitudinal cohort of chronical renal disease. The predictor items were the baseline characteristics of the patients and the outcome of the routine blood testing. Interest established the presence or absence of ESKD at five years. Multiple imputation was used to impute the missing data. This was done by training and evaluating five machine learning algorithms on fivefold cross-validation. The Kidney Failure Risk Equation or KFRE was used in the comparison of the performance of each model.

Alturki et al. [46] talked about using cutting-edge ML on a dataset about chronic renal disease from the University of California, UC Irvine ML page. TrioNet, a combination model, was shown by the experts. that takes three ML methods to produce unbelievable predictions of chronic kidney disease. Also, SMOTE addressed the issue of class-imbalance and K-Nearest Neighbour imputer addressed the issue of missing data. The usefulness of the suggested model was assessed during an extensive comparison with the alternative machine learning techniques. The TrioNet presented, a combination of K-Nearest Neighbor imputer and SMOTE demonstrated a higher performance than previous models being identified as having the highest accuracy rate of 98.97 in the identification of chronic kidney disease.

Pal [47] made and tried a prediction model that could tell people who had chronic kidney disease what would happen. In medicine, ML techniques are used a lot to classify diseases and predict what will happen with them. Three ML models were used to look at a 25-feature dataset about chronic kidney disease that was stored in the UCI ML Repository. Then, the bagging ensemble method was used to make the model's output better. The groups of data about chronic renal disease were used to train the ML models. In addition, the Kidney Disease Collection had features and groups that did not follow a straight line.

Ebiaredoh-Mienye et al. [48] devised an effective method for diagnosing CKD by integrating an information-gain-dependent feature selection technique with a cost-sensitive AdaBoost classifier methodology. It was compared to well-known classifiers and a new model that was suggested as a way to identify chronic kidney disease. The smaller set of features was then used to make the suggested cost-sensitive AdaBoost, which was better at classifying than the other classifiers. The test results also show that The feature selection had a positive impact on the performance of the various classifiers.

We can draw conclusions about the analysis of the research works of the previous researchers that many efforts are already being performed to use ML to guess how chronic kidney disease will progress. The amount of datasets, their quality, and the time at which they are collected are also important factors in improving the performance of the models. Also, most of the studies that have been done so far have been about evaluating person strategies or small groups of strategies to diagnose and treat chronic renal disease. Accordingly, chronic renal disease is to be identified and treated properly with a more comprehensive and comparative scientific approach.

## MATERIAL AND METHODS

### Proposed Methodology

In our study where we are dealing with the classification methodology (presented in the fig. 1 and fig. 2) of severe kidney ailment, we concurrently employ six common Machine learning methods. These ML techniques are The following algorithms are utilized: RF, logistic regression, Naive Bayes, KNN, SVM, and feed forward neural network. These ML-based methods were made and are now being used to sort patients into groups based on their serious kidney conditions, especially to see if they have chronic kidney disease. The ML processes involved in the current research have been selected based on the popularity in predicting severe renal illness and the classification results that have been achieved in the already conducted studies [25, 49, 50]. Moreover, our methodological solution, which concurrently uses six various machine learning methods, is more holistic and comparative in nature and has the capability to conduct the systematic evaluation and navigation of the multifaceted physiological indicators of chronic kidney disease. This proves particularly helpful in determining whether the patient is affected by the chronic kidney disease or not as we pay special attention to the aspect of the clinical applicability as opposed to the algorithmic optimization (which had been chiefly prioritized by the corresponding literatures).

### Dataset Briefing

A dataset of people with CKD was used in the study. It was found in a ML library. of UCI [51] that was obtained in a healthcare facility and donated in 2015. The dataset in consideration contained four hundred specimens. In each specimen of the discussed dataset, there are twenty-four predictive factors or traits (13 of which are category-specific parameters and 11 are numeric-typed parameters), along with a categoric-specific reaction parameter (category). Each of the categories has 2 types of values: one type of data is the patient having There is data on patients who do not have chronic kidney disease (150 out of 400 specimens) and data on patients who do have chronic kidney disease (250 out of 400 specimens).

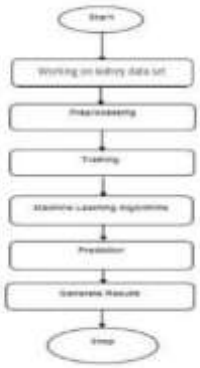

Fig. 1. Flow representation of our improved methodology for diagnosis of CKD using ML techniques.

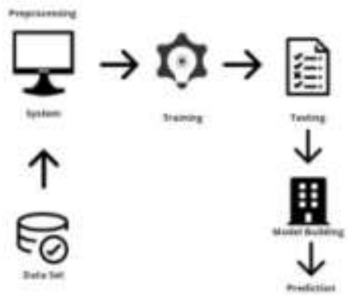

Fig. 2. Architectural representation of our improved methodology for diagnosis of CKD using ML techniques.

**Data Preparation for model application**

Between the category parameters, the missing values of the obtained source chronic renal disease dataset were appropriately managed and initially filled-in. KNN imputation that identified This study looked at the K full instances with the smallest Euclidean distance between them and the instances with missing values. The most prevalent class of the associated parameter has been utilized in K complete cases, to fill the values missing in the categorized parameters. Conversely, the values missing in numeric-typed parameters have been filled in by the median value of the respective parameter in K complete cases. The K-Nearest Neighbour technique was used to fill the gaps in the acquired source chronic renal disease dataset as individuals with similar medical histories are expected to share similar physiological data. As an example, there are certain physiological data that have to be constant in a given range in healthy persons. On the same note, these sort of interventions to the ailing are supposed to be comparable to those of individuals with ailments of the same nature. Specifically, variations in data of physiological observations cannot be significant when comparing individuals in a similar environment.

**Applied ML techniques for prediction of Kidney Ailments**

The following is a list of the six ML techniques that we used in our better way of diagnosing chronic kidney disease.

**Random Forest**

RF is another ML technique which is an extension of bagging. As opposed to exhaustively using its features to produce trees, it introduces an extra step of randomly selecting traits. Preferably, the random forest should be used when many randomized observations are to be dealt with. Suppose that the collection to be trained has M attributes and N occurrences. First, training set is selected to a randomly chosen sample with appropriate replacement. M attributes are randomly selected and the node is partitioned divided recursively by the attribute with the best partition. This approach will make sure the data is right and make it easy to fill in any gaps in the numbers. When used with ML, it can be used to solve problems with regression and classification. In the RF idea, different classification methods are put together to solve a difficult problem and make the models work better. This is called collaborative learning.

**Logistic Regression**

One of the most popular ML techniques is logistic regression, which is part of the field of ML. It has been applied in projecting the category-dependent parameter of a specified set of unique factors. It will be able to forecast the result of a parameter that is based on a category. The resultant decision must, consequently, be unique or category-specific. Such outcomes could be true or false, 0 or 1, Yes or No etc. However, it does not provide the specific numbers such as zero and one but rather provides the probability values in terms of zero and one. The optimum factors to classify can be found in this way as well as the optimum classification of measures depending on the types of information.

**Naive Bayes**

NB is a simple algorithm which makes use of certain probabilistic conditions. The algorithm applied in this kind of approach is the probabilistic tabulation which is updated by the training data. To make a prediction of a new measurement, it is required to research the possibilities of categories, and the tabulation that is based on probabilities is constructed on specific figures. In practice, there is no likelihood that any input characteristic is not dependent on other considerations. The reasons as to why it was selected in the classification endeavors are as follows: it is simple to apply, demonstrates high results, needs minimal training data, scales linearly with the number of information pointers and predictors, solves multiple-class as well as binary classification tasks and makes probable forecasts, and is resistant to irrelevant characteristics.

**K-Nearest Neighbour**

KNN uses a repository containing information points that are organized into categories and performs classification efforts. The strategy is meant to categorize every single bit of information which gets into it as a problem of classification. It doesn't make any assumptions about how the underlying data is spread out, which is why it's called non-parametric. The principal reasons why its choice was made in the classification efforts are the following: it is a simple and low-cost implementation strategy; its structure is inexpensive to construct; it is a highly flexible categorization method ideal in complex categories and it can record a variety of labels on a wider category group.

**Support Vector Machine**

Machine help Tasks such as classification and regression may be done with the help of the ML. This technique requires that a hyperplane form the basis required to define the boundary of the decision. When there is a need to divide any group of objects of different categories, a decision plane is utilized. Complex mathematical equations known as kernels are needed to segregate objects of various categories should the items not be linearly segregated. It tries to classify items using sampling of training dataset. The key reasons why it was chosen in the classification efforts are that it can deal with both semi-arrange and organized data and it can deal with complex functions provided that the kernel can purchase an appropriate function.

**Feed Forward Neural Network**

FFNN are created using computerized neural networks with nodes that do not create looping. This kind of network is also referred to as a multiple-level network because it only passes information in one direction only. The nodes which provide input receive data when data flows through them and then leave through layers of concealment. There are no connections with networks that may be used to send data back to the outcome node. These networks are able to modify their weights during is training depending on a protocol known as the delta protocol which allows the networks to compare their real results with the ones they expected. These networks are made up of artificial neurons that are ultimately created out of real neurons. It is composed of two major components which calculate mathematical operations.

**Implementation Overview**

As a part of the implementation phase of our improved strategy to detect chronic renal disease with the use of ML techniques, we have come up with modules of two different entities, the user and the system. We have made four modules in the serving system and nine in the serving user.

**Our Novelty**

Overall, the novelty of this work lies in simultaneous use of six different machine learning techniques thus allowing the usage of a non-isolated, methodological approach; the greatest focus on clinical applicability, which ensures less complex medical recognitions; and the development of an all-encompassing cum comparative machine learning-based solution, which uncovers the most favorable methodology and, consequently, previously unknown patterns in kidney health data.

**Advantages**

The possible advantages of implementing our improved methodology in the diagnosis of CKD based on ML are as follows:

Early Detection: ML systems can look at a lot of patient data, like test results, medical histories, and imaging data, to find patterns and things that put people at risk for getting chronic kidney disease. This can be used to find the sickness early on. age and thus the medical practitioners can intervene and initiate treatment when the disease is easier to manage.

Personalized Treatment Plans: People who have chronic renal disease can get personalized care plans with the help of ML. ML could improve the health of patients. by suggesting personalized medicine and interventions which have a higher likelihood to succeed in each given person, as it evaluates patient-specific information including genetics, lifestyle, and reaction to treatment.

Predictive Analytics: According to the previous statistics, machine learning algorithms can estimate the way a disease will progress and the way a patient would perform. This predictive capacity can be used to help medical professionals make informed choices on treatment options and time by reducing complications and improving the quality of care received by patients with chronic renal disease.

Monitoring and Alerts: ML-based monitoring systems can be used to constantly monitor the health indicators of a patient and indicate any abnormalities. This real-time monitoring can alert the medical professionals of potential issues by allowing the early intervention and reducing hospitalization rates of the complications related to the chronic renal disease.

Data-driven Research: ML can help to speed up the study of chronic kidney disease through the analysis of very large volumes of data and the identification of new associations and patterns. Researchers can also apply machine learning to discover new biomarkers, risk factors, and treatment methods that will enhance our understanding of the disease and potentially lead to new breakthroughs in the treatment of the disease**.**

## RESULT AND DISCUSSIONS

This part shows the findings of every model of our improved approach to the use of ML methods to diagnose chronic kidney disease.

### Model-wise Performance Evaluation

In this scenario, each model's confusion matrix and ROC curve are utilized to evaluate how well a few classifiers perform. Random forest, logistic regression, naive bayes, KNN, SVM, and feed forward neural network are some examples.

**Random Forest**

Figure 3 shows the confusion matrix of the RF classifier.

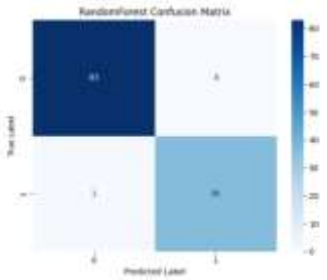


Fig. 3. “Confusion Matrix for Random Forest classifier.”

The graph showing the ROC curve for the Random Forest classifier is signified in the below fig. 4.

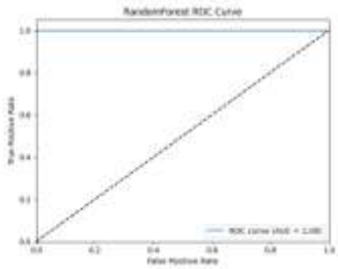


Fig. 4. “ROC Curve for Random Forest classifier.”

**Logistic Regression**

The confusion matrix for the Logistic Regression classifier is depicted in the below fig. 5.

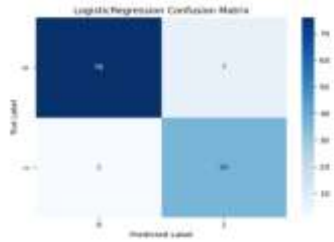


Fig. 5. “Confusion Matrix for Logistic Regression classifier.”

The graph showing the ROC curve for the Logistic Regression classifier is indicated in the below fig. 6.

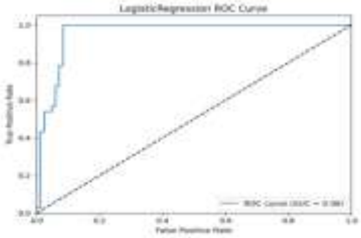


Fig. 6. “ROC Curve for Logistic Regression classifier.”

**Naive Bayes**

Figure 7 below shows the Naive Bayes classifier's uncertainty matrix.

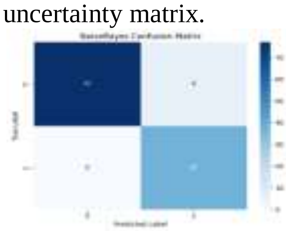


Fig. 7. “Confusion Matrix for Naive Bayes classifier.”

Figure 8 depicts the ROC curve for the Naive Bayes classifier.

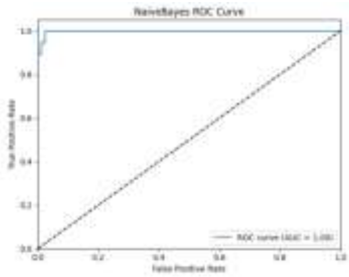


Fig. 8. “ROC Curve for Naive Bayes classifier.”

**K-Nearest Neighbour**

Figure 9 shows the confusion matrix for the KNN method.

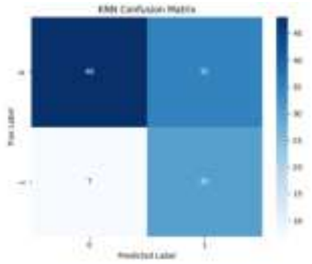

Fig. 9. Confusion Matrix for K-Nearest Neighbour classifier.

The ROC curve for the KNN classifier is shown in the picture below (fig. 10).

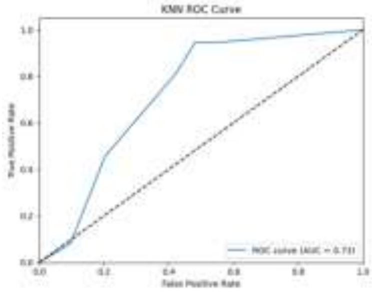

Fig. 10. ROC Curve for K-Nearest Neighbour classifier.

**Support Vector Machine**

Figure 11 shows the SVM classifier's uncertainty matrix.

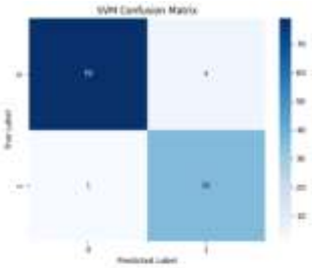

Fig. 11. Confusion Matrix for support vector machine classifier.

The ROC curve for the SVM classifier can be seen in the picture below (fig. 12).

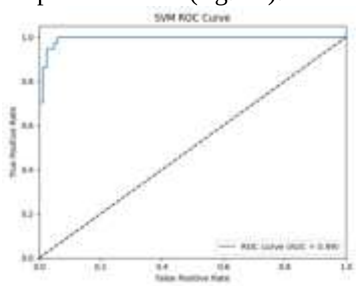

Fig. 12. ROC Curve for SVM classifier.

**Feed Forward Neural Network**

Figure 13 depicts a confusion matrix for the neural network algorithm.

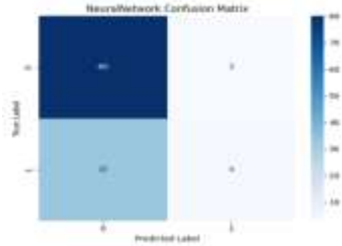

Fig. 13. “Confusion Matrix for Neural Network classifier.”

The ROC curve for the neural network classifier is shown in the picture below (fig. 14).

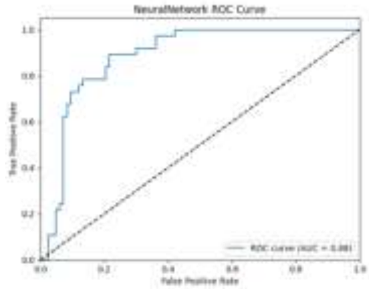

Fig. 14. “ROC Curve for Neural Network classifier.”

## Comparison of all Models

Here, we have tabulated the values of performance metrics like accuracy and ROC AUC for the simultaneously implemented six machine learning approaches in the below table I.

TABLE I. COMPARATIVE INVESTIGATION OF MODELS

| Model | Accuracy | ROC AUC |
|---|---|---|
| Random Forest | **0.9917** | **1.0000** |
| SVM | 0.9583 | 0.9932 |
| Naïve Bayes | 0.9500 | 0.9980 |
| Logistic Regression | 0.9333 | 0.9577 |
| KNN | 0.6500 | 0.7260 |
| Neural Network | 0.6167 | 0.4347 |

Table I above reflects that random forest technique was superior to the other ML methods that were being used simultaneously. Further, the performance metrics-based comparison has been indicated in figures 15 and 16 below.

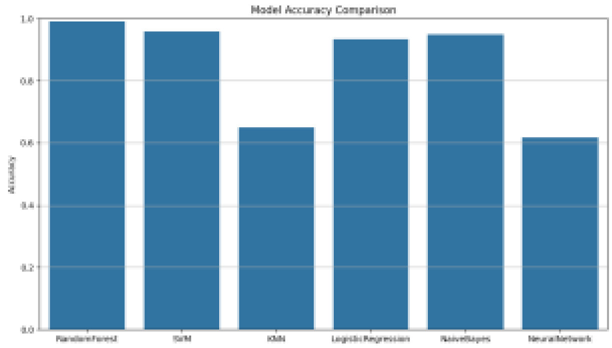


Fig. 15. Model-Wise Accuracy Comparison Chart.

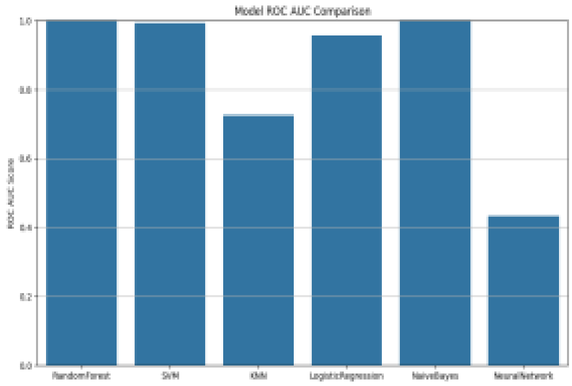


Fig. 16. Model-Wise ROC AUC Comparison Chart.

## Comparison Chart of Detection Results

The below chart displayed in the below fig. 17 depicts the comparison chart of the chronic kidney disease detection result of the patients.

Fig. 17. Comparison Chart of detection result of chronic kidney disease.

## Conclusion

Six well-known ML algorithms were used in this study: RF, LR, naive bayes, KNN, SVM, and feed forward neural network. These ML methods were created to sort people with serious kidney diseases into groups based on whether they have chronic kidney disease or not. The parts of our plan that dealt with material analysis and data imputation worked out well. The models were accurate enough after KNN interpolation was used to fill in the missing values in the UCI dataset without any help from a person. The method found trends in kidney health data that had not been seen before. Because of this, we think that using it to accurately identify chronic kidney disease will have a good effect on it.

## References

[1] Z. Chen, Z. Zhang, R. Zhu, Y. Xiang, P. B. J. C. Harrington, and I. L. Systems, "Diagnosis of patients with chronic kidney disease by using two fuzzy classifiers," vol. 153, pp. 140-145, 2016.

[2] A. Subasi, E. Alickovic, and J. Kevric, "Diagnosis of chronic kidney disease by using random forest," in CMBEBIH 2017: Proceedings of the International Conference on Medical and Biological Engineering 2017, 2017, pp. 589-594: Springer.

[3] L. Zhang et al., "Prevalence of chronic kidney disease in China: a cross-sectional survey," vol. 379, no. 9818, pp. 815-822, 2012.

[4] A. Singh, G. Nadkarni, O. Gottesman, S. B. Ellis, E. P. Bottinger, and J. V. J. J. o. b. i. Guttag, "Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration," vol. 53, pp. 220-228, 2015.

[5] A. M. Cueto-Manzano, L. Cortés-Sanabria, H. R. Martínez-Ramírez, E. Rojas-Campos, B. Gómez-Navarro, and M. J. A. o. m. r. Castillero-Manzano, "Prevalence of chronic kidney disease in an adult population," vol. 45, no. 6, pp. 507-513, 2014.

[6] H. Polat, H. Danaei Mehr, and A. J. J. o. m. s. Cetin, "Diagnosis of chronic kidney disease based on support vector machine by feature selection methods," vol. 41, pp. 1-11, 2017.

[7] C. Barbieri et al., "A new machine learning approach for predicting the response to anemia treatment in a large cohort of End Stage Renal Disease patients undergoing dialysis," vol. 61, pp. 56-61, 2015.

[8] V. Papademetriou et al., "Chronic kidney disease, basal insulin glargine, and health outcomes in people with dysglycemia: the ORIGIN Study," vol. 130, no. 12, pp. 1465. e27-1465. e39, 2017.

[9] N. R. Hill et al., "Global prevalence of chronic kidney disease–a systematic review and meta-analysis," vol. 11, no. 7, p. e0158765, 2016.

[10] M. M. Hossain et al., "Mechanical anisotropy assessment in kidney cortex using ARFI peak displacement: Preclinical validation and pilot in vivo clinical results in kidney allografts," vol. 66, no. 3, pp. 551-562, 2018.

[11] M. Alloghani, D. Al-Jumeily, T. Baker, A. Hussain, J. Mustafina, and A. J. Aljaaf, "Applications of machine learning techniques for software engineering learning and early prediction of students' performance," in Soft Computing in Data Science: 4th International Conference, SCDS 2018, Bangkok, Thailand, August 15-16, 2018, Proceedings 4, 2019, pp. 246-258: Springer.

[12] D. Gupta, S. Khare, and A. Aggarwal, "A method to predict diagnostic codes for chronic diseases using machine learning techniques," in 2016 International Conference on Computing, Communication and Automation (ICCCA), 2016, pp. 281-287: IEEE.

[13] L. Du, C. Xia, Z. Deng, G. Lu, S. Xia, and J. J. I. j. o. m. i. Ma, "A machine learning based approach to identify protected health information in Chinese clinical text," vol. 116, pp. 24-32, 2018.

[14] R. Abbas, A. J. Hussain, D. Al-Jumeily, T. Baker, and A. Khattak, "Classification of foetal distress and hypoxia using machine learning approaches," in Intelligent Computing Methodologies: 14th International Conference, ICIC 2018, Wuhan, China, August 15-18, 2018, Proceedings, Part III 14, 2018, pp. 767-776: Springer.

[15] M. Mahyoub, M. Randles, T. Baker, and P. Yang, "Comparison analysis of machine learning algorithms to rank alzheimer's disease risk factors by importance," in 2018 11th International Conference on Developments in eSystems Engineering (DeSE), 2018, pp. 1-11: IEEE.

[16] Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, and H. J. F. i. g. Tang, "Predicting diabetes mellitus with machine learning techniques," vol. 9, p. 515, 2018.

[17] Z. Gao, J. Li, J. Guo, Y. Chen, Z. Yi, and J. J. I. A. Zhong, "Diagnosis of diabetic retinopathy using deep neural networks," vol. 7, pp. 3360-3370, 2018.

[18] E. Alickovic and A. J. J. o. m. s. Subasi, "Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier," vol. 40, no. 4, p. 108, 2016.

[19] Z. Masetic, A. J. C. m. Subasi, and p. i. biomedicine, "Congestive heart failure detection using random forest classifier," vol. 130, pp. 54-64, 2016.

[20] M. Patrício et al., "Using Resistin, glucose, age and BMI to predict the presence of breast cancer," vol. 18, pp. 1-8, 2018.

[21] R. J. Kate, R. M. Perez, D. Mazumdar, K. S. Pasupathy, V. J. B. m. i. Nilakantan, and d. making, "Prediction and detection models for acute kidney injury in hospitalized older adults," vol. 16, pp. 1-11, 2016.

[22] N. Park et al., "Predicting acute kidney injury in cancer patients using heterogeneous and irregular data," vol. 13, no. 7, p. e0199839, 2018.

[23] X. Wang, Z. Wang, J. Weng, C. Wen, H. Chen, and X. J. I. a. Wang, "A new effective machine learning framework for sepsis diagnosis," vol. 6, pp. 48300-48310, 2018.

[24] Y. Chen et al., "Machine-learning-based classification of real-time tissue elastography for hepatic fibrosis in patients with chronic hepatitis B," vol. 89, pp. 18-23, 2017.

[25] D. A. Debal and T. M. J. J. o. B. D. Sitote, "Chronic kidney disease prediction using machine learning techniques," vol. 9, no. 1, p. 109, 2022.

[26] R. C. Das, M. C. Das, M. A. Hossain, M. A. Rahman, M. H. Hossen, and R. Hasan, "Heart disease detection using ML," in 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), 2023, pp. 0983-0987: IEEE.

[27] S.-J. Sammut et al., "Multi-omic machine learning predictor of breast cancer therapy response," vol. 601, no. 7894, pp. 623-629, 2022.

[28] R. S. Chaulagain, F. T. Liza, S. Chunduri, X. Yuan, and M. J. A. I. S. t. p. Lang, "Achieving the performance of global adaptive routing using local information on dragonfly through deep learning," 2020.

[29] Z. Yao et al., "Machine learning for a sustainable energy future," vol. 8, no. 3, pp. 202-215, 2023.

[30] M. Khan, R. C. Das, J. Casey, B. L. Reese, B. Akintunde, and A. K. J. A. A. Pathak, "Near room temperature magnetocaloric properties in Ni deficient (Mn0. 525Fe0. 5) Ni0. 975Si0. 95Al0. 05," vol. 12, no. 3, 2022.

[31] M. K. Hossain et al., "Combined DFT, SCAPS-1D, and wxAMPS frameworks for design optimization of efficient Cs 2 BiAgI 6-based perovskite solar cells with different charge transport layers," vol. 12, no. 54, pp. 35002-35025, 2022.

[32] M. K. Hossain et al., "A review of applications, prospects, and challenges of proton-conducting zirconates in electrochemical hydrogen devices," vol. 12, no. 20, p. 3581, 2022.

[33] M. K. Hossain et al., "Design and simulation of Cs2BiAgI6 double perovskite solar cells with different electron transport layers for efficiency enhancement," vol. 37, no. 5, pp. 3957-3979, 2023.

[34] P. L. Donti, J. Z. J. A. R. o. E. Kolter, and Resources, "Machine learning for sustainable energy systems," vol. 46, no. 1, pp. 719-747, 2021.

[35] M. G. Bastos and G. M. J. B. J. o. N. Kirsztajn, "Chronic kidney disease: importance of early diagnosis, immediate referral and structured interdisciplinary approach to improve outcomes in patients not yet on dialysis," vol. 33, pp. 93-108, 2011.

[36] A. Farjana et al., "Predicting chronic kidney disease using machine learning algorithms," in 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), 2023, pp. 1267-1271: IEEE.

[37] J. W. Stanifer et al., "The epidemiology of chronic kidney disease in sub-Saharan Africa: a systematic review and meta-analysis," vol. 2, no. 3, pp. e174-e181, 2014.

[38] S. Abd ElHafeez, D. Bolignano, G. D'Arrigo, E. Dounousi, G. Tripepi, and C. J. B. o. Zoccali,

"Prevalence and burden of chronic kidney disease among the general population and high-risk groups in Africa: a systematic review," vol. 8, no. 1, p. e015069, 2018.

[39] M. D. Molla et al., "Assessment of serum electrolytes and kidney function test for screening of chronic kidney disease among Ethiopian Public Health Institute staff members, Addis Ababa, Ethiopia," vol. 21, pp. 1-11, 2020.

[40] A. Agrawal, H. Agrawal, S. Mittal, and M. Sharma, "Disease prediction using machine learning," in Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT), 2018, pp. 26-27.

[41] G. Murugesan et al., "Fuzzy logic-based systems for the diagnosis of chronic kidney disease," vol. 2022, p. 2653665, 2022.

[42] E. M. Senan et al., "Diagnosis of chronic kidney disease using effective classification algorithms and recursive feature elimination techniques," vol. 2021, no. 1, p. 1004767, 2021.

[43] S. K. Dey, K. M. M. Uddin, H. M. H. Babu, M. M. Rahman, A. Howlader, and K. A. J. I. S. w. A. Uddin, "Chi2-MI: A hybrid feature selection based machine learning approach in diagnosis of chronic kidney disease," vol. 16, p. 200144, 2022.

[44] M. M. Hassan et al., "A comparative study, prediction and development of chronic kidney disease using machine learning on patients clinical records," vol. 3, no. 2, pp. 92-104, 2023.

[45] Q. Bai, C. Su, W. Tang, and Y. J. S. r. Li, "Machine learning to predict end stage kidney disease in chronic kidney disease," vol. 12, no. 1, p. 8377, 2022.

[46] N. Alturki et al., "Improving Prediction of Chronic Kidney Disease Using KNN Imputed SMOTE Features and TrioNet Model," vol. 139, no. 3, 2024.

[47] S. J. B. M. Pal and Devices, "Chronic kidney disease prediction using machine learning techniques," vol. 1, no. 1, pp. 534-540, 2023.

[48] S. A. Ebiaredoh-Mienye, T. G. Swart, E. Esenogho, and I. D. J. B. Mienye, "A machine learning method with filter-based feature selection for improved prediction of chronic kidney disease," vol. 9, no. 8, p. 350, 2022.

[49] K. J. C. k. d. p. b. o. n. B. t. Priyanka, "Science bc," pp. 1653-9, 2019.

[50] S. Kapoor, R. Verma, and S. Panda, "Detecting kidney disease using Na ve bayes and decision tree in machine learning," 2019.

[51] D. Dua and C. J. U. h. a. i. u. e. m. Graff, "UCI machine learning repository, 2017," vol. 7, no. 1, p. 62, 2017.

1.