

# Chronic Kidney Disease Identification in HIV Infected Patients using Machine Learning.

Review 2

Date : 23/12/22

1CR19CS078

KAVYA SONI

1CR19CS019

ANAMIKA MAJEE

1CR19CS023

ARPITA D

**Under the guidance of:**

Name: Mrs. Ashwathi

Designation: Assistant professor

# Abstract



- ▶ The Chronic Kidney Disease (CKD) is one of the deadliest diseases that slowly damages human kidney. The disease remains undetected in its early stage and the patients can only realise the severity of the disease when it gets advanced people infected with HIV often suffer from CKD.
- ▶ In our proposed system we compare and analyse different Machine Learning Algorithm such as Naive Bayes, KNN, Decision Tree Classifier, SVM, Logical Regression and SGD and to predict CKD for individual human.





- ▶ Introduction
- ▶ Problem Statement
- ▶ Objectives
- ▶ Literature Survey
- ▶ Sources
- ▶ Comparison
- ▶ Conclusion
- ▶ References

# Introduction



- ▶ There are some symptoms which shows kidneys are beginning to fail like muscle cramps, nausea and weight loss, swelling in your feet and ankles, too much urine or not enough urine, shortness of breath, trouble sleeping, fever and vomiting.
- ▶ Risk factors of CKD are diabetes, smoking, lack of sleeping, hyper tension, improper diet, etc. Among them diabetes is the more dangerous factor. At the last stage, the patient must take dialysis or do kidney transplantation. One of the best ways to reduce this death rate is early treatment.

# Problem Statement



- ▶ To implement Chronic Kidney Disease Identification in HIV infected patients using Machine Learning Algorithm.

# Objective



- ▶ CKD remains undetected in its early stage and the patients can only realise the severity of the disease when it gets advanced.
- ▶ Hence, detecting such disease at earlier stage is a key challenge now.
- ▶ This project helps to raise awareness among people and to promote early diagnosis.
- ▶ Early prediction and proper treatments can possibly stop, or slow the progression of this chronic disease.



# Literature Review

# Literature Survey : Sources



In order to get required knowledge about various concepts related to the present application, existing literature were studied. Some of the important conclusions were made through those are listed below :

## **1. Prediction of Chronic Kidney Disease using Random Forest Machine Learning Algorithm** **Author : Prof.Manish Kumar**

In this paper, CD is being predicted using a Machine Learning Technique called Random Forest Algorithm. Random Forests are a combination of tree predictors so that all the trees depend on random vector sampled autonomously. Many other algorithms like SVM and Naive Bayes were also used, but Random Forest was the most accurate.



# Literature Survey : Sources



## 2. Chronic Kidney Disease Prediction using Machine Learning Models Author: Prof. S. Revathy

In this paper, CKD prediction was made using Decision Trees, Random Forrest and Support Vector Machine algorithms. All the three algorithms were tested to check the highest accurate algorithm. Random Forrest is the most accurate followed by SVM and decision trees.

## 3. Comparative study of classification Algorithms in Chronic Kidney Disease Author: Prof. Pratibha Devishri S

In this paper, CKD prediction is done using performance measures like Kappa statistics, Receiver Operating Characteristic, Mean Absolute Error and Root mean squared Error using WEKA. Classifiers to be examined are created using Decision stamp, Rep tree, IBK, K-star, SGD and SMO. It is understood that Decision stump and Rep tree are giving better results.

# Literature Survey : Sources

## **4. A survey on Chronic Kidney Disease detection using Novel Methods Author: Prof. K Dharmarajan**

In this paper, CKD is being predicted using multiple Machine Learning Techniques. Some of the techniques are Naive Bayes, SVM, MBPN, LDA, KNN, etc. This survey confirmed that all these techniques successfully predicted with reasonable margin of error.

## **5. Rule Induction and Prediction of Chronic Kidney Disease Using Boosting Classifiers, Ant-Miner and J48 Decision Tree Author: Prof. Arif-Ul-Islam and Shamim H. Ripon**

In this paper, CKD is being predicted using Boosting classifiers like AdaBoost and rule induction techniques like Ant-Miners. Boosting algorithm is an ensemble type machine learning algorithm which converts weak classifiers to strong model to achieve a better accuracy. Researches showed that many data mining techniques had been applied for CKD classification. Among those algorithms, AdaBoost classifier and J48 rule induction method performed well.



# Relevance of Problem

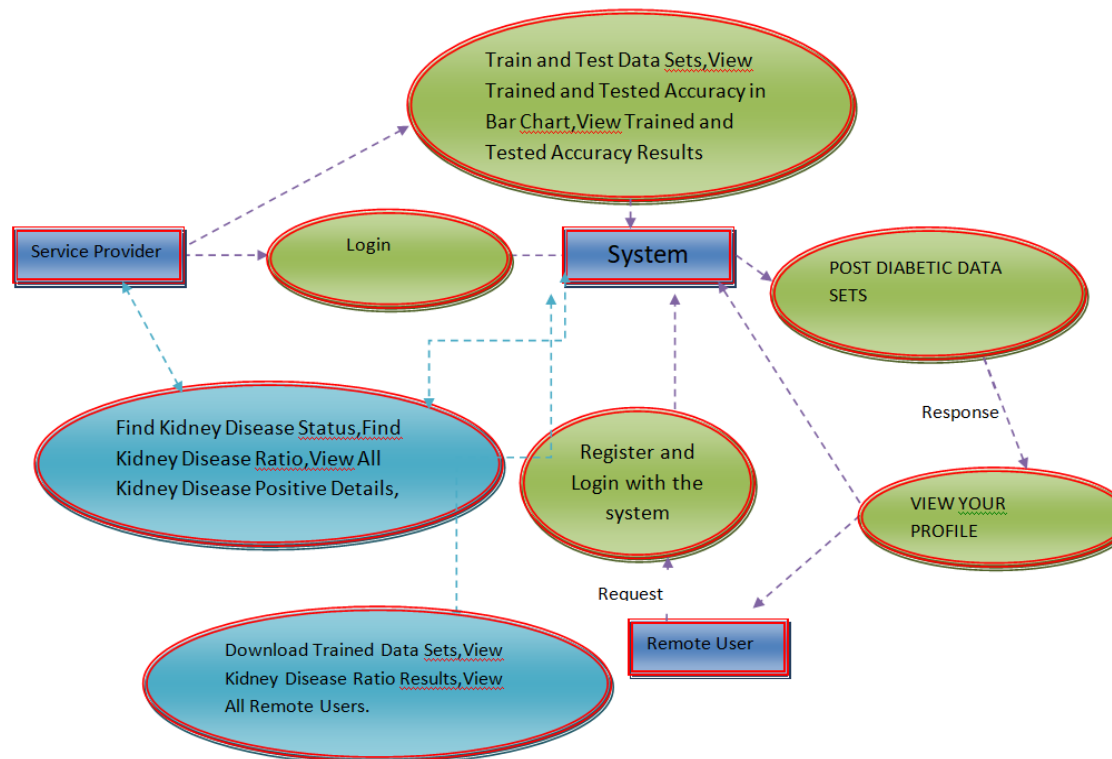


- This would help detect the chances of a person having CKD further on in his life which would be really helpful and cost effective people.
- This model could be integrated with normal blood report generation, which could automatically flag out if there is a person at risk.
- Patients would not have to go to a doctor unless they are flagged by the algorithms.
- This would make it cheaper and easier for the modern busy person.
- CKD ranges from mild to severe all 5 stages of CKD can be predicted.

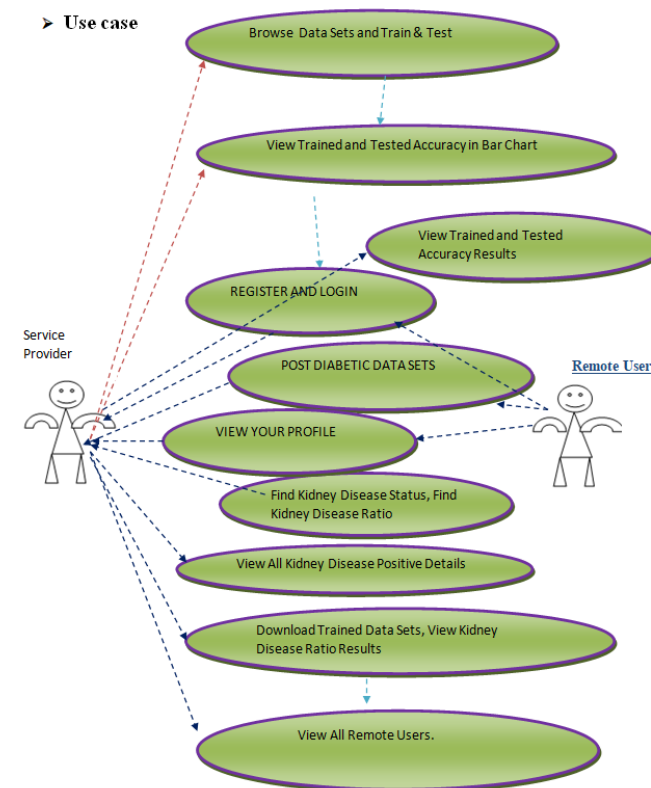
# Block Diagram for Summary



## ➤ Data Flow Diagram :

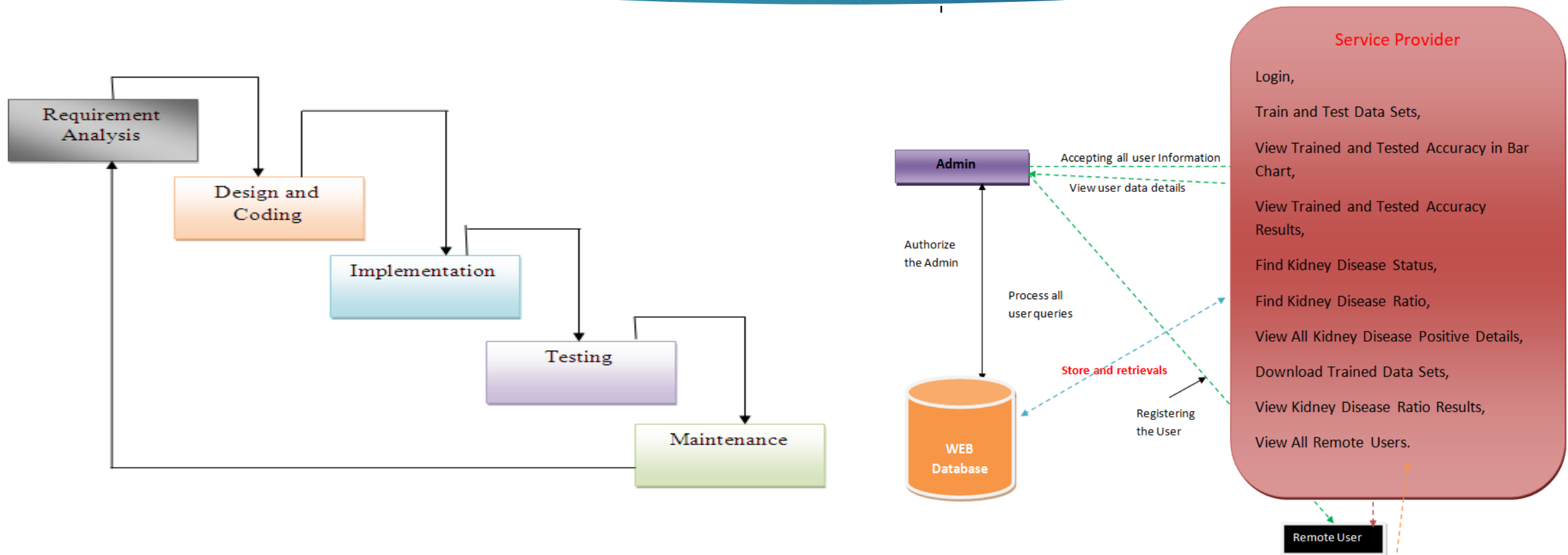


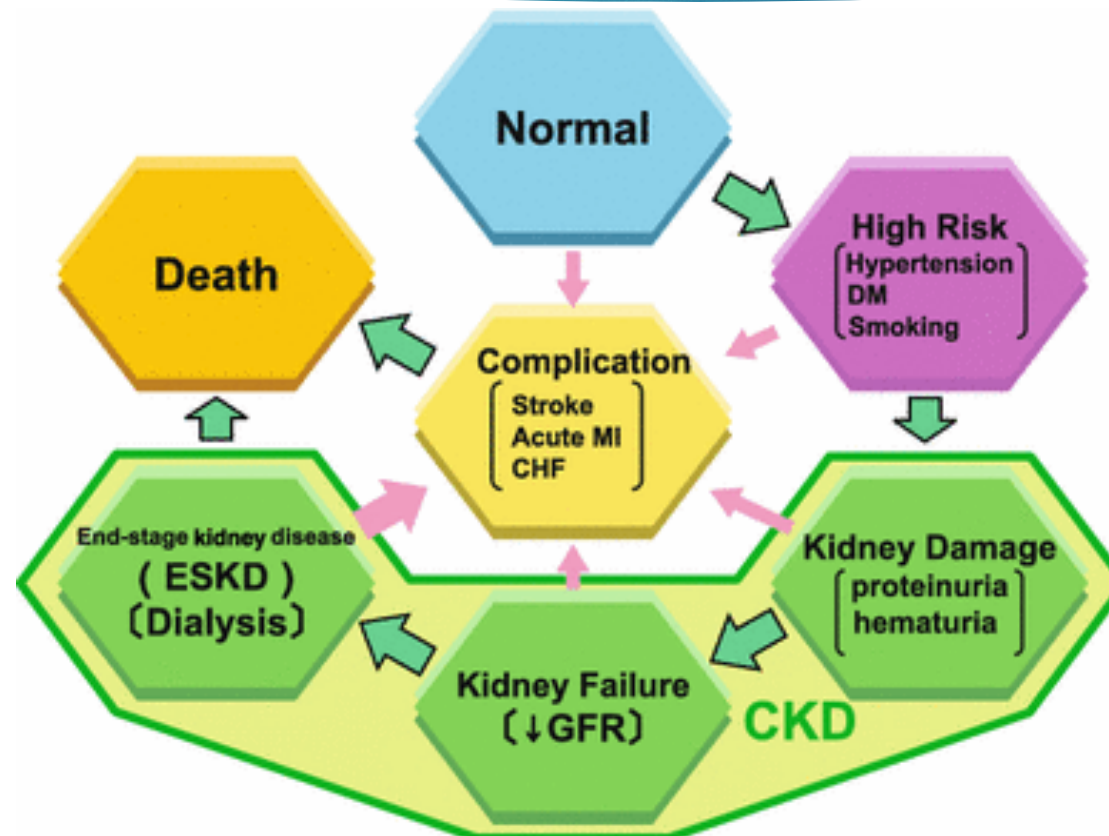
## ➤ Use case





# Block Diagram for Summary







# Research Gap



- ▶ We will enhance the dataset feature space to obtain more accurate results by machine learning techniques and add more folds of datasets.
- ▶ The unwanted presence of missing and outlier values in the training data often reduces the accuracy of a model or leads to a biased model. It leads to inaccurate predictions.
- ▶ An accurate predictive model from different machine learning algorithms, trained on a feature rich datasets comprising of 35 different attributes.
- ▶ Tuning of different classifiers will find the optimum value for each parameter to improve the accuracy of the model.

# Conclusion



- ▶ This system presented the best prediction algorithm to predict CKD at an early stage. The dataset shows input parameters collected from the CKD patients and the models are trained and validated for the given input parameters. K Nearest Neighbours Classifier, Naive Bayes, Logical Regression and SVM, SGD models are constructed to carry out the diagnosis of CKD. The performance of the models is evaluated based on a variety of comparison metrics are being used, namely Accuracy, The results of the research showed that logical regression and SGD model has better predicts CKD in comparison to the other models and also help to predict chronic kidney disease in HIV infected patient.



# References



- M.P.N.M. Wickramasinghe, D.M. Perera and K.A.D.C.P Kahandawaarachchi, "Dietary prediction for patients with Chronic Kidney Disease (CKD) by considering blood potassium level using machine learning algorithms," 2017 IEEE Life Sciences Conference (LSC), Sydney, NSW, 2017, pp. 300-303.
- U.N. Dulhare and M. Avesha, "Extraction of action rules for chronic kidney disease using Naive baves classifier," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, 2016. pp.I-S.
- R. Devika, S.V. Avilala and V. Subramaniaswamy, "Comparative Study of Classifier for Chronic Kidney Disease prediction using Naive Bayes, KNN and Random Forest," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, pp. 679-684.
- G. Kaur and A. Sharma, "Predict chronic kidney disease using data mining algorithms in hadoop." " 2017 International Conference on Inventive Computeine and Informatics (ICICI), Coimbatore, 2017, pp. 973-979.