



INFORMATICS
INSTITUTE OF
TECHNOLOGY

UNIVERSITY OF
WESTMINSTER

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER, UK

6 COSC023C.Y - Final Year Project

“CKD Predictor”

Predicting Chronic Kidney Disease using Machine Learning

A dissertation by

Navin Rerukaana

(IIT ID : 2018604 | UoW ID : W1742854)

Supervised by

Miss Gayashini Ratnayake

Submitted in partial fulfillment of the requirements for the

BSc(Hons) Business Information Systems degree

at the University of Westminster.

School of Computing and Engineering University of Westminster

Abstract

Chronic kidney disease (CKD) is a long-term disease that causes the kidneys to not work as they should. This disease can affect anyone, but it is more common for people who are older and for people who are from South Asia. Kidney disease affects 10% of the human population. CKD slowly and progressively loses the functions of the kidney over time, and eventually the person who has CKD may develop kidney failure.

The main reason for implementing a system like this is to prevent CKD patients from getting into End Stage Renal Disease (ESRD), which means the person's kidneys stop working permanently, needing either a kidney transplant or a regular course of long-term treatment to stay alive. The main reasons for a patient getting ESRD is by failing to detect the disease at an early stage, not following proper diet plans even after identifying the disease, getting advice from inexperienced doctors, and not having proper follow up care.

This system will help patients regularly get CKD predictions while they are at home, receive diet recommendations based on their level of disease to control or prevent getting the disease, and recommend specialized doctors (Neurologist) to get guidance from. With the use of machine learning techniques with the highest accuracy, this system will help people identify the disease before or at an earlier stage, which will help the patient prevent or slow down the progression of CKD.

When implementing the machine learning model, 10 different models were created to pick the model with the most accuracy. A dataset with more than 400 records was used when implementing the model. To test the model, 40% of the dataset was used, and the rest, 60%, was used to train the model. Out of the 10 models created, the author has chosen the Random Forest classifier as the prediction model with an accuracy of 99.38%.

The primary goal of this research is to create a system that will assist patients in detecting CKD in its early stages, receiving the appropriate treatment to manage it without developing end-stage renal disease (ESRD), and living a healthy life.

Keywords - Chronic Kidney Disease, End Stage Renal Disease, Diet plans, Machine Learning, Random Forest