# A Report

# On

# Proposed Title: Chronic Kindey Disease risk prediction using Machine Learning


**Submitted by:**

**Harsh Kumar Sharma (20205073)**

**Rishav Raj (20205122)**

**Sathpati Deepak (20205142)**


**B.Tech (VII SEM)**

**Session 2023-24**



**Supervised by:**

**DR. SATISH CHANDRA**

**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING**

**MOTILAL NEHRU NATIONAL INSTITUTE OF TECHNOLOGY, ALLAHABAD**

**PRAYAGRAJ-211004, INDIA**

## ABSTRACT:

This study will highlight the importance of combining machine learning with expert knowledge when trying to predict Chronic Kidney Disease (CKD). To develops a web application with user interface to prompt out whether a patient is prone to CKD or not. To necessitate this, the prominent features will be selected using feature selection techniques and classification will be preformed by widely used classifiers. Each feature will be computed by a score using feature Selection methods. Different Classification Algorithms will be used to predict the class labels. The model will be trained with data set and using K Fold Cross validation method the dataset will be partitioned into training and testing data. It will try to achieve its maximum accuracy.

## INTRODUCTION:

Chronic kidney disease (CKD) means your kidneys are damaged and can't filter blood the way they should. The disease is called "chronic" because the damage to your kidneys happens slowly over a long period of time. This damage can cause wastes to build up in your body.

The two main causes of CKD are diabetes and high blood pressure. Diabetes is elevated blood sugar levels, harming the blood vessels, eyes, kidneys, and heart. Ineffective management of high blood pressure can contribute significantly to heart attack, stroke, and chronic renal illness. Because a person may only survive without their kidneys for an average of 18 days, dialysis and kidney transplants are in great demand. It is crucial to possess reliable techniques for CKD early prediction. [1]

Machine learning can be used to predict a positive CKD status and the stages of CKD. When it comes to making predictions based on historical data using classification and regression techniques, machine learning captures a significant portion of artificial intelligence.

## OBJECTIVE:

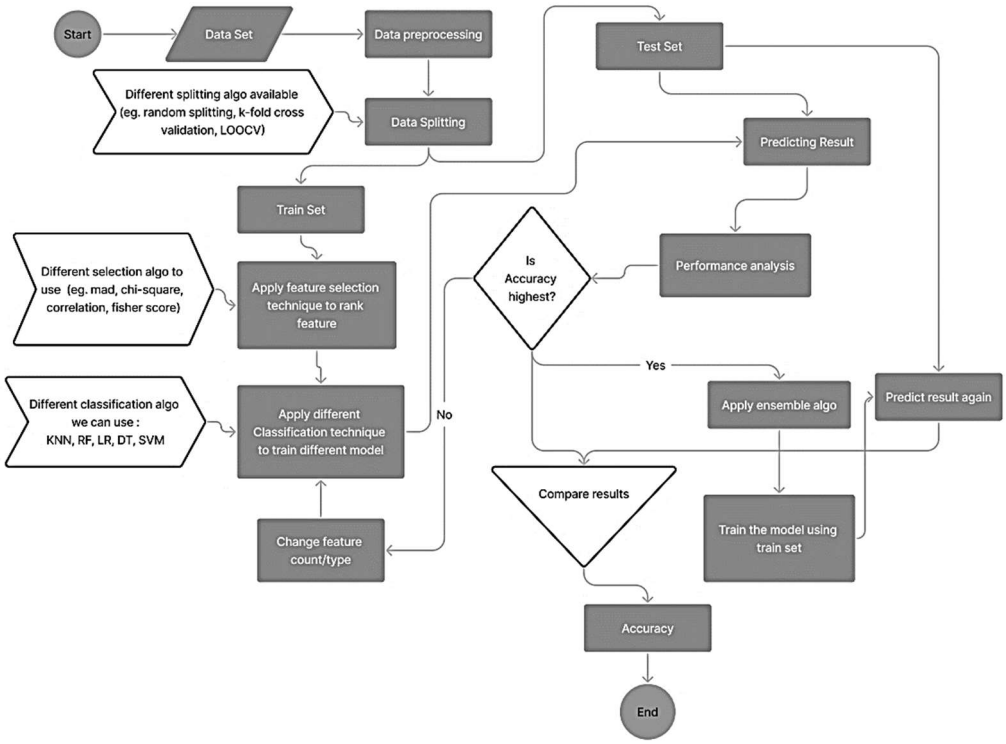
1. To develop model with higher accuracy using the identified features and classifier combination.
2. To develop a web application for diagnosing the risk factor of kidney.

## METHODOLOGY:

The project will use the dataset provided by UCI [2] repository for training and testing the model. It will first pre-process the dataset to remove redundant and faulty data, then split the dataset to train and test data using data dividing algorithm. Then train data will be used to do feature selection by using different algorithm and comparing the result's accuracy, to rank the features according to it. This will help in finding the best set of features and classifier. It will then use ensemble algorithm to combine different classifier to increase the accuracy of final model.

The project will be done in Jupyter Notebook with the help of python as the programming language and sklearn as the library for using different Machine Learning algorithms. The web page will be made using React.

FLOW DIAGRAM FOR PROPOSED WORK :




TIMELINE:

| | | |
|---|---|---|
| **Literature Survey** | 16 Aug-30 Sep | Doing literature Survey to analyze the preexisting model and come up with a better model |
| **Learning of AI/ML** | 16 Aug-16 Oct | To learn about different Machine Learning algorithms and integrating in with a programming language. |
| **Datasets** | 17 Oct-20 Oct | To collect the real time data from the UCI repository of patients and do data pre-processing. |
| **Feature Selection** | 21 Oct- 21 Nov | To determine the parameters that affect the trend of chronic kidney disease using the rank method |
| **Optimal Classification** | 21 Oct- 21 Nov | To train the model using different classification methods (KNN, Random Forest, decision tree) and determine the best among them. |
| **Final model and user interface** | 21 Oct- 21 Nov | To determine the evaluation functions (accuracy, precision, and recall) of the model using classification and feature selection and further using ensemble model to predict the result. Finally, creating a web page for user to interact with. |
| **Preparation of Project Report** | 21 Nov- 30 Nov | To do a through documentation on the findings of the project. |

REFERENCES:


1. **Dataset: Centers for Disease Control and Prevention. Chronic Kidney Disease in the United States, 2019. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2019.**
2. **Kelly M., Longjohn R., Nottingham K., The UCI Machine Learning Repository, https://archive.ics.uci.edu**
3. **Haratian, A., Maleki, Z., Shayegh, F. et al. Detection of factors affecting kidney function using machine learning methods. Sci Rep 12, 21740 (2022).**
4. **Bai, Q., Su, C., Tang, W. et al. Machine learning to predict end stage kidney disease in chronic kidney disease. Sci Rep 12, 8377 (2022).**
5. **Elhoseny, M., Shankar, K. & Uthayakumar, J. Intelligent Diagnostic Prediction and Classification System for Chronic Kidney Disease. Sci Rep 9, 9583 (2019).**