JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Liver Disease Prediction Using Machine Learning

[1]Miss.Printi Yadav,[2]Miss.Sanika Yadav,[3]Miss.Ayodhya Bachche,[4]Miss.Shreya Solankure.

[5] Mr. Rahul S. Kumbhar.

[1]student, [2] Student, Student ,[4] Student,[5]Guide
[1] Department Of Computer Engineering ,
[1]D. Y. Patil Technical Campus, Talsande , Kolhapur, India

Abstract : Recently liver diseases are becoming most lethal disorder in a number of countries. The count of patients with liver disorder has been going up because of alcohol intake, breathing of harmful gases, and consumption of food which is spoiled and drugs. Liver patient data sets are being studied for the purpose of developing classification models to predict liver disorder. This data set was used to implement prediction and classification algorithms which in turn reduces the workload on doctors. In this work, we proposed apply machine learning algorithms to check the entire patient's liver disorder.

Machine Learning is a process which is used to discover patterns in huge data/ large data set to enable decision, thereby allowing machines to go through a learning process (i.e. supervised, unsupervised and semi-supervised or reinforced). The data set used in this paper is Liver Patient taken from UCI Repository (i.e. Supervised Learning).

IndexTerms **- Liver, diseases, Liver Disease Machine Learning, health**

I. INTRODUCTION

Liver is your body's second-largest organ. The liver is a large, meaty organ that sits on the right side of the belly. Weighing about 3 pounds; the liver is reddish-browning colour and feels rubbery to the feel. The liver has two large sections, called the right and the left lobes. The gallbladder sits below the liver, along with parts of the pancreas and intestines. The liver and these organs behaviour together to digest, absorb, and process food. The liver's main is to strain the blood coming from the digestive tract, before passing into the rest of the body.The liver also detoxifies chemicals and metabolizes drugs.

learning techniques have demonstrated encouraging results in predicting and categorizing liver diseases based on patient data. By utilizing sophisticated algorithms to analyze and learn from large datasets, these techniques can identify patterns and anticipate outcomes. The employment of machine learning techniques in liver disease prediction and classification is a dynamic area of research, with continual advancements being made to enhance accuracy and decrease healthcare costs.

II. MOTIVATION

Scarring in your liver blocks the flow of blood and oxygen through your liver tissues. This slows your liver's ability to process your blood, metabolize nutrients and filter out toxins. Cirrhosis reduces your liver's ability to produce bile and essential blood proteins. Scar tissue can also compress blood vessels running through your liver, including the important portal vein system, leading to a condition called portal hyper tension.

Cirrhosis is relatively common and is a significant cause of hospitalization and death, especially after middle age. That's because it develops gradually over time. In the United States, cirrhosis affects about 0.25% of all adults and about 0.50% of adults between the ages of 45 and 54. Each year, about 26,000 deaths in the United States are attributed to cirrhosis, and these rates are rising. Cirrhosis is global health concern.deaths in the United States are attributed to cirrhosis, and these rates are rising. Cirrhosis is global health concern.

III. LITERATURE REVIEW

Liver Diseases Prediction is developed by a set of Machine Learning Approaches [1]In this paper, the researcher used a set of algorithms to predict liver disease by classifying a data set (ILPD) such as Decision Tree, Perceptron, Random Forest, K-Nearest Neighbor, Support vector machine, with feature selection and without feature selection, and it indicates that the best performance was the algorithm KNN with an accuracy of 74% with selecting the features, and the accuracy of the performance of the algorithms without selecting the features was as follows: Decision Tree 60%, Perceptron 39%, Random Forest 64%, K-Nearest Neighbor 66%, Support vector machine 71%, And with feature selection was as follows Decision Tree 72%, Perceptron 66%, Random Forest 73%, K- Nearest Neighbor 74%, Support vector machine 72% After studying this research, it is clear from the results it reached that it still needs improvement[1].

Prediction of Liver Malady Using Advanced Classification Algorithms In this paper, the author used SVM to predict liver disease using the ILPD dataset and obtained an accuracy of 78%. But in this research, the researcher did not perform his experiment on a set of algorithms, but only one algorithm, and also did not predict the severity of liver disease [2]..

P. Rajeswari,G.Sophia Reena et al.,[2010]has proposed the data classification is based on liver disorder. The training dataset is developed by collecting data from UCI repository consists of 345 instances with 7 different attributes. This paper deals with results in the field of data classification obtained with Naïve Bayes algorithms. FTtree algorithms, and KStar algorithms and on the whole performance made know FT Tree algorithm when tested on liver disease datasets, time taken to run the data for result is fast when compare to other algorithm with accuracy of 97.10%Based on the experimental results the classification accuracy is found to be better using FT Tree algorithm[3].

Dhamo dharan et.al has predicted three major liver diseases such as Liver cancer, Cirrhosis and Hepatitis with the help of distinct symptoms. They used Naïve Bayes and FT Tree algorithms for disease prediction. Comparison of these two algorithms has been done based on their classification accuracy measure. From the experimental results they concluded the Naïve bayes as the better algorithm which predicted diseases with maximum classification accuracy than the other algorithm [4].

IV. ANALYSIS

Table 1.1: Count Plot show is the rate of liver patients

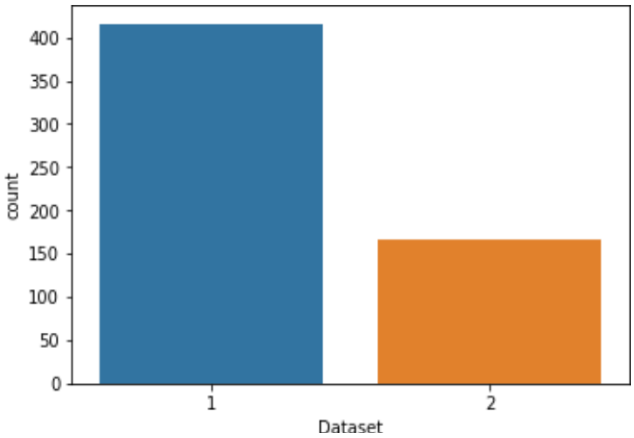


Table 1.1: Multi-Collinearity found on our dataset

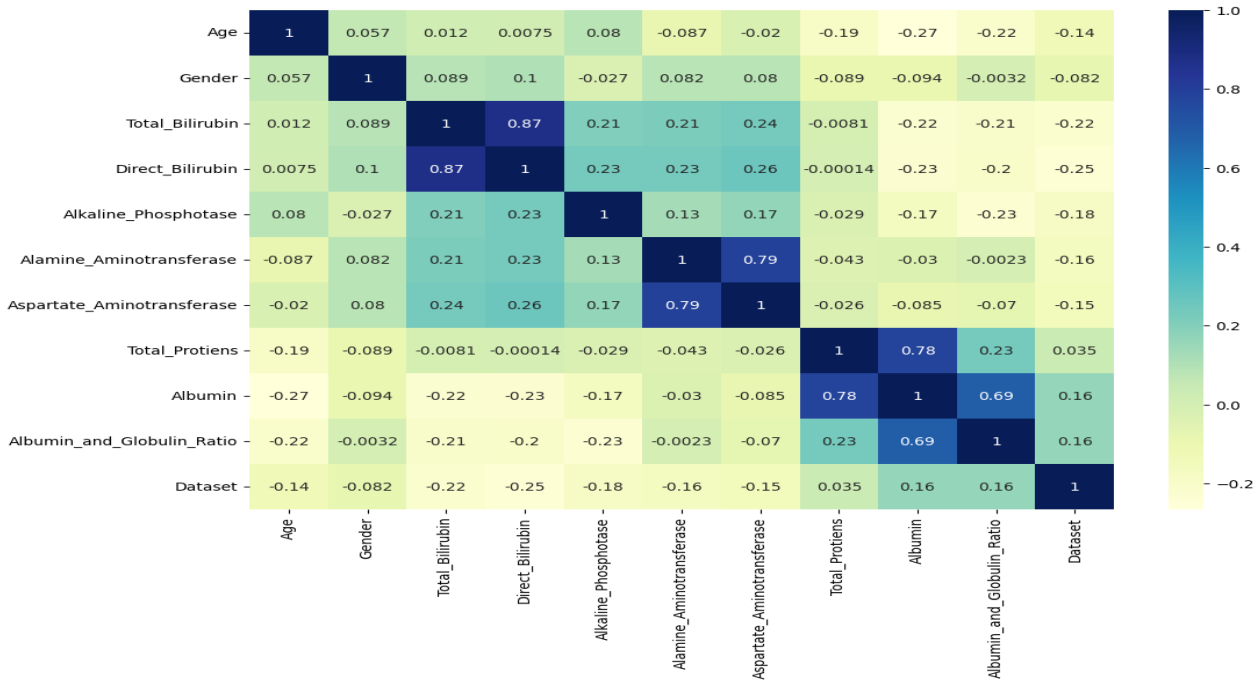
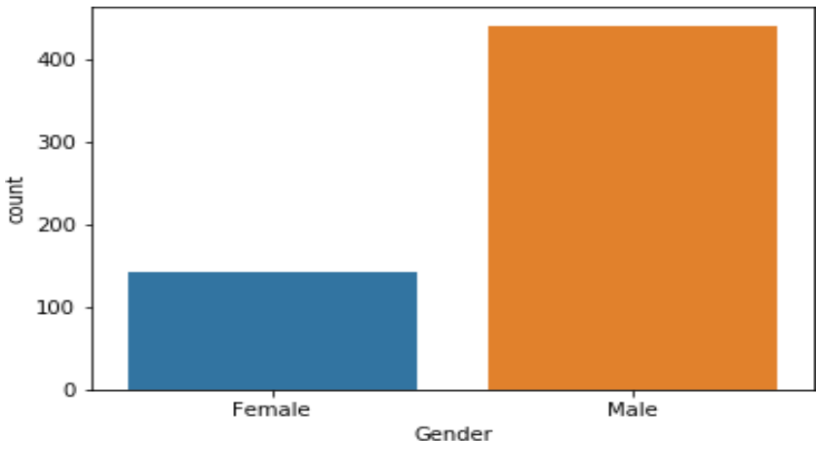


Table 1.3: Count Plot shows the ratio of gender of liver patients.



V. METHODOLOGY

Main objective of this study is to identify that whether the patient has liver disease or not .Some of the parameter are used for predicting the liver disease and compare the performance of the various decision tree techniques. Wake is a data mining tool which is written in java and developed at Waikato.

WEKA is a very efficient data mining tool to classify the accuracy by applying different algorithm approaches and compare on the basis of datasets [12].It is also a good tool for build new machine learning schemes. There result found from the liver disease dataset by using Weka tool are in section4.10-foldcrossvalidationperformedonthe dataset.

This data set contains 416 liver patient records and 167 non liver patient records collected from North East of Andhra Pradesh, India. The "Dataset" column is a class label used to divide groups into liver patient (liver disease) or not (no disease). This data set contains 441 male patient records and 142 female patient records. Use these patient records to determine which patients have liver disease and which ones do not.

V. MODULES

A. Module 1 *: **Data Collection***

In this experiment, we collect a dataset from the UCI Machine Learning Repository. In addition, the original dataset was collected from the northeast of Andhra Pradesh, India [7]. This dataset consists of 583 liver patient's data whereas 75.64% male patients and 24.36% are female patients. This dataset has contained 11 particular parameters whereas we choose 10 parameters for our further analysis and 1 parameter as a target class. Such as,

I. Age: Age of the patient

II. Gender: Gender of the Patients

III. TB: Total Bilirubin

IV. DB: Direct Bilirubin

V. Alkphos: Alkaline Phosphotase

VI. Sgpt: Alamine Aminotransferase

VII. Sgot: Asparatate Aminotransferase

VIII. TP: Total Proteins

IX. ALB: Albumin

X. AG Ratio: Albumin and Globulin Ratio

XI. Selector field used to split the data into two sets

B. Module 2: ***Data Pre-processing***

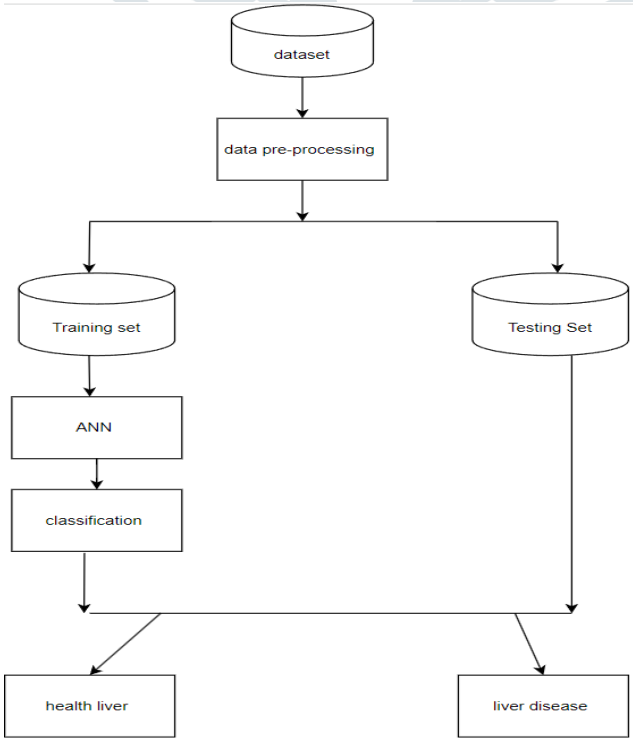
In this study, we analyzed 583 liver patient's data whereas 416 samples are liver patient and 167 samples are non-liver patients. The ratio of total liver patients is presented in table1.2 Moreover, from the liver patient's dataset, (table1.1) 441 are male samples and 112 are female samples were taken for analysis.
The heat map is shown in (table1.3) appear to have some correlated parameters. Some of these columns have a low correlation. Therefore, we omitted some of the features for better prediction of liver disease.

C. Module 2 : ***Description of the classification Algorithms***

Artificial Neural Networks (ANN)

Artificial neural network (ANN) model involves computations and mathematics, which simulate the human–brain processes. Many of the recently achieved advancements are related to the artificial intelligence research area such as image and voice recognition, robotics, and using ANNs.

VI. ARCHITECTURE



VIII. SCOPE

i. Diagnostic Biomarkers: Biomarkers play a crucial role in predicting liver disease.
ii. Telemedicine platforms and remote monitoring devices enable healthcare providers to remotely assess patients' liver health.
iii. PSO has been widely used for feature selection to improveliver classification performance.
iv. Advanced Imaging Techniques: High-resolution imaging combined with artificial intelligence algorithms will provide more accurate and early detection of liver abnormalities.

IX. CONCLUSION

There are many criterions for evaluating the selected feature subset, here this thesis used features such as Total bilirubin, Direct_ bilirubin, Total_protiens, Albumin to evaluate the performance of different classification algorithm.
We will thus utilise the proportion of people who get the condition as both positive and negative data. machine learning techniques have demonstrated encouraging results in predicting and categorizing liver diseases based on patient data.

REFERENCES

[1] "M. S. Azam, A. Rahman, S. M. H. S. Iqbal" Prediction of liver diseases by using few machine learning based approaches, Aust. J. Eng. Innov. Technol, vol. 2,pp. 85–90, 2020.
[2] "K. Sravani, G. Anushna, I. Maithraye, P. Chetan," Prediction of Liver Malady Using Advanced Classification Algorithms, In Machine Learning Technologies and application, Springer,Singapore, pp.39-49, 2021.
[3] "P. Rajeswari ,G. Sophia Reena" , Analysis of Liver Disorder Using Data Mining Algorithm, Global Journal of Computer Science and Technology,2010.
[4]"Ebenezer Obaloluwa Olaniyi shaman Aadnan,"Dz Liver Disease Diagnosis Base don Neural Networked , Advances in Computational Intelligence, Proceedings of the 16th International Conference on Neural Networks (NN'15), November7-9,2015.
[5]Dhamo dharan.S,Liver Disease Prediction Using Bayesian Classification,Special Issue,4th National Conference on Advanced Computing ,Applications & Technologies, ,may 2014, pageno 1-3.
[6] Bendi Venkata Ramana, Surendra. Prasad Babu. M, Venkateswarlu. N.B, A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis, International Journal of Database Management Systems (IJDMS), Vol.3, No.2,May2011
[7]A. S. Aneesh kumar and C. J. Venkates waran,"Estimating the Surveillance of Liver Disorder using Classification Algorithms" , International Journal of Computer Applications(0975 –8887), Vol. 57, no. 6, (2012) Guide Mr. Kumbhar R. S. Page 11 of 12 Student
[8]P. Porwal et al., "Indian Diabetic Retinopathy Image Dataset (IDRiD): A Database for Diabetic Retinopathy Screening Research",Data,vol. 3, no. 3, 2018.
[9]U Rajendra Acharya,Vinitha Sree Subbhuraam, Recardo Ribeiro,Ganapathy Krishna murthi,RuiTato Marinho,Joao Sanches, etal.,"Data mining framework for fatty liver disease classification in ultra sound : ahy brid feature extra action paradigm ", Medical physics, vol. 39,no. 7Part1, pp. 4255-4264,2012.
[10]Tapas Ranjan Baitharua , Subhendu Kumar Panib,Dz Analysis of Data Mining Techniques For Healthcare Decision Support System Using Liver Disorder Dataset dz, International Conference on Computational Modeling and Security,2016.
[11] L.A. Auxilia, "Accuracy prediction using machine learning techniques for indian patient liver disease," in 2018 2nd International Conference on Trends in Electronics and Informatics(ICOEI).IEEE,2018,pp.45–50.