

# **LITERATURE SURVEY**

## **Statistical Machine Learning approaches to Liver disease prediction**

### **1. A Comparative Study on Liver Disease Prediction Using Supervised Machine Learning Algorithms**

**A.K.M Sazzadur Rahman, F. M. Javed Mehedi Shamrat, Zarrin Tasnim, Joy Roy, Syed Akhter Hossain**

**ResearchGate – 2019**

Chronic Liver Disease is the leading cause of global death that impacts the massive quantity of humans around the world. This disease diagnosis is very costly and complicated. Therefore, this paper evaluates the performance of different Machine Learning algorithms in order to reduce the high cost of chronic liver disease diagnosis by prediction. Six machine learning techniques have been applied including Logistic Regression, K Nearest Neighbors, Decision Tree, Support Vector Machine, Naïve Bayes, and Random Forest. The performance was evaluated on different measurement techniques such as accuracy, precision, recall, f-1 score, and specificity and the result was that LR achieved the highest accuracy.

### **2. Machine learning-based liver disease diagnosis: A systematic review**

**Rayyan AzamKhan, Yigang Luo, Fang Xiang Wu**

**ScienceDirect – 2022**

This paper mainly focuses on the computer-aided diagnosis of hepatic lesions in view of diffuse- and focal liver disorders. This is based on three image acquisition modalities: ultrasonography, computed tomography, and magnetic resonance imaging. Insightful analysis is presented for each preliminary step, particularly preprocessing, attribute analysis, and classification techniques to accomplish clinical diagnostic tasks. In preprocessing denoising, deblurring, and segmentation methods are used. Denoising is mainly performed with nonlinear models.

### **3. Diagnosing of Liver Disease Prediction in Patients using combined Machine Learning Models**

**Chokka Anuradha, D Swapna, Balamuralikrishnan Thati**

**IEEE**

In the human body one of the most important organs is the liver. If the regular functionality of the liver is disturbed then this condition is called disease-affected liver. Therefore, an early stage of disease detection is more important which helps in disease prevention at starting stage with small medications. But, it is too difficult to identify Liver disease at the early stages because symptoms are very less at the starting stage. Lab results with physical examination are involved in the Traditional methods. This paper aims to represent a Diagnosing for Liver disease prediction in Patients using Combined Machine Learning Models. Optimized three machine learning algorithms are used for the accurate diagnosis of liver disease and they are Artificial Neural Networks (ANN), Decision Trees, and K-Nearest Neighbors (KNN). With the help of these algorithms, given data is classified and results are produced. The future data is predicted with the help of past and present data. The accuracy results are produced by comparing three classification algorithms.

### **4. Statistical Machine Learning Approaches to Liver Disease Prediction**

**Fahad Mostafa, Easin Hasan, Morgan Williamson, Hafiz Khan**

**MDPI - 2021**

ML algorithms are trained to detect the possibility of liver disease to assist healthcare workers. Correlation of chosen variable with the risk of liver disease is performed to train the model. ML methods were able to identify the liver disease with high accuracy. The PCA results showed five important factors for liver disease diagnosis: AST, ALT, GGT, BIL, and ALP. In a real situation, a clinician can strongly suspect liver disease using only these five variables, as they are very descriptive for liver function. The ratio of ALT and AS denotes the cause of a liver injury. GGT and ALP increase in circulation with the severity of a liver injury. Additionally, the injury proximity to the bile duct is determined by the concentration of ALP. This study shows several machine learning approaches with PCA, which outperformed the classification. Among three ML classification methods, the performance of SVM and RF is better than ANN.

## **5. Liver Disease Prediction System using Machine Learning Techniques**

**Rakshith D B, Mrigank Srivastava, Ashwani Kumar, Gururaj S P**

**IJERT – 2021**

In this paper risk of liver disease for a person is predicted based on the blood test report results of the user. With the dataset used for this project, 100 % accuracy is obtained for SVM model. The data preprocessing was done using Jupyter Notebook and Desktop Application was Implemented using Synder IDE. The programming language which was used is python and machine learning Sklearn was used to build the model using classification algorithm like KNN, SVM, Naive Bayes and ANN.

## **6. Statistical Machine Learning Approaches to Liver Disease Prediction**

**Robin Biju**

**International Journal of Scientific Research and Engineering Development – 2022**

This study attempts to find an appropriate machine learning algorithm that can determine whether a person has liver disease or not given a dataset containing biological and diagnostic data of 583 Indian patients. Using certain characteristics such as total bilirubin, direct bilirubin, alkaline phosphatase, total protein, albumin, and globulin, this software can determine whether a patient has liver disease or not.

## **Conclusions:**

The mechanisms that are currently used in the prediction of liver disease are prone to have different levels of accuracy and effectiveness. Different diseases demand accuracy of a different set of parameters and might not demand the same set of inferences, throughout more than a single case. In the near future, the study reflects that there was a decent amount of accuracy that was achieved. However, the agenda of our paper is to improvise on those lines and come up with better accuracy standards. The following are some of the clear limitations that have been observed, in order to account for innovation in this paper, having brought about the connotation of improvising on these lines.

- When it comes to the classification process, it is not necessary that the cohesion that a classifier shares with a particular set of data should stand viable for the rest of the training set. This is to imply that there are some classifiers that don't stand fit to the data set in the context.
- There are certain methodologies that are incompatible and non-cohesive when it comes to the collection of real-time data and the implementation procedures of the same.
- Some of the machine learning approaches that are being considered, do not stand viable for a large volume of data.