Research Article

iMedPub Journals
http://www.imedpub.com

American Journal of Computer Science and Engineering Survey

2020
Vol. 8 No.2:7

DOI: 10.36648/computer-science-engineering-survey.08.02.07.

Early Prediction of Chronic Kidney Disease by using Machine Learning Techniques

Bidri Deepika*, Vasudeva Rao KR, Dharmaj N Rampure, Prajwal P and Devanand Gowda G

Department of Computer Science & Engineering, PES College of Engineering, Mandya, India

***Corresponding author:** Deepika Bidri, Assistant Professor, Department of Computer Science & Engineering, PES College of Engineering, Mandya, India, E-mail: bdeepika@pesce.ac.in

Citation: Deepika B, Rao VKR, Rampure DN, Prajwal P, Gowda DG, et al (2020) Early Prediction of Chronic Kidney Disease by using Machine Learning Techniques. Am J Comput Sci Eng Surv Vol. 8 No. 2: 07.

Abstract

There are many people who are suffering from chronic kidney diseases worldwide. Due to the several risk factors like food, environment and living standards many people get diseases suddenly. Diagnosing of chronic kidney diseases is generally invasive, costly, time-consuming and often risky. That is why many patients reach late stages of it without treatment, especially in those countries where the resources are limited. Therefore, the early detection strategy of the disease remains important, particularly in developing countries, where the diseases are generally diagnosed in later stages. Finding a solution for the above-mentioned problems and riding out from disadvantages became a strong motive to conduct this study. Chronic Kidney Disease (CKD) is one of the types of kidney disease, which results in a gradual loss of kidney function. This phenomenon can be observed over a period of months or years due to several living conditions of patients. The goal is to build a real time application by using the machine learning techniques (Naive Bayes and KNN algorithms), to detect the CKD at an early stage.

Keywords: Chronic Kidney Disease (CKD); Naive bayes; K-Nearest neighbor; Machine learning

Received: September 04, 2020; **Accepted:** September 18, 2020; **Published:** September 25, 2020

Introduction

Currently, there are many people who are suffering from chronic kidney diseases worldwide. Due to the several risk factors like food, environment and living standards many people get diseases suddenly. Diagnosing of chronic kidney diseases is generally invasive, costly, time consuming and often risky. That is why many patients reach latest age so fit without treatment, especially in those countries where there sources are limited. Therefore, the early detection strategy of the disease remains important, particularly in developing countries, where the diseases are generally diagnosed in later stages. Finding a solution for the above-mentioned problems and riding out from disadvantages became a strong motive to conduct this study. Chronic Kidney Disease (CKD) is one of the types of kidney disease, which results in a gradual loss of kidney function. This phenomenon can be observed over a period of months or years due to several living conditions of patients [1]. Kidney disease is broadly classified into acute kidney injury and chronic kidney disease. Acute kidney injury is sudden damage to the kidneys. In many cases it will be short term but, in some people, it may lead to long-term chronic kidney disease. Chronic kidney disease (CKD) means the 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. The main causes are damaged blood vessels of the kidneys due to High Blood Pressure and Diabetes. The CKD is also called a chronic kidney failure where according current medical statistics the 10% of the population worldwide is affected by CKD [2,3]. There were approximately 58 million deaths in the year of 2005 worldwide. According to the World Health Organization (WHO), 35 million attributed to chronic diseases. Currently it is estimated that one in five men, and one in four women aged 65 through 74 are going to be affected by CKD worldwide. Diagnosing CKD usually starts with clinical data, lab tests, imaging studies and finally biopsy. In this study, by using the machine learning techniques, we are proposing cheap, simple and non-invasive tests that can be performed easily. The data has been obtained from the dataset which is obtained from UCI machine learning repository for CKD patients [4]. By this strategy, we hope to produce "down- staging" (increasing in the proportion of CKD detected at an early stage) of the disease to stages that are more amenable to curative treatment.

Literature Survey

Kunwar, et al. entitled "Chronic Kidney Disease Analysis is Using Data Mining Classification Techniques" published in 2016. 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 [5]. Amirgaliyev, et al. entitled "Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods" published in 2015. 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 [6]. Devika, et al. entitled "Comparative Study of Classifier for Chronic Kidney Disease Prediction Using Naive Bayes, KNN and Random Forest" published in 2019. 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 [7]. Avci E et al. entitled "Performance Comparison of Some Classifiers on Chronic Kidney Disease Data" published in 2018. 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 [8]. Dulhare, et al. entitled "Extraction of Action Rules for Chronic Kidney Disease using Naive Bayes Classifier" published in 2017. 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 [9]. Aljaaf, et al. entitled "Early Prediction of Chronic Kidney Disease Using Machine Learning Supported by Predictive Analytics" published in 2018.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 [10].

Materials and Methods

Proposed method

Chronic kidney disease (CKD) has become a global health issue and is an area of concern. It is a condition where kidneys become damaged and cannot filter toxic wastes in the body. The proposed system predominantly focuses on predicting this life-threatening disease Chronic Kidney Disease (CKD) using Classification algorithms (KNN and Naive Bayes). Proposed system is automation for chronic kidney disease prediction using classification techniques and supervised learning algorithms. The data for dataset is obtained from UCI machine learning repository which contains 25 parameters (features) including the class (CKD or NOT CKD) is given in **Table 1**.

Table 1: Dataset parameters.

Attribute name	Value range	Description
Age	2,…,90	Age
BP	50,…,180	Blood Pressure
SG	1.005, 1.010, 1.015, 1.020, 1.025	Specific Gravity
Al	0,1,2,3,4,5	Albumin
Su	0,1,2,3,4,5	Sugar
RBC	2,1,…,8	Red Blood Cells
PC	Normal, abnormal	Pus Cell
PCC	Present, not present	Pus Cell Clumps
Ba	Present, not present	Bacteria
BGR	22,…,490	Blood Glucose Random
BU	1.5,,…,391	Blood Urea
SC	0.4,…,76	Serum Creatinine
Sod	4.5,…,163	Sodium
Pot	2.5,…,47	Potassium
Hemo	3.1,…,17.8	Haemoglobin
PCV	9,…,54	Packed Cell Volume
WBC	2200,…,26400	White Blood Cell Count
RBC	2.1,…,8	Red Blood Cell Count
Htn	Yes, no	Hypertension
DM	Yes, no	Diabetes Mellitus
CAD	Yes, no	Coronary Artery Disease
Appet	Good, poor	Appetite
PE	Yes, no	Pedal Edema
Ane	Yes, no	Anemia
Class	Ckd, not ckd	Class

The proposed system uses Naive Bayes and K-Nearest Neighbor algorithms for classification. Also, it predicts the stage of the CKD. The entire system is built as a real time application which contains 4 modules (Admin, Receptionist, Doctor and Patient).

The workflow of the proposed system is shown in **Figure 1**.

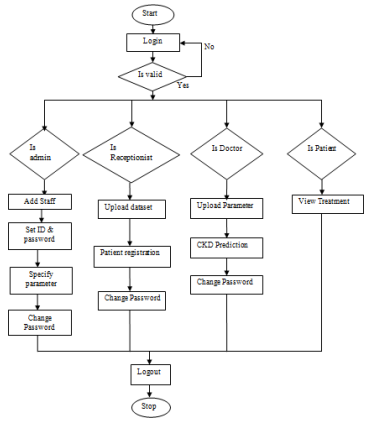

Figure 1: Flow chart of the proposed system.

Results

The application uses KNN and Naive Bayes Algorithms for classification. The application has Admin module which is the main module to maintain the application. After admin's successful login he can add doctors and receptionists. The receptionist will add the training dataset (old patient) and register's the new patient. Doctor can analyze whether a patient have CKD or not and also determine the CKD stage if patient having CKD. Also, doctors have an option to upload treatment details for particular patient. The patient can view his treatment details by logging in to the application. When multiple patient data is analyzed, we got 97% of accuracy using KNN is shown in **Figure 2** and 91% of accuracy using Naive Bayes is shown in **Figure 3**. Graph Comparison of KNN and Naive Bayes is shown in **Figure 4**.

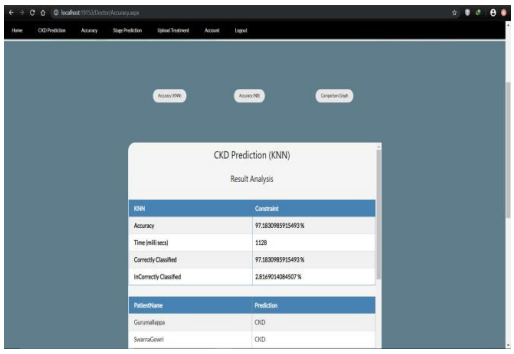

Figure 2: KNN accuracy.

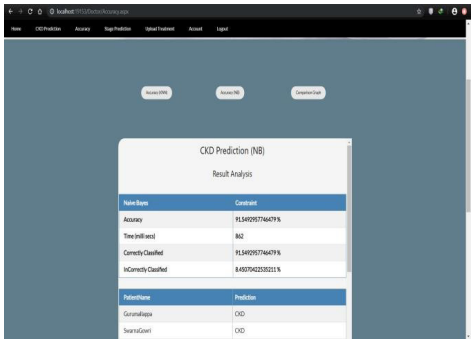

Figure 3: Naive Bayes accuracy.

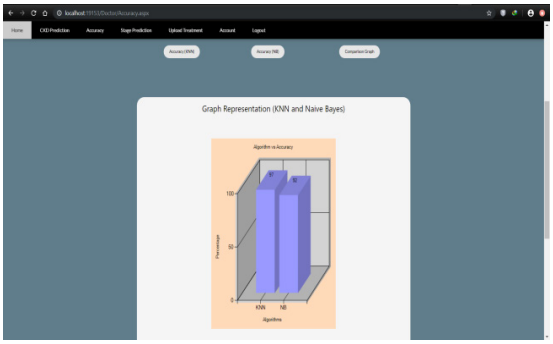

Figure 4: Graph comparison of KNN and Naive Bayes.

Conclusion

This project is a medical sector application which helps the medical practitioners in predicting the CKD disease based on the CKD parameters. It is automation for CKD disease prediction and it identifies the disease, its stages in an efficient and economically manner. It is successfully accomplished by applying the KNN and Naive Bayes algorithms for classification. This classification technique comes under data mining technology. This algorithm takes CKD parameters as input and predicts the disease based on old CKD patient's data.

Future Enhancements

SMS/Email module

In the proposed system, admin assigns Id and password for doctors and receptionists and is intimated manually, so we can add SMS/Email module as a future enhancement where doctors and receptionists receive an SMS or Email regarding the Id and password.

Query module

We can add the query module as a future enhancement to the application where doctor, receptionist and admin of the application can interact with each other.

References

1. Noia TD, Ostuni VC, Pesce F, Binetti G, Naso D, et al. (2013) An end stage kidney disease predictor based on an artificial neural network ensemble. Expert Syst Appl 40: 4438–4445.
2. Levey AS, Atkins R, Coresh RJ, Cohen EP, Collins AJ, et al. (2007) Chronic kidney disease as a global public health problem: approaches and initiatives - a position statement from Kidney Disease Improving Global Outcomes. Kidney Int 72: 247-259.
3. Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, et al. (2013) Chronic kidney disease: global dimension and perspectives. Lancet 382: 260- 272.
4. Soundarapandian P (2015) Chronic Kidney Disease Data Set. UCI Machine Learning Repository, Center for Machine Learning and Intelligent Systems, Irvine, CA: University of California, School of Information and Computer Science.

5. Kunwar V, Chandel K, Sai Sabitha A, Bansal A (2016) Chronic Kidney Disease Analysis Using Data Mining Classification Techniques. 2016 6th International Conference-Cloud System and Big Data Engineering.

6. Amirgaliyev Y, Shamiluulu S, Serek A (2018) Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods. 2018 IEEE 12th International Conference on Application of Information and Communication Technologies (AICT).

7. Devika R, Sai Vaishnavi A, Subramaniyaswamy V (2019) Comparative Study of Classifier for Chronic Kidney Disease Prediction Using Naive Bayes, KNN and Random Forest. 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC).

8. Avci E, Karakus S, Ozmen O, Avci D (2018) Performance Comparison of Some Classifiers on Chronic Kidney Disease Data. 2018 6th International Symposium on Digital Forensic and Security (ISDFS).

9. Dulhare UN, Ayesha M (2016) Extraction of Action Rules for Chronic Kidney Disease using Naive Bayes Classifier. 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC).

10. Aljaaf AJ, Al-Jumeily D, Haglan HM, Alloghani M, Baker T, et al. (2018) Early prediction of chronic kidney disease using machine learning supported by predictive analytics. 2018 IEEE Congress on Evolutionary Computation (CEC).