Accuracy Prediction and Classification Using Machine Learning Techniques for Liver Disease

Tanneeru Rachana *Somu LeelaMadhav* *Mr.Rajkumar*

SRMIST *SRMIST* *SRMIST*

Student *Student* *M.tech, Asst Professor*

Abstract-

Patients with liver disease have been continuously increasing because of excessive consumption of alcohol, inhale of harmful gases, intake of contaminated food, pickles and drugs. Liver patient datasets are investigated for building classification models in order to predict liver disease. This dataset was used to evaluate prediction algorithms in an effort to reduce burden on doctors. Chronic liver disease refers to disease of the liver which lasts over a period of six months. So in that, we will take results of how much percentage patients get disease as a positive information and negative information. Using classifiers, we are processing liver disease percentage and values are shown in a confusion matrix. We proposed a various classification scheme which can effectively improve the classification performance in the situation that training dataset is available.

Keyword:liver disease,confusion matrix,ROC-AUC.

I.INTRODUCTION

The point of this task is to some degree reduce the time delay caused because of the superfluous forward and backward transporting between the healing centre and the pathology lab .l nitia lly ,the work has been done in recognizing the sicknesses like coronary illness , Parkinson's from different highlights (!). For this situation , a machine learning calculation will be prepared to foresee a liver ailment in patients. We have been given a settled number of highlights for every piece of information , and our point will be to prepare an assortment of Supervised Learning calculations on this information , so that, when another information point emerges, our best performing classifier can be utilized to classify the information point as a positive case or negative case.

The specialists can enter the patient's report as information. (6)Using the framework we are therefore looking if the patient is having a liver infection or not just by examining the accessible dataset. The relatable data is collected and the test takes place accordingly. We are catching the test report contribution from MIT and storing it in a table. We have made a table with the current dataset which has every one of the records. We have utilized confusion matrix and ROC-AUC to make a model.

II. MACHINE LEARNING TECHNIQUES

I.Logical Regression:It is mainly used for the classification of Machine Learning (6).In this only specific values or categories are allowed.It also uses independent predictors to predict the value.To find the predicted value between O and 1 we use a function called sigmoid function. Following is the graph related to the logical regression:

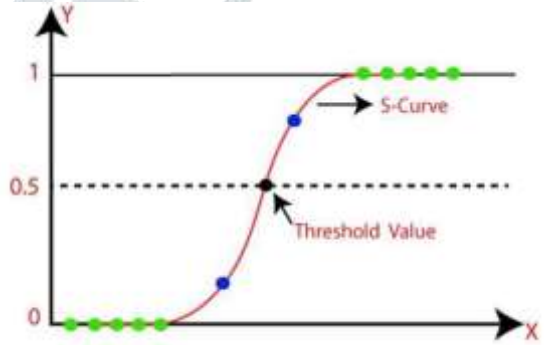

Fig no1:Logical Regression Graph

2.SVM:Support Vector Machine is a machine learning technique .SVM algorithm ,we plot each item data as a point in n-dimensional space.This model is mainly is basically a representation of variety of hyperplanes in a 3D plane.The hyperplanes are

generated by the SVM so that they can reduce the error.The basic idea of SVM is shown through the

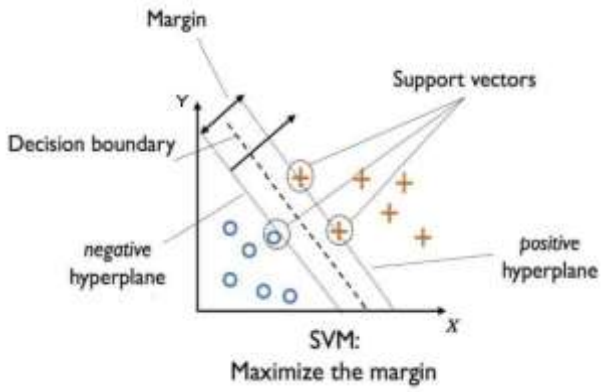

Fig no2:SVM Graph

3. **Random Forest:Random** Forest consists of a large number of individual decision trees.By using this technique it only gives an accurate result from the given data sets.This model is very complex in nature.

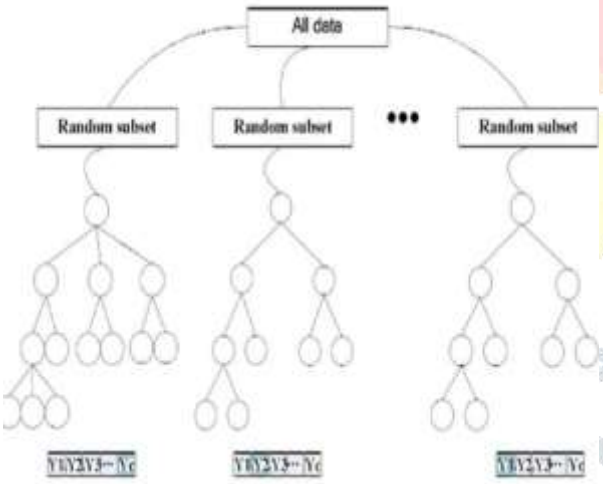

Fig no3:Random Forest

III. ARCHITECTURE OF CLASSIFICATION

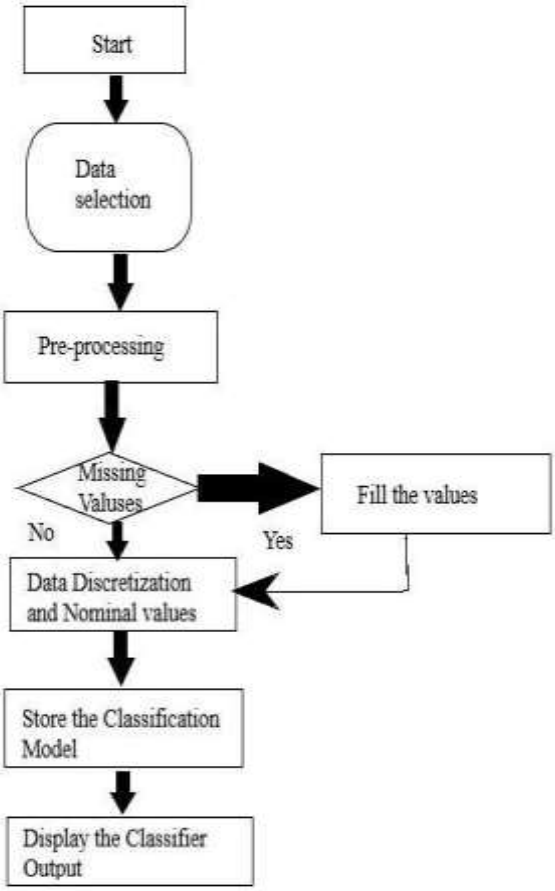

Fig no 4 :Architecture of Classification of liver Disease

This mainly prioritizes on which type of datasets are available and based on the classification is done accordingly.The relevant technique is identified and the classification of the model is stored and after that it is displayed as the output or the result.(1)(6)

IV.TABULATION

Below is the datasets which are taken from the UCI ,(3)where the information about the patients is given in detailed ,the few are as follows:

Table I.Datasets of Liver Patients

I			I
I			I
I			I
I			I

Using the method name confusion matrix ,a matrix is drawn from the clear explanation of the data.

- **Confusion Matrix-A** confusion matrix is a summary of prediction on the basis of classification problem.The key of confusion matrix is that the number of correct and incorrect predictions are summarized with count values and broken down by each class.(l)It actually gives the gives the performance accuracy of the techniques which are used.

Table 2.Confusion Matrix Prediction

	Predicted class I	

Where Class1=positive Class2=negat ive,T=true value,F =false value

For accuracy prediction ,

accuracy=TP+TN/TP+TN+FN+FP .....(1)

For precisi on,

precisio n=TP/TP+TF..... (2)

V.GRAPHS

The graphs here are mainly the data set of the liver disease patients arranged according to their internal organ enzymes level .(3)Following is the dataset of the total proteins present in a liver infected person and a non infected person,

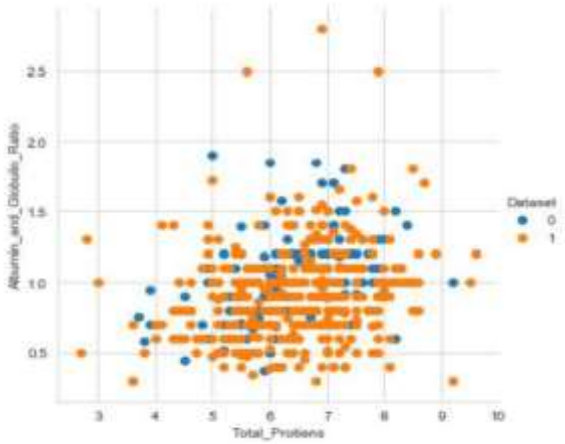

Fig no 5 :Representation of proteins from Dataset

This is the overall representation of the liver patients who are infected and also not infected.With the help of the colour used we can accurately identify the liver infected person.And also to study this graph a method named ROC-AUC graph is used for studying.

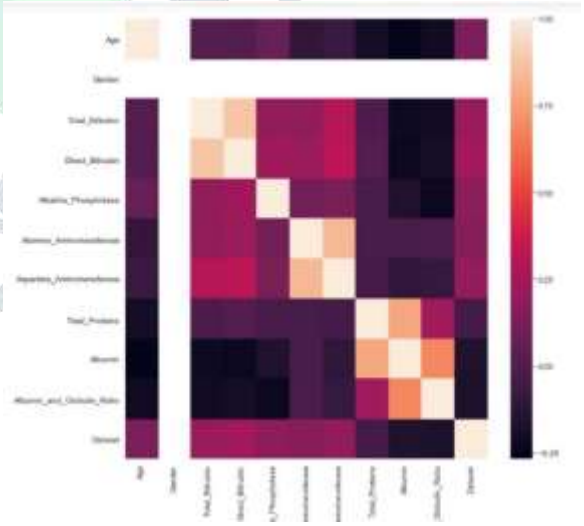

Fig no 6:Representation of all the dataset provided by

- **ROC graph-It** is also known as receiver operating characteristics graphs which are mainly used for organizing classifiers and visualizing their performance.In recent years,the use of these graphs have been increased in machine learning and data mining research.

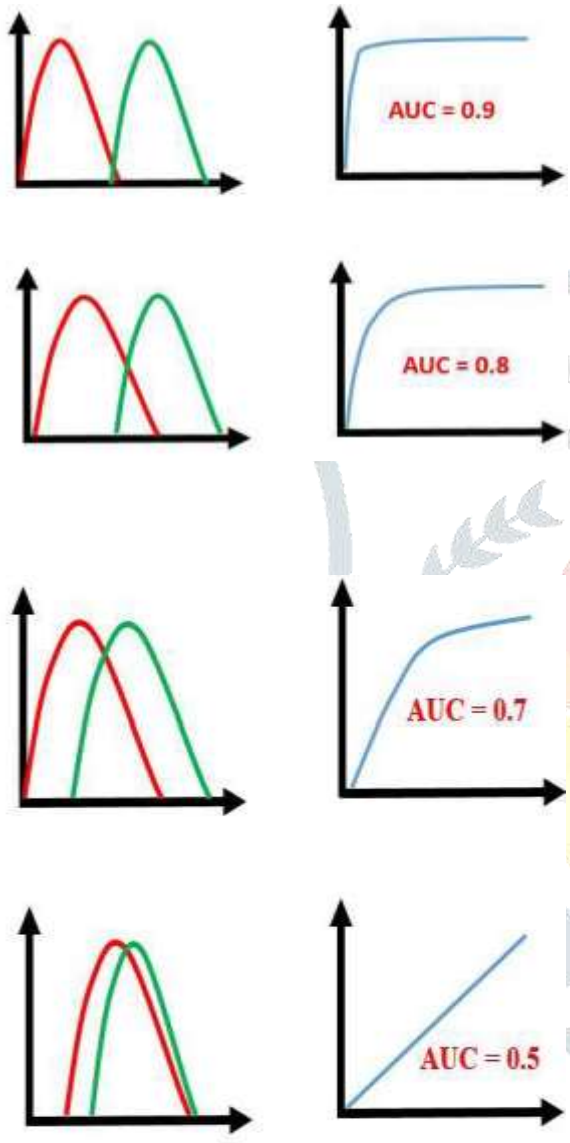

Fig no 8:The overall AUC Graph details from the techniques used

VII.CONCLUSION

This paper has discussed the important techniques which are used in machine learning for the prediction of liver disease.With the help of confusion matrix and ROC graph we get to predict the liver disease of a large dataset and the accurate information can be obtained.The graph that is used gives the detailed explanation of the liver disease patients.

V.REFERENCE

(1) machine learning confusion-matrix-machine-leam ing

(2)h ttps://doi.org /10.1016/j.cmpb.2018.12.032

(3) https://www.kaggle.com/harisyammnv/liver-d isease-prediction

(4) https://www.ncbi.nlm.nih .gov/pubmed/23963359

(5) https://www.kaggle.com/sanjames/liver-patients-a nalysis-prediction-accuracy

(6) https://blogs.oracle.com/bigdata/machine-leaming-techniques

(7) https://medium.com/greyatom/lets-leam-about-auc-roc-curve-4a94b4d88l 52

(8) Siegelman, E.S. and Rosen, M.A. Imaging of hepatic steatosis. Semin Liver Dis. 2001; 21: 71-80

(9) Tiikkainen, M., Tammine n, M., Hakkine n, **A.M.** et al. Liver-fat accumulation and insulin resistance in obese women with previous gestational diabetes. Obes Res. 2002; 10: 859-867

(l0)Kotronen, A., Vehkavaara, S., Seppala-L ind roos, A. et al. Effect of liver fat on insulin clearance. Am J Physiol Endocrinol Metab. 2007; 293: El 709-El715

(11) "Liver function tests: MedlinePlus Medical Encyclopedia". www .nlm.nih.gov. Retrieved 2015-06-20.