LITERATURE SURVEY

EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING

---

1. Qin, Jiongming, et al. "A machine learning methodology for diagnosing chronic kidney disease." *IEEE Access* 8 (2019): 20991-21002.

In this paper, the CKD dataset was from the UCI which has a large number of missing values was used. KNN imputation was used to fill the missing values. Six machine learning algorithms (logistic regression, random forest, support vector machine, k-nearest neighbour, Naive Bayes classifier and feed forward neural network) was developed for the analysis of prediction. The random forest model achieved the best performance with 99.75% accuracy. An integrated model by combining both logistic regression and random forest was also built. This model achieved an accuracy of 99.83% after ten times of simulation.

2. Almansour, Njoud Abdullah, et al. "Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study." *Computers in biology and medicine* 109 (2019): 101-111.

In this paper, the data of 400 patients with 24 attributes related to CKD was used. The missing values in the dataset were replaced by the mean of the respective attributes. The DL and ML techniques such as Artificial Neural Network (ANN) and Support Vector Machine (SVM) were used for prediction. The optimal parameters and features for both the models were determined by tuning the parameters. The ANN achieved better performance than SVM with the accuracy of 99.75% while the latter with only 97.75%.

3. Vásquez-Morales, Gabriel R., et al. "Explainable prediction of chronic renal disease in the colombian population using neural networks and case-based reasoning." *Ieee Access* 7 (2019): 152900-152910.

In this paper, demographic data of two different populations: people diagnosed with CKD and others without CKD is used. A neural network-based classifier is used to predict whether a person is at a risk of developing CKD. The model achieved an accuracy of 95%. A Case-Based Reasoning(CBR) is used as a twin system for the proposed paradigm for the explanation of CKD predictions. The system was also used to test on population, where 7% of the total population in Colombia were identified as being at risk of developing CKD.

4. Tekale, Siddheshwar, et al. "Prediction of chronic kidney disease using machine learning algorithm." International Journal of Advanced Research in Computer and Communication Engineering 7.10 (2018): 92-96.

In this paper, the data of 400 patients with 24 attributes related to CKD was used. Only the 14 optimal attributes for the prediction of CKD were considered. The various machine learning models like Decision Tree, and SVM was built and compared. The SVM achieved an accuracy of 96.75% while the Decision Tree algorithms with 91.75%. SVM performs better than the Decision Tree algorithms but it is time-consuming the Decision Tree algorithms.

5. Chimwayi, Kerina Blessmore, et al. "Risk level prediction of chronic kidney disease using neuro-fuzzy and hierarchical clustering algorithm (s)." (2017).

In this paper, 10 optimal features contributing to CKD were considered. A Neuro-Fuzzy (fusion of neural networks with fuzzy logic) algorithm was developed to predict the risk of CKD among patients. The model achieved an accuracy of 97%. The prediction results are grouped to identify the percentage of patients at high risk for kidney disease who are more likely to develop diabetes. Using hierarchical clustering, the three clusters formed to show a strong relationship between chronic kidney disease and diabetes.

6. Dulhare, Uma N., and Mohammad Ayesha. "Extraction of action rules for chronic kidney disease using Naïve bayes classifier." ***2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*****. IEEE, 2016.**

In this paper, the UCI dataset for CKD with 25 attributes. The missing values in the dataset were replaced using the mode value of the attributes taken across the dataset. A naive Bayes Classifier was built to predict CKD. The 'One Rule' algorithm is used for dimensional reduction which reduced almost 80% of the attributes in the dataset. Normal Naive Bayes Classifier is compared with a Naive Bayes Classifier with OneR where the latter improved the accuracy by 12.5%.

7. Rubini, L. Jerlin, and P. Eswaran. "Generating comparative analysis of early Stage prediction of Chronic Kidney Disease." ***International Journal of Modern Engineering Research (IJMER)*** **5.7 (2015): 49-55.**

In this paper,a a new chronic kidney disease dataset with three classifiers such as radial basis function network, multilayer perceptron, and logistic regression was proposed. Highest accuracy is achieved by MLP(99.75%) RBF Network followed by RBK Network(98.5%) and then Logistic Regression(97.5%). Type I error, Type II error, type I error rate, Type II error rate,

sensitivity, specificity, F-score and kappa values are also predicted using these classifiers. Among all these classifiers Multilayer perceptron classifier gave good accuracy.

8. Rubini, L. Jerlin, and P. Eswaran. "Generating comparative analysis of early stage prediction of Chronic Kidney Disease." ***International Journal of Modern Engineering Research (IJMER)*** **5.7 (2015): 49-55.**

In this paper, a decision support system to predict Chronic Kidney Disease by comparing the performance of Support Vector Machine and K-Nearest Neighbour classifiers based on calculated accuracy, precision, recall and f measure values was built. From the analysis they concluded that KNN classifier performed better than SVM classifier.

9. Dubey, Abhinandan. "A classification of ckd cases using multivariate k-means clustering." ***International Journal of Scientific and Research Publications*** **5.8 (2015): 1-5.**

In this paper, an adopted K-means Clustering algorithm with a single mean vector of centroids, to classify and make clusters of varying probability of likeliness of suspect being prone to CKD. They observed and stated that the suspects falling in clusters K1 or K3 are surely suffering from CKD. The probability of a suspect lying in K2 cluster to fall in the class of CKD is 0.50545, which implies that the suspect cannot be classified by their L-factor classifier. However, suspects from clusters K1 & K3 were found to be falling in CKD class with full probability.

10. Neves, José, et al. "A soft computing approach to kidney diseases evaluation." ***Journal of medical systems*** **39.10 (2015): 1-9.**

In this paper, the dataset consists 24 attributes, forming five main categories. A hybrid decision support system is developed by Neves and his team, allowed to consider incomplete, unknown, and even contradictory information. This is complemented with an approach to computing centered on Artificial Neural Networks, in order to weigh the Degree-of-Confidence in terms of reasoning procedures and knowledge representation based on Logic Programming. Their study involved 558 patients with an age average of 51.7 years and the chronic kidney disease was observed in 175 cases.