DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING

Author name:Marwa Almasoud. Year:2015

Abstract—Chronic kidney disease (CKD) is one of the most critical health problems due to itsincreasing prevalence. In this paper, we aim to test the ability of machine learning algorithms for the prediction of chronic kidney disease using the smallest subset of features. Several statistical tests have been done to remove redundant features such as the ANOVA test, the Pearson's correlation, and the Cramer's V test. Logistic regression, support vector machines, random forest, and gradient boosting algorithms have been trained and tested using 10-fold cross-validation. We achieve an accuracy of 99.1 according to F1-measure from Gradient Boosting classifier. Also, we found that hemoglobin has higher importance for both random forest and Gradient boosting in detecting CKD. Finally, our results are among the highest compared to previous studies but with less number of features reached so far. Hence, we can detect CKD atonly $26.65 by performing three simple tests.

Author name:Surya Krishnamurthy Year:22 March 2021

Abstract: Chronic kidney disease (CKD) represents a heavy burden on the healthcare system becauseof the increasing number of patients, high risk of progression to end-stage renal disease, and poorprognosis of morbidity and mortality. The aim of this study is to develop a machine-learning modelthat uses the comorbidity and medication data obtained from Taiwan'sNational Health InsuranceResearch Database to forecast the occurrence of CKDwithin the next 6 or 12 months before itsonset, and hence its prevalence in the population. A total of 18,000 people with CKD and 72,000people without CKD diagnosis were selected using propensity score matching. Their demographic,medication and comorbidity datafrom their respective two-year observation period were used tobuild a predictive model. Among the approaches investigated, the Convolutional Neural Networks(CNN) model performed best with a test set AUROC of 0.957 and 0.954 for the 6-month and 12-month predictions, respectively. The most prominent predictors in the tree-based models wereidentified,including diabetes mellitus, age, gout, and medications such as sulfonamidesandangiotensins. Themodel proposed in this study could be a useful toolforpolicymakers in predicting the trends of CKDin the population. The models can allowclose monitoring of people at risk, early detection of CKD,
better allocation of resources, and patient-centric management.

Author name:ALVARO SOBRINHO Year:2006

Abstract: The high incidence and prevalence of chronic kidney disease (CKD), often caused by latediagnoses, is a critical public health problem, especially in developing countries such as Brazil. CKDtreatment therapies, such as dialysis and kidney transplantation, increase the morbidity and mortality rates,besides the public health costs. This study analyses the usage of machine learning techniques to assistin the early diagnosis of CKD in developing countries. Qualitative and quantitative comparative analysesare, respectively, conducted using a systematic literature review and an experiment with machine learningtechniques, with the k-fold cross-validation method based on the Weka software and a CKD dataset.These analyses enable a discussion on the suitability of machine learning techniques for screening forCKD risk, focusing on low-income and hard-to-reach settings of developing countries, due to the specificproblems faced by them, e.g., inadequate primary health care. The study results

show that the J48 decisiontree is a suitable machine learning technique for such screening in developing countries, due to the easyinterpretation of its classification results, with 95.00% accuracy, reaching a nearly perfect agreement withan experienced nephrologist's opinion. Conversely, random forest, naive Bayes, support vector machine,multilayer perceptron, and Qk-nearest neighbor techniques, respectively, yield 93.33%, 88.33%, 76.66%,75.00%, and 71.67% accuracy, presenting at least moderate agreement with the nephrologist, at the cost of a more difficult interpretation of the classification results.

Author name:Anusor Charleonnan

Abstract:Predictive analytics for healthcare using machine learning is a challenged task to help doctors decide the exact treatments for saving lives. In this paper, we present machine learning techniques for predicting the chronic kidney disease using clinical data. Four machine learning methods are explored including K-nearest neighbors (KNN), support vector machine (SVM), logistic regression (LR), and decision tree classifiers. These predictive models are constructed from chronic kidney disease dataset and the performance of these models are compared together in order to select the best classifier for predicting the chronic kidney disease.

Author name:Vijendra Singh

Abstract:Diabetes and high blood pressure are the primary causes of Chronic Kidney Disease (CKD). Glomerular Filtration Rate (GFR) and kidney damage markers are used by researchers around the world to identify CKD as a condition that leads to reduced renal function over time. A person with CKD has a higher chance of dying young. Doctors face a difficult task in diagnosing the different diseases linked to CKD at an early stage in order to prevent the disease. This research presents a novel deep learning model for the early detection and prediction of CKD. This research objectives to create a deep neural network and compare its performance to that of other contemporary machine learning techniques. In tests, the average of the associated features was used to replace all missing values in the database. After that, the neural network's optimum parameters were fixed by establishing the parameters and running multiple trials. The foremost important features were selected by Recursive Feature Elimination (RFE). Hemoglobin, Specific Gravity, Serum Creatinine, Red Blood Cell Count, Albumin, Packed Cell Volume, and Hypertension were found as key features in the RFE. Selected features were passed to machine learning models for classification purposes. The proposed Deep neural model outperformed the other four classifiers (Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Logistic regression, Random Forest, and Naive Bayes classifier) by achieving 100% accuracy. The proposed approach could be a useful tool for nephrologists in detecting CKD.

Authorname:RameshChandraPoonia Year:2002

Abstract :Kidney disease is a major public health concern that has only recently emerged. Toxins are removed from the body by the kidneys through urine. In the early stages of the condition, the patient has no problems, but recovery is difficult in the later stages. Doctors must be able to recognize this condition early in order to save the lives of their patients. To detect this illness early on, researchers have used a variety of methods. Prediction analysis based on machine learning has been shown to be more accurate than other methodologies. This research can help us to better understand global disparities in kidney disease, as well

as what we can do to address them and coordinate our efforts to achieve global kidney health equity. This study provides an excellent feature-based prediction model for detecting kidney disease. Various machine learning algorithms, including k-nearest neighbors algorithm (KNN), artificial neural networks (ANN), support vector machines (SVM), naive bayes (NB), and others, as well as Re-cursive Feature Elimination (RFE) and Chi-Square test feature-selection techniques, were used to build and analyze various prediction models on a publicly available dataset of healthy and kidney disease patients. The studies found that a logistic regression-based prediction model with optimal features chosen using the Chi-Square technique had the highest accuracy of 98.75 percent.

Author name:Pankaj Chittora Year:2010

Abstract:Chronic Kidney Disease is one of the most critical illness nowadays and proper diagnosis is required as soon as possible. Machine learning technique has become reliable for medical treatment. With the help of a machine learning classifier algorithms, the doctor can detect the disease on time. For this perspective, Chronic Kidney Disease prediction has been discussed in this article. Chronic Kidney Disease dataset has been taken from the UCI repository. Seven classifier algorithms have been applied in this research such as artificial neural network, C5.0, Chi-square Automatic interaction detector, logistic regression, linear support vector machine with penalty L1 & with penalty L2 and random tree. The important feature selection technique was also applied to the dataset. For each classifier, the results have been computed based on (i) full features, (ii) correlation-based feature selection, (iii) Wrapper method feature selection, (iv) Least absolute shrinkage and selection operator regression, (v) synthetic minority over-sampling technique with least absolute shrinkage and selection operator regression selected features, (vi) synthetic minority over-sampling technique with full features. From the results, it is marked that LSVM with penalty L2 is giving the highest accuracy of 98.86% in synthetic minority over-sampling technique with full features. Along with accuracy, precision, recall, F-measure, area under the curve and GINI coefficient have been computed and compared results of various algorithms have been shown in the graph. Least absolute shrinkage and selection operator regression selected features with synthetic minority over-sampling technique gave the best after synthetic minority over-sampling technique with full features. In the synthetic minority over-sampling technique with least absolute shrinkage and selection operator selected features, again linear support vector machine gave the highest accuracy of 98.46%. Along with machine learning models one deep neural network has been applied on the same dataset and it has been noted that deep neural network achieved the highest accuracy of 99.6%.

Authorname:SiddheshwarTekal Year:2018

Abstract: In today's era everyone is trying to be conscious about health although due to workload and busy schedule one gives attention to the health when it shows any symptoms of some kind. But CKD is a disease which doesn't shows symptoms at all or in some cases it doesn't show any disease specific symptoms it is hard to predict, detect and prevent such a disease and this could be lead to permanently health damage, but machine learning can be hope in this problem.it is best in prediction and analysis. By using data of CKD patients with 14 attributes and 400 record we are going to use various machine learning techniques likeDecision Tree, SVM, etc. To build a model with maximum accuracy of predicting whether CKD or not and if yes then its Severity.

Author name:Abhimanyu Agarwala Year:2013

Abstract:Chronic Kidney Disease (CKD) is one of the deadliest diseases that slowly damageshumankidney. The disease remains undetected in its early stage and the patients can only realize the severity of the disease when it gets advanced. Hence, detecting such disease at earlier stage is a key challenge now. Machine Learning is one of the emerging field used in the health sectors for the diagnosis of different diseases. In this paper, we compute, analyze and compare between Machine Learning classification approaches to determine which classification approach is the optimal for the prediction of CKD. K-Nearest Neighbor Classifier, Decision Tree Classifier, GaussianNB, Logical Regression and Artificial Neural Network (ANN) are some renowned machine learning methods which were selected to train the model and based on these results, we can compare and determine which among the following Machine Learning Methods can predict the possibility of CKD at themost accurate level. From this comparative analysis, Logical Regression is found to be the best approach to predict CKD. Methods can predict the possibility of CKD at the most accurate level. From this comparative analysis, Logical Regression is found to be the best approach to predict CKD.

Author name:FuzheMaa

Abstract:The prevalence of chronic kidney disease (CKD) increases annually in the present scenario of research. One of the sources for further therapy is the CKD prediction where the Machine learning techniques become more important in medical diagnosis due to their high accuracy classification ability. In the recent past, the accuracy of classification algorithms depends on the proper use of algorithms for feature selection to reduce the data size. In this paper, Heterogeneous Modified Artifical Neural Network (HMANN) has been proposed for the early detection, segmentation, and diagnosis of chronic renal failure on the Internet of Medical Things (IoMT) platform. Furthermore, the proposed HMANN is classified as a Support Vector Machine and Multilayer Perceptron (MLP) with a Backpropagation (BP) algorithm. The proposed algorithm works based on an ultrasound image which is denoted as a preprocessing step and the region of kidney interest is segmented in the ultrasound image. In kidney segmentation, the proposed HMANN method achieves high accuracy and significantly reducing the time to delineate the contour.

Author name:Madhivanthana Year:2020

Abstract:Chronic kidney disease (CKD) is among the highest twenty causes of death worldwide and affects some 10 percent of the globe adult population. CKD could be a disorder that disrupts traditional urinary organs operation. glomerular Filtration Rate (GFR) and kidney damage markers are employed by researchers round the world to spot CKD as a condition that ends up in reduced urinary organ function over time. An individual with CKD encompasses a higher probability of dying young. Doctors face a tough task in identification the various illnesss connected to CKD at associate early stage so as to stop the disease. This project presents a completely unique deep learning model for the early detection and prediction of CKD. The objectives to make a deep neural network and compare its performance to that of other contemporary machine learning techniques. In tests, the typical of the associated options was accustomed to replace all missing values within the data set. After that, the neural network's optimum parameters were mounted by establishing the parameters and running multiple trials. The proposed Deep neural model outperformed the other 3 classifiers (Support Vector Machine (SVM), call Tree and Random Forest) by achieving smart accuracy. The planned approach can be a great tool for nephrologists in detecting CKD.

Author name:Vijendra Singh. Year:2021.Vijendra Singh

Abstract:Chronic Kidney Disease (CKD) has become a prevalent disease nowadays, affecting people globally around the world. Accurate prediction of CKD progression over time is essential for reducing its associated mortality and morbidity rates. This paper proposes a fast, novel hybrid approach to diagnose Chronic Renal Disease. The proposed approach is based on the optimization of SVM classifier with the hybridized dimensionality reduction approach to identify the most informative parameters for CKD diagnosis. It handles the selection of features through two steps. The first one is a filter-based approach using ReliefF method to assign weights and ranks to each feature of the dataset. The second step is the dimensionality reduction of the best-selected subset by means of PCA, a feature extraction technique. For faster execution of datasets, simultaneous execution on multiple processors is employed. The proposed model achieved the highest prediction accuracy of 92.5% on the clinical CKD dataset compared to existing methods - 'CFS+SVM' (60.45%), 'ReliefF + SVM'(86%), 'MIFS + SVM' (56.72%), 'ReliefF + CFS + SVM' (54.37). The proposed work is also examined on the benchmarked Chronic Kidney Disease Dataset and achieved classification accuracy of 98.5% compared to the accuracy with other methods -'CFS+SVM' (92.7%), 'ReliefF + SVM' (89.6%), 'MIFS + SVM' (94.7%). The experimental outcomes positively demonstrate that the proposed hybridized model is effective in undertaking medical data classification tasks and is, therefore, a promising tool for the diagnosis of CKD patients. The proposed approach is statistically validated with the Friedman test with significant results compared to other techniques. The proposed approach also executes in the least time with improved prediction accuracy and competes with and even outperforms other methods in the literature.