Prediction of Chronic Kidney Disease Using Machine Learning

Himani Hatwar1, Divya Chhaprwal2, Dvarkesh Rokade3, Santosh Kale4

Department of Computer Science
NBN Sinhgad School of Engineering, Pune, Maharashtra, India

Abstract: Chronic Kidney Disease is one of the most serious illnesses nowadays, and it is vital to have a good diagnosis as soon as possible. Machine learning has proven to be effective in medical therapy. The doctor can diagnose the ailment early with the use of machine learning classifier algorithms. This article has examined Chronic Kidney Disease prediction from this standpoint. The Chronic Kidney Disease dataset was obtained from the University of California at Irvine's repository. The artificial neural network, C5.0, Chi-square. After effectively filling out the incomplete data set, 4 machine learning algorithms (Random forest, Support vector machine, K-nearest neighbor and Decision tree) were used to establish models. The dataset was also subjected to the significant feature selection technique. The results were computed for each classifier using (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, and (vi) synthetic minority over-sampling technique with full features. The results show that in synthetic minority over-sampling technique with full features, LSVM with penalty L2 has the maximum accuracy of 98.86 percent. Along with precision, recall, F-measure, and area, accuracy, precision, recall. Among these machine learning models, random forest achieved the best performance with 99.75% diagnosis accuracy. Hence, we speculated that this methodology could be applicable to more complicated clinical data for disease diagnosis.

Keywords: Chronic Kidney Disease, Machine Learning, Prediction.

REFERENCES