Decision Tree to Guide Chronic Kidney Disease Patients at Incipient Stage

Krupanidhi SREERAMA*

Consultant Distinguished Professor, Dept. of Biotechnology, Vignan's Foundation for Science, Technology and Research, Vadlamudi 522213 AP, India
E-mail: sknbirac@gmail.com

* Author to whom correspondence should be addressed; Mobile: +91 9440984668

Abstract
Background: The etiologies of kidney disease are sequential and interlinked. They are known as disease markers. The Decision Tree representing initial symptoms needs to be popularized to seek prophylactic measures and educate the community health workers. *Material and Methods*: Chronic Kidney Disease dataset containing two classes, namely CKD and not CKD was retrieved from UC Irvine Machine Learning Repository available from https://archive.ics.uci.edu/ml/index.php. This dataset includes 400 instances and 25 attributes related to the CKD. The CKD dataset was saved in ARFF format. J48 algorithm was chosen from the WEKA software tool to develop the Decision Tree. *Results*: The Decision Tree comprises root nodes, branches, internal nodes, and leaf nodes. The root node began in J48 classifier with 'sc' (serum creatinine). This node gave two branches, of which one branch led to the internal node 'pe' (pedal edema) (sc<=1.2), and the other (sc>1.2) ended with a leaf node showing the condition CKD. The internal node 'pe' further yielded two branches, the branch with pedal edema terminated with CKD and the branch with no pedal edema led to the internal node. Internal node with 'dm' (diabetes mellitus) showed two more branches namely 'yes', which lead to CKD, and branch 'no' 'dm' led to another internal node namely hemo (haemoglobin). In continuation, the nodes for yet another two disease markers, namely hemoglobin and specific gravity of serum of people prone to kidney diseases, were shown at the terminal end of the Decision Tree. *Conclusion*: The Decision Tree developed for the CKD dataset by using J48 classifier would guide prospective patients with their clinical data reports.

Keywords: Chronic Kidney Disease; UCIML repository; WEKA; J48 algorithm; Decision Tree