

Improving the Accuracy of Prediction in Liver Disease Using Machine Learning Techniques

By: Othman Ali Mahmood Ayoub Agha
Supervisors: Dr. Yousef Elsheikh & Dr.Fadi Almasalha.

Abstract

These days, people are working hard to care for their health to live a comfortable life. But people do not care when they work as they eat any food, they forget to live a healthy lifestyle, and they do not do any physical activity, all of which led to the infection of liver disease at an early age.

The liver is the biggest organ after the skin responsible for clotting blood oxygen, removing harmful chemicals that your body makes such as toxins from the blood. The liver makes a liquid called bile, which helps humans break down fat from food and it stores glucose, which gives the human a fast energy boost when need it. Liver disease occurs when the liver does not work well and loss many of function. The most common cause of liver disease infection by viruses, immune system abnormality, and Genetics. Predicting the risk of liver disease may help save the patient from liver disease.

Living a healthy lifestyle, eat a healthy diet and get regular exercise, keep, and maintain weight under control, and use medications wisely all help prevent liver disease.

The prediction of Liver disease is a complicated process depending on many factors, such as Diabetes type two, obesity, family history of liver disease, exposure to proven chemicals or toxins, the risk of prediction with the liver disease increases even if they are young more than non-diabetic patients. Sometimes many factors together increase the risk very much, and each factor alone increases the risk too, therefore the effects of each factor need to be understood, and the effect of each factor with other factors also needs to be understood. Therefore, using machine learning techniques might be promising to address the complicated process of liver disease prediction.

In this thesis, we will fill this gap by proposing a machine learning based mechanism to improve the accuracy of predicting liver disease. This mechanism utilizes some feature processing and selection techniques as well as ensemble learning approach to achieve the best results in comparison.

The obtained results from this thesis were encouraging and showed accuracy improvement in liver disease predication compared to other studies in the literature that used the same Dataset named Indian patient liver disease Dataset taken from UCI Repository. The best accuracy obtained for liver disease prediction using Naïve Bayes was 73.87%, and by using K-Nearest Neighbour was 76.84%.