Chronic Renal Disease Prediction using Clinical Data and Different Machine Learning Techniques

Md. Mohsin Sarker Raihan1 , Eshtiak Ahmed2 , Asif Karim3 , Sami Azam*4 , M. Raihan5 , Laboni Akter6 and Md. Mehedi Hassan7 Department of Biomedical Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh1,6 Faculty of Infomration Technology and Communication Sciences, Tampere University, Tampere, Finland2 College of Engineering, IT and Environment, Charles Darwin University, Darwin, NT, Australia3,4 Department of Computer Science and Engineering, North Western University, Khulna, Bangladesh5,7 Emails: msr.raihan@gmail.com1 (ORCID: 0000-0002-0401-312X)1 eshtiak.ahmed@tuni.fi2 , asif.karim@cdu.edu.au3 , sami.azam@cdu.edu.au4 , raihanbme@gmail.com5 (ORCID: 0000-0003-1072-3555)5 , laboni.kuet.bme@gmail.com6 (ORCID: 0000-0002-0063-3188)6 , mehedihassan@ieee.org7 (ORCID: 0000-0002-9890-0968)7

Abstract—Chronic Renal Disease (CRD) or Chronic Kidney Disease (CKD) is defined as the continuous loss of kidney function. It's a long-term condition in which the kidney or renal doesn't work properly, gets damaged and can't filter blood on a regular basis. Diabetes, high blood pressure, swollen feet, ankles or hands and other disorders can cause chronic renal disease. By gradual progression and lack of treatment, it can lead to kidney failure. A prior prognosis of CKD can nourish the quality of life to a higher range in such circumstances and can enhance the attribute of life to a larger province. Now a days, bioscience is playing a significant role in the aspect of diagnosing and detecting numerous health conditions. Machine Learning (ML) as well as Data Mining (DM) methods are playing the leading role in the realm of biosciences. Our objective is to predict and diagnose (CKD) with some machine learning algorithms. In this study, an attempt to diagnose chronic renal disease has been taken with four ML algorithms named XGBoost, Adaboost, Logistic Regression (LR) as well as Random Forest (RF). By using decision tree-based classifiers and analyzing the dataset with comparing their performance, we attempted to diagnose CKD in this study. The results of the model in this study showed prosperous indications of a better prognosis for the diagnosis of kidney diseases. Considering and contemplating the performance analysis, it is accomplished that Random Forest ensemble learn ing algorithm provides better classification performance than other classification methods.

Index Terms—*Chronic Kidney Disease, CKD, XGBoost, Ad aboost, Logis-tic Regression, Random Forest.*