EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING

(APPLIED DATA SCIENCE)

IBM PROJECT-TEAMID : PNT2022TMID51040

TEAMLEAD

VIDHYA V(953619104048).

TEAM MEMBERS

SANGEETHA P(953619104035)

VITHYA THARANI M(953619104054)

YUVAASRI R K(953619104055)

BACHELOR OF ENGINEERINGN

COMPUTER SCIENCE AND ENGINEERING

RAMCO INSTITUTE OF TECHNOLOGY

ANNAUNIVERSITY-CHENNAI 600025

LITERATURE REVIEW

1. An end stage kidney disease predictor based on an artificial neural networks ensemble :

Author:T. Di Noia et al
Year:2013

T. Di Noia et al. [1], 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, an F-measure. The presented software tool has been made available as both an Android mobile application and online web application.Eventually leading to ESKD in up to 40 percent of patients after 20 years from the renal biopsy (Manno et al.,).

2. Performance evaluation on machine learning classification techniques for disease classification and forecasting through data analytics for chronic kidney disease (CKD)

Author name:W. Gunarathne, K. Perera and K. Kahandawaarachchi
Year:2017

W. Gunarathne, K. Perera, and K. Kahandawaarachchi [2], Microsoft Azore has been used to predict the patient status of CKD. By considering 14 attributes out of 25, they compared four different algorithms, which were Multiclass Decision Forest, Multiclass Decision Jungle, Multiclass Decision Regression, and Multiclass Neural Network. After comparison, they found that Multiclass Decision Forest performed the best with 99.1% accuracy.

3. Applying Machine Learning Techniques for Predicting the Risk of Chronic Kidney Disease

Author :K. A. Padmanaban and G. Parthiban
Year:2016

K. A. Padmanaban and G. Parthiban [3] aimed in their work to detect chronic kidney disease for diabetic patients using machine learning methods. In their research, they used 600 clinical records collected from a leading Chennai-based diabetes research center. The authors have tested the dataset using the decision tree and Naïve Bayes methods for classification using the WEKA tool. They concluded that the decision tree algorithm outweighs the Naïve Bayes with an accuracy of 91%.

4. Early prediction of chronic kidney disease using machine learning supported by predictive analytics

Author : A. J. Aljaaf et al
Year:2018

A. J. Aljaaf et al. [4] examined in their study the ability of four machine learning (ML) models for early prediction of CKD, which were: support vector machine (SVM), classification and regression tree (CART), logistic regression (LR), and multilayer perceptron neural network (MLP). By using the CKD dataset from UCI and seven features out of 24, they compared the performance of these ML models. The results showed that the MLP model had the highest AUC and sensitivity. It was also noticeable that logistic regression almost had the same performance as MLP but with the advantage of the simplicity of the LR algorithm. Therefore, in their study, we can use the LR algorithm as a start or a benchmark and then use more complex algorithms.

5. Comparison and development of machine learning tools in the prediction of chronic kidney disease progression

Author:J. Xiao et al
Year:2019

J. Xiao et al. [5] in their study established and compared nine ML models, including LR, Elastic Net, ridge regression lasso regression SVM, RF, XGBoost, knearest neighbor and neural network to predict the progression of CKD. They used available clinical features from 551 CKD follow-up patients.AUC represents the probability of random positive and random negative.They conclude that linear models have the overall predictive power with an average AUC above 0.87 and precision above 0.8 and 0.8, respectively.

6. Performance Comparison of Some Classifiers on Chronic Kidney Disease Data:

Author: Avci E, Karakus S, Ozmen O, AvciD
Year:2018

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 [6].

7. Extraction of Action Rules for Chronic Kidney Disease using Naive Bayes Classifier

Author:Dulhare UN, Ayesha M
Year:2016

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 [7].

8. Chronic Kidney Disease Analysis Using Data Mining Classification Techniques

Author:Kunwar V, Chandel K, Sai Sabitha A, Bansal A
Year:2016

Kunwar, et al. entitled "Chronic Kidney Disease Analysis is Using Data Mining Classification Techniques" published. 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 [8].

9. Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods

Author:Amirgaliyev Y, Shamiluulu S, Serek A
Year:2018

Amirgaliyev, et al. entitled "Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods" published. 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 [9].

10. Comparative Study of Classifier for Chronic Kidney Disease Prediction Using Naive Bayes

Author:Devika R, Sai Vaishnavi A, Subramaniyaswamy V
Year: 2019

Devika, et al. entitled "Comparative Study of Classifier for Chronic Kidney Disease Prediction Using Naive Bayes, KNN and Random Forest" published . 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 [10].

11. Chronic kidney disease diagnosis usingdecision tree algorithms

Author: Ilyas, H.; Ali, S.; Ponum, M.; Hasan, O.; Mahmood, M.T.; Iftikhar, M.; Malik, M.H.
Year:2021

Chronic Kidney Disease (CKD), i.e., gradual decrease in the renal function spanning over a duration of several months to years without any major symptoms, is a life-threatening disease. It progresses in six stages according to the severity level. It is categorized into various stages based on the Glomerular Filtration Rate (GFR), which in turn utilizes several attributes, like age, sex, race and Serum Creatinine. Among multiple available models for estimating GFR value, Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI), which is a linear model, has been found to be quite efficient because it allows detecting all CKD stages [11].

12. Machine learning algorithm for early detection of end-stage renal disease

Author: Segal, Z.; Kalifa, D.; Radinsky, K.; Ehrenberg, B.; Elad, G.; Maor, G.; Lewis, M.; Tibi, M.; Korn, L.; Koren, G
Year :2020

Z. Segal et al. [12] presented a machine learning technique based on an ensemble tree (XGBoost) for the early diagnosis of renal illness. The presented model was compared against Random Forest, CatBoost, Regression with Regularization. The proposed model showed better performance in all matrices,including c-statistics 0.93, sensitivity 0.715, and specificity 0.958.

REFERENCES

1. T. Di Noia et al, "An end stage kidney disease predictor based on an artificial neural networks ensemble," Expert Syst. Appl., vol. 40, (11), pp. 4438-4445, 2013.
2. W. Gunarathne, K. Perera and K. Kahandawaarachchi, "Performance evaluation on machine learning classification techniques for disease classification and forecasting through data analytics for chronic kidney disease (CKD)," in Bioinformatics and Bioengineering (BIBE), 2017 IEEE 17th International Conference On, 2017.
3. K. A. Padmanaban and G. Parthiban, "Applying Machine Learning Techniques for Predicting the Risk of Chronic Kidney Disease," Indian Journal of Science and Technology, vol. 9, (29), 2016.
4. A. J. Aljaaf et al, "Early prediction of chronic kidney disease using machine learning supported by predictive analytics," in 2018 IEEE Congress on Evolutionary Computation (CEC), 2018.
5. J. Xiao et al, "Comparison and development of machine learning tools in the prediction of chronic kidney disease progression," Journal of Translational Medicine, vol. 17, (1), pp. 119, 2019.
6. 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).
7. 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).
8. 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.
9. 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).
10. 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)
11. Ilyas, H.; Ali, S.; Ponum, M.; Hasan, O.; Mahmood, M.T.; Iftikhar, M.; Malik, M.H. Chronic kidney disease diagnosis using decision tree algorithms. BMC Nephrol. 2021, 22, 273.
12. Segal, Z.; Kalifa, D.; Radinsky, K.; Ehrenberg, B.; Elad, G.; Maor, G.; Lewis, M.; Tibi, M.; Korn, L.; Koren, G. Machine learning algorithm for early detection of end-stage renal disease.