LITERATURE SURVEY ON EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING AND INFORMATION GATHERING

Chronic kidney disease (CKD) is a significant public health problem worldwide, especially for low and medium in come countries. Chronic kidney disease (CKD) means that the kidney does not work as expected and cannot correctly filter blood. About 10% of the population worldwide suffers from (CKD), and millions die each year because they cannot get affordable treatment, with the number increasing in the elderly. According to the Global Burden Disease 2010 study conducted by the International Society of Nephrology, chronic kidney disease (CKD) has been raised as an important cause of mortality worldwide with the number of deaths increasing by 82.3% in the last two decades [1, 2]. Also, the number of patients reaching end-stage renal disease (ESRD) is increasing, which requires kidney transplantation or dialysis to save patients' lives [1, 3, 4]. CKD, in its early stages, has no symptoms; testing may be the only way to find out if the patient has kidney disease.Early detection of CKD in its initial stages can help the patient get effective treatment and then prohibit the progression to ESRD [1].

It is argued that every year, a person that has one of the CKD risk factors, such as a family history of kidney failure, hypertension, or diabetes, get checked. The sooner they know about having this disease, the sooner they can get treatment. To raise awareness and to encourage those who are most susceptible to the disease to perform the tests periodically, we hope that the disease can be detected with the least possible tests and at low cost. So, the objective of this research is to provide an effective model to predict the CKD by least number of predictors. In this paper, Section II reviews various research works that target the diagnosis of CKD using different intelligent techniques. Section III presents the dataset source and description. Section IV presents the methodology used for the prediction, including the data preprocessing steps and the modeling stage. Section V shows the results of the experiment and discusses the performance of ML algorithms in detecting CKD. Finally, Section VI includes the conclusion and future work of this work.

Related Work:

In recent years, few studies have been done on the classification or diagnosis of chronic kidney disease. In 2013, T. Di Noia et al. [5], presented a software tool that used the artificial neural network ANN to classify patient status, which is likely to lead to end-stage renal disease (ESRD). The classifiers were trained using the data collected at the University of Bari over a 38-year period, and the evaluation was done based on precision, recall, and F-measure. The presented software tool has been made available as both an Android mobile application and online web application. Using data from Electronic Health Records (EHR) in 2014, H. S. Chase et al. [6] identified two groups of patients in stage 3: 117 progressor patients (eGFR declined >3ml/min/1.73m2 /year) and 364 non-progressor patients (eGFRdeclined <1 ml/min/1.73m2) .Where GFR is a glomerular filtration rate that commonly used to detect CKD.

Based on initial lab data recorded, the authors used Naïve Bayes and Logistic Regression classifiers to develop a predictive model for progression from stage 3 to stage 4. They compared the

metabolic complications between the two groups and found that phosphate values were significantly higher, but bicarbonate, hemoglobin, calcium, and albumin values were significantly lower in progressors compared to nonprogressors, even if initial eGFR values were similar. Finally, they found that the probability of progression in patients classified as progressors was 81% (73% – 86%) and nonprogressors was 17% (13% – 23%).

REFERENCES(BIBLIOGRAPHY)

[1] J. Radhakrishnan et al, "Taming the chronic kidney disease epidemic: a global view of surveillance efforts," Kidney Int., vol. 86, (2), pp. 246-250, 2014.
[2] R. Lozano et al, "Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010," The Lancet, vol. 380, (9859), pp. 2095-2128, 2012.