

Introducing Novel Classification Methodology to Detect Kidney Disease Patterns in Sri Lanka

K. A. Ayesha Chathurangi^{a,1}, R. M. Kapila, Tharanga Rathnayaka^b

^a*Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka*

^b*Department of Physical Sciences and Technology, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka*

The healthcare sector has vast amount of medical data which are not properly analyzed and mined to discover useful information and interesting patterns. Applying data mining techniques on such domain can help medical practitioners to predict even the crucial diseases with ease. This study introduced a novel kidney disease classification methodology in Sri Lankan domain using data mining techniques. Basically there are two types of kidney diseases that can be found in Sri Lanka namely Chronic Kidney Disease (CKD) and Acute Kidney Disease (AKD). The aim of this work is building a model to predict whether a person has a risk on having a kidney disease or not and a model for CKD prediction. The data collected from 108 patients are used to train and test the models. Random Forest algorithm and a multilayered feed forward neural network were used to build the models. Result of this study is a modified Artificial Neural Network with 2 hidden layers to detect kidney disease which gives 0.80952 accuracy and a model with the combination of Random Forest algorithm and Artificial Neural Network with 3 hidden layers for CKD prediction which gives 0.81395 accuracy for testing data. The constructed models give high accuracy and minimum error rate when comparing with the other data mining algorithms.

Keywords: Data mining; Random Forest; Neural Network; Algorithm

¹ Corresponding author: K.A. Ayesha Chathurangi Tel.: +94-71-600-4258
E-mail address: ayesachathurangi1110@gmail.com