EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING

TEAM ID:

PNT2022TMID15929

TEAM MEMBERS:

1. RAGHUL V (927619BEC4155)

2. NITHISHKUMAR S (927619BEC4136)

3. RAGHURAJ S (927619BEC4156)

4. ROAHIT S (927619BEC4166)

INDUSTRY MENTORS NAME:

1. SOWJANYA

2. SANDEEP DOODIGANI

FACULTY MENTOR NAME:

1. Mr. S. MEIVEL

ABSTRACT:

Chronic kidney disease (CKD) is among the top 20 causes of death worldwide and affects approximately 10% of the world adult population. CKD is a disorder that disrupts normal kidney function. Due to the increasing number of people with CKD, effective prediction measures for the early diagnosis of CKD are required. The novelty of this study lies in developing the diagnosis system to detect chronic kidney diseases. This study assists experts in exploring preventive measures for CKD through early diagnosis using machine learning techniques. The mean and mode statistical analysis methods were used to replace the missing numerical and the nominal values. To choose the most important features, Recursive Feature Elimination (RFE) was applied. Four classification algorithms applied in this study were support vector machine (SVM), k-nearest neighbors (KNN), decision tree, and random forest. The random forest algorithm outperformed all other applied algorithms, reaching an accuracy, precision, recall, and F1-score of 100% for all measures. CKD is a serious life-threatening disease, with high rates of morbidity and mortality. Therefore, artificial intelligence techniques are of great importance in the early detection of CKD. These techniques are supportive of experts and doctors in early diagnosis to avoid developing kidney failure.

LITERATURE REVIEW:

A recent development of machine learning techniques and data mining has led to an interest in implementing these techniques in various fields. The medical field is not exclusion and the increasing requirements for chronic kidney disease treatments have risks as the data required for treatment plays a major role. That has led to an interest in developing current methods of prediction.

AUTHOR: Kunwar, et al. entitled "Chronic Kidney Disease Analysis is Using Data Mining Classification Techniques" published in 2016.

DESCRIPTION: Data mining is the process of extracting hidden information from massive dataset, categorizing valid and unique patterns in data. There are many data mining techniques like clustering, classification, association analysis, regression etc. The objective of the paper is to predict Chronic Kidney Disease (CKD) using classification techniques like Naive Bayes and Artificial Neural Network (ANN). The experimental results implemented in Rapid Miner tool show that Naive Bayes produce more accurate results than Artificial Neural Network.

AUTHOR: Amirgaliyev, et al. entitled "Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods" published in 2015.

DESCRIPTION: Currently, there are many people in the world suffering from chronic kidney diseases worldwide. Due to the several risk factors like food, environment and living standards many people get diseases suddenly without understanding of their condition. In this research study, the effects of using clinical features to classify patients with chronic kidney disease by using support vector machines algorithm is investigated. The chronic kidney disease data set is based on clinical history, physical examinations, and laboratory tests.

AUTHOR: Devika, et al. entitled "Comparative Study of Classifier for Chronic Kidney Disease Prediction Using Naive Bayes, KNN and Random Forest" published in 2019.

DESCRIPTION: Chronic Kidney disease defines constrains which affects your kidneys and reduces your potential to stay healthy. Machine learning is an important task as it benefits many applications, varied knowledge mining classification approaches and machine learning algorithms are applied for

prediction of chronic diseases. Therefore, this paper examines the performance of Naive Bayes, K-Nearest Neighbor (KNN) and Random Forest classifier on the basis of its accuracy, preciseness and execution time for CKD prediction. Finally, the outcome after conducted research is that the performance of Random Forest classifier is finest than Naive Bayes and KN

AUTHOR: Dulhare, et al. entitled "Extraction of Action Rules for Chronic Kidney Disease using Naive Bayes Classifier" published in 2017.

DESCRIPTION: Chronic kidney disease (CKD), also known as chronic renal disease, which is a progressive loss in kidney function over a period of months or years. It is defined by the presence of kidney damage or decreased glomerular filtration rate (GFR). The estimated prevalence of CKD is about 9-13 % in the general adult population. Chronic Kidney Disease is a silent condition. Signs and symptoms of CKD, if present, are generally not specific in nature and unlike several other chronic diseases (such as congestive heart failure and chronic obstructive lung disease), they do not reveal a clue for diagnosis or severity of the condition. Early detection and treatment can often keep chronic kidney disease from getting worse.

AUTHOR: Aljaaf, et al. entitled "Early Prediction of Chronic Kidney Disease Using Machine Learning Supported by Predictive Analytics" published in 2018.

DESCRIPTION: Chronic Kidney Disease is a serious lifelong condition that induced by either kidney pathology or reduced kidney functions. Early prediction and proper treatments can possibly stop, or slow the progression of this chronic disease to end-stage, where dialysis or kidney transplantation is the only way to save patient's life. In this study, we examine the ability of several machine-learning methods for early prediction of Chronic Kidney Disease.

AUTHOR: Avci E et al. entitled "Performance Comparison of Some Classifiers on Chronic Kidney Disease Data" published in 2018.

DESCRIPTION: In this study, dataset named "Chronic Kidney Disease" obtained from UCI database is used. The dataset consists of 400 individual's information and contains 25 features dataset was classified according to whether it is chronic kidney disease using Naive Bayes (NB), K-Star, Support Vector Machines (SVM) and J48 classifiers used in data mining.