Analysis of the Attributes of Liver Cirrhosis using the Machine Learning Tools

Gireesh NAMINENI[a] and Krupanidhi SREERAMA[b,*]

[a] Department of Electronics and Communication Engineering, Mohan Babu University, Tirupati, India
[b] Department of Biotechnology, Vignan's Foundation for Science, Technology and Research, Vadlamudi, 522213 AP, India
E-mail: naminenigireesh@gmail.com; (*) sknbirac@gmail.com

* Author to whom correspondence should be addressed;

Abstract

Background: The liver controls the metabolic process of our physiological system. The hepatic portal system is the pathway via which food and other ingredients we ingest enter the liver. The liver is a digestive gland or organ that secretes and controls bilirubin, transaminase enzymes, and several other biochemical components. As a result, the liver is a crucial organ in human physiology. Various natural and acquired disorders connected with the liver, such as inflammation, fibrosis, jaundice, cirrhosis, and others, cause liver damage and ultimately death. *Methods*: The cirrhosis dataset was obtained from the URL https://www.kaggle.com/datasets/fedesoriano/cirrhosis-prediction-dataset?resource=download.
WEKA, an ML software package, was used to analyze the dataset. The Random Forest algorithm (a tree classifier) and the Multilayer Perceptron algorithm (a function classifier) were selected in the WEKA tool to examine the eight attributes in the cirrhosis dataset. The chosen eight features are mostly responsible for diagnosing liver cirrhosis. Ascites, hepatomegaly, edema, bilirubin, albumin, alkaline phosphatase, serum glutamic oxaloacetic transaminase, and prothrombin are among them. The first 199 instances from the dataset were chosen for the current investigation. *Results*: The Random Forest algorithm produced a correlation coefficient of 0.9664 with a least mean absolute error of 0.2284, indicating that the data in the cirrhosis dataset is uniform and corresponds to the disease phases of cirrhosis. The multilayer perceptron technique produced a correlation coefficient of 0.3226 with a mean absolute error of 0.8921, indicating that the Random Forest approach is well suited to investigate the accuracy of the attributes of cirrhosis. *Conclusion*: Eight clinical features that made up the etiology of cirrhosis were taken into account in the current study, and the Random Forest algorithm confirmed their statistical validity.

Keywords: Cirrhosis; Random Forest; Multilayer Perceptron; WEKA; Healthcare digital ecosystem

Appl Med Inform 45(Suppl. S1) September/ 2023

S24