IBM NALAIYATHIRAN:EARLY DETECTION OF CHRONIC KIDNEY DISEASE

TEAM ID:PNT20022TMID25007

LITERATURE SURVEY

[1]Prediction of Chronic Kidney Disease - A Machine Learning Perspective:

P. Chittora *et al*., "Prediction of Chronic Kidney Disease - A Machine Learning Perspective," in *IEEE Access*, vol. 9, pp. 17312-17334, 2021, doi: 10.1109/ACCESS.2021.3053763.

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 oversampling 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%

[2]Chronic Kidney Disease Prediction using Machine Learning Ensemble Algorithm

Nikhila, "Chronic Kidney Disease Prediction using Machine Learning Ensemble Algorithm," ***2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*****, 2021, pp. 476-480, doi: 10.1109/ICCCIS51004.2021.9397144.**

Chronic Kidney Disease is one among the non-contagious illnesses that affect most of the individual in the world. The main factors of risk for the Chronic Kidney Disease are Diabetes, Heart Ailment, Hypertension. The Chronic Kidney Disease shows no symptoms in the early stages and most of the cases are diagnosed in the advanced stage. This leads to delayed treatment to the patient which may be fatal. Machine learning technique provides an efficient way in the prediction of Chronic Kidney Disease at the earliest stage. In this paper, four ensemble algorithms are used to diagnose the patient with Chronic Kidney Disease at the earlier stages. The machine learning models are evaluated based on seven performance metrics including Accuracy, Sensitivity, Specificity, F1-Score, and Mathew Correlation Coefficient. Based on the evaluation the AdaBoost and Random Forest performed the best in terms of accuracy, precision, Sensitivity compared to Gradient Boosting and Bagging. The AdaBoost and Random Forest also showed the Mathew Correlation Coefficient and Area Under the curve scores of 100%. The machine learning model proposed in this paper will provide an efficient way to prevent Chronic Kidney diseases by enabling the medical practitioners to diagnose the disease earlier.

[3]Chronic Kidney Disease Prediction Using Machine Learning Methods

I. U. Ekanayake and D. Herath, "Chronic Kidney Disease Prediction Using Machine Learning Methods," *2020 Moratuwa Engineering Research Conference (MERCon)*, 2020, pp. 260-265, doi: 10.1109/MERCon50084.2020.9185249

Chronic Kidney Disease (CKD) or chronic renal disease has become a major issue with a steady growth rate. A person can only survive without kidneys for an average time of 18 days, which makes a huge demand for a kidney transplant and Dialysis. It is important to have effective methods for early prediction of CKD. Machine learning methods are effective in CKD prediction. This work proposes a workflow to predict CKD status based on clinical data, incorporating data prepossessing, a missing value handling method with collaborative filtering and attributes selection. Out of the 11 machine learning methods considered, the extra tree classifier and random forest classifier are shown to result in the highest accuracy and minimal bias to the attributes. The research also considers the practical aspects of data collection and highlights the importance of incorporating domain knowledge when using machine learning for CKD status prediction.

[4]Predicting and Staging Chronic Kidney Disease using Optimized Random Forest Algorithm

S. Samet, M. R. Laouar and I. Bendib, "Predicting and Staging Chronic Kidney Disease using Optimized Random Forest Algorithm," *2021 International Conference on Information Systems and Advanced Technologies (ICISAT)*, 2021, pp. 1-8, doi: 10.1109/ICISAT54145.2021.9678441

The silent killer Chronic Kidney Disease (CKD) in wealthy countries and listed with the leading causes of death in impoverished countries. Because of its rising incidence, CKD is included in the most serious public health problems. It is apparent that early detection of CKD may reduce the severity of damage in maturity. The patient must go to a diagnostic facility and consult with a doctor. This significant issue has been solved with the introduction of machine learning. This study's main objective is to build a model that can reliably predict a person's risk of acquiring CKD. Data mining and machine learning techniques have been widely employed for forecasting chronic renal disease, but little research has been done mixing imputation approaches at the pre-processing stage and feature selection strategy so that classification accuracy will be enhanced. The CKD Database, which is used in the experiments and consists of 400 records with 25, is accessible through UCI's machine learning repository. It does, however, have a large number of missing values, which is why we proposed combining several missing data imputation strategies to solve the problem. The chi-square test was used to select features in this work. A supervised machine learning classification model called Random Forest (RF) is utilized and optimized with gridsearch to diagnose CKD at an early stage. Following a cross-validation procedure with 5 folders, several metrics were utilized to evaluate the model. Our RF had a 99.24% accuracy. The model's best result is created by considering the 10 best-selected features. When compared to previous studies, our results are among the best for assessment metrics and the ranking accuracy. However, with only fewer features. In practice, some decision assistance for renal illness' diagnosis, prevention, and prediction are provided by this study.

[5]Chronic kidney disease prediction using machine learning techniques

G. Nandhini and J. Aravinth, "Chronic kidney disease prediction using machine learning techniques," *2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, 2021, pp. 227-232, doi: 10.1109/RTEICT52294.2021.9573971.

Early diagnosis and characterization are the important components in determining the treatment of chronic kidney disease (CKD). CKD is an ailment which tends to damage the kidney and affect their effective functioning of excreting waste and balancing body fluids. Some of the complications included are hypertension, anemia (low blood count), mineral bone disorder, poor nutritional health, acid base abnormalities, and neurological complications. Early and error-free detection of CKD can be helpful in averting further deterioration of patient's health. These chronic diseases are prognosticated using various types of data mining classification approaches and machine learning (ML) algorithms. This Prediction is performed using Random Forest (RF) Classifier, Logistic Regression (LR) and K-Nearest Neighbor (K-NN) algorithm and Support Vector Machine (SVM). The data used is collected from the UCI Repository with 400 data sets with 25 attributes. This data has been fed into Classification algorithms. The experimental results show that K-NN, LR, SVM hands out an accuracy of 94%, 98% and 93.75% respectively. The RF classifier gives out a maximum accuracy of 100%

INFORMATION GATHERING :

https://www.niddk.nih.gov/health-information/kidney-disease/chronic-kidney-disease-ckd/what-is-chronic-kidney-disease

https://bmcnephrol.biomedcentral.com/articles/10.1186/s12882-021-02474-z#:~:text=The%20best%20way%20to%20measure,creatinine%20value%20of%20a%20person.

https://www.sciencedirect.com/science/article/pii/S2352914818302387#:~:text=The%20stages%20of%20Chronic%20Kidney,are%20(see%20Table%201)%3A