

CHRONIC KIDNEY DISEASE PREDICTION USING MACHINE LEARNING

BACKGROUND STUDY

The prevalence of CKD is on the rise worldwide. In the United States, an estimated 37 million people have CKD, with the majority of cases being undiagnosed. CKD is also a leading cause of death, particularly in older adults. There are several risk factors for CKD, including diabetes, high blood pressure, obesity, and smoking. Despite efforts to improve early detection and management of CKD, many people with the disease do not receive appropriate care, and there is a need for better methods of identifying and managing CKD.

ABSTRACT

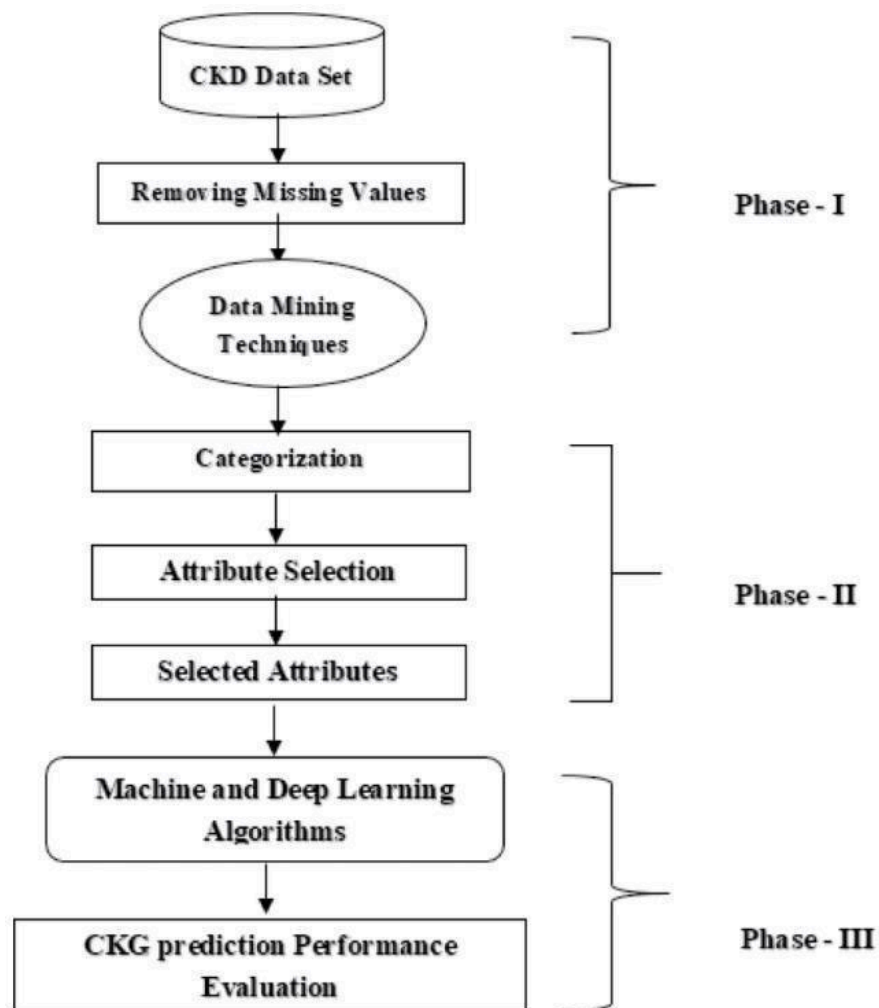
Chronic Kidney Disease (CKD) or chronic renal disease has become a major issue with a steady growth rate. A person can only survive without kidneys for an average time of 18 days, which makes a huge demand for a kidney transplant and Dialysis. It is important to have effective methods for early prediction of CKD. Machine learning methods are effective in CKD prediction. This work proposes a workflow to predict CKD status based on clinical data, incorporating data preprocessing, a missing value handling method with collaborative filtering and attributes selection. Out of the 11 machine learning methods considered, the extra tree classifier and random forest classifier are shown to result in the highest accuracy and minimal bias to the attributes. The research also considers the practical aspects of data collection and highlights the importance of incorporating domain knowledge when using machine learning for CKD status prediction.

OBJECTIVE OF STUDY

The primary objective of this study is to develop and evaluate machine learning models for predicting the risk of CKD based on patient data. Specifically, we aim to:

- Develop models that can accurately predict the presence of CKD based on demographic clinical, and laboratory data.
- Evaluate the performance of different feature extraction techniques for identifying the most important predictors of CKD.
- Compare the performance of different machine learning algorithms for predicting CKD risk.
- Investigate the impact of sample size and data imbalance on model performance.
- Provide insights into the most important factors associated with CKD risk and identify opportunities for improving CKD diagnosis and management.

METHODOLOGY BLOCK DIAGRAM OF CKD



Submitted By:

20MH1A4210 - Eethakota Ayyappa Karthik
20MH1A4239 - Mutha Surya Bhavani
20MH1A4252 - Singineedi Charan Sagar Aswanth
20MH1A4263 - Vikas Sai Kumar Balla

Under The Guidance Of
Mr. K N S K Santhosh. M.Tech(PhD)
Associate Professor
Department of CSE(AI & ML/IOT)
Aditya College Of Engineering