LITERATURE SURVEY FOR STATISTICAL MACHINE LEARNING APPROACHES FOR LIVER DISEASE PREDICTION

Introduction

Liver diseases disturb the normal function of the liver. One of the main reasons for liver diseases is a large amount of alcohol consumption. Early prediction of liver disease using classification algorithms is an efficacious task that can help doctors to diagnose the disease within a short duration of time. Discovering the existence of liver disease at an early stage is a complex task for doctors. The main objective of this project is to analyze the parameters of various classification algorithms and compare their predictive accuracies so as to find out the best classifier for determining liver disease.

The life of humans living without liver tumors is one of the fundamental care of human livelihood. Therefore, for better care, the detection of liver disease at a primitive phase is necessary. For medical experts, predicting the illness in the early stages due to subtle signs is a very difficult task. Many, when it is too late, the signs become evident. The current work aims to augment the perceived nature of the liver disease by means of machine learning methods to solve this epidemic. The key purpose of the present work focused on algorithms for the classification of healthy people from liver datasets. Centered on their success variables, this research also aims to compare the classification algorithms and to provide prediction accuracy results.

Existing Solutions

- A number of network models based on neural have been developed in recent research assisting physicians with liver diagnosis in the medical field, such as the diagnostic support system, the expert system, the perceptive diagnostic model, and the hybrid recommendation framework.
- Christopher N suggested a system for the diagnosis of medical diseases, taking into account six benchmarks: liver, hepatitis heart, diabetes, breast, and lymph disorders. The researchers developed WSO and C4.5-based systems.
- Ramana has also conducted an acritical study on the diagnosis of liver diseases.

Drawbacks and Limitations

- From Christopher's Observation, the researchers developed WSO and C4.5-based systems, with a precision of 64.60 percent with 19 liver disorder dataset rules and 62.89 percent with 43 WSO and C4.5 rules, respectively. In the evaluation of identified categorization techniques.
- From Ramana's view, On the Naïve Bayes classifier, the authors gained 51.59 percent accuracy, 55.94 percent on the C4.5 algorithm, 66.66 percent with respect to BPNN, 62.6 percent with respect to Knowledge discovery, and sixty-two percent accuracy with respect to vector machine support algorithm.

Proposed Solutions

- SVM is a technique of supervised learning that pertained to classification as well as regression. It has effective performance in generalization. Moreover, when the algorithm requires input space with high dimensions, there is no requirement to add a previous understanding. This helps make it a very effective classifier for quality. The primary purpose of the SVM classifier is to classify between groups of various classifications by choosing the best classifier function in the training data. A generalized linear method of classification is SVM. At the same time, the geometric margin is maximized and the classification error is minimized.

- This Project examines data from liver patients concentrating on relationships between a key list of liver enzymes, proteins, age, and gender using them to try and predict the likeliness of liver disease. Here we are building a model by applying various machine learning algorithms to find the best accurate model. And integrate into flask-based web applications. Users can predict the disease by entering parametersin the web application.

Conclusion

SVM, Logistic Regression, comprises two main machine learning techniques used. Using all the models, the prediction analysis has been implemented and their performance has been assessed. The probability of liver disease prediction was attained with an accuracy of 96%. In the future, the present scenario can be compared with other techniques such as naïve Bayes classification, Random forest,etc.

References

- Yongjun Piao, Minghao Piao, Keun Ho Ryu, **MULTICLASS CANCER CLASSIFICATION USING A FEATURE SUBSET BASED ENSEMBLE FROM microRNA EXPRESSION PROFILES, COMPUTERS IN BIOLOGY AND MEDICINE 80** -2017.
- Sanjay Kumar, Sarthar Katyal, **EFFECTIVE ANALYSIS AND DIAGNOSIS OF LIVER DISORDER BY DATA MINING** – 2018.
- Ajay S Singh, MD Irfan, Abishek**, PREDICTION OF LIVER DISEASE USING CLASSIFICATION ALGORITHMS** – 2018
- Md Fazle, SM Mahedy Hasan, Arifa , Md Asif, Md kamrul, **PREDICTION OF LIVER DISORDERS USING ML ALGORITHMS**-2020
- C Geetha, AR Arunachalam, **EVALUATION BASED APPROACHES FOR LIVER DISEASE PREDICTION USING MACHINE LEARNING ALGORITHMS** – Jan 2021.

PROBLEM STATEMENT

To develop a Machine Learning Model for Liver Disease Prediction at an early stage by applying various machine learning algorithms and finding an accurate model. After that, it is integrated into the flask-based web application where users can predict the disease by entering parameters in the web application.