

LITERATURE SURVEY

1) A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis

Authors : Bendi Venkata Ramana, Prof. M.Surendra Prasad Babu, Prof. N. B. Venkateswarlu

Year : 2011

ML algorithms used : Naïve Bayes, C4.5, Backward propagation, KNN and SVM

Dataset used : AP liver dataset and UCLA liver dataset

Observation : This paper concentrates on performance of classification algorithms With different feature combinations such as SGOT, SGPT and ALP with the selected datasets. Naive Bayes, C4.5 KNN, Backward propagation and SVM has 95.07, 96.27, 96.93, 97.47, & 97.07% accuracy respectively.

Conclusion : KNN, Backward propagation and SVM are giving better results. AP dataset is better than UCLA for all the selected algorithms.

2) Liver disease prediction using machine learning

Authors : Vasan Durai, Suyan Ramesh, Dinesh Kalthireddy

Year : 2019

ML algorithms used : SVM,NB & J48

Dataset used : UCI repository

Observation : The patient data sets are analyzed for the predictability of the subject to have a liver disease based purely on a widely analyzed classification model. The J48 algorithm has better feature selection with 95.04% accuracy.

Conclusion : J48 algorithm has an accuracy rate of 95.04%.

3) Application of Machine Learning Techniques for Clinical Predictive Modeling: A Cross-Sectional Study on Nonalcoholic Fatty Liver Disease in China

Authors : Han Ma, Cheng-fu Xu, Zhe Shen, Chao-hui Yu, and You-ming Li

Year : 2018

ML algorithms used : Using 11 classification algorithms

Dataset used : First Affiliated Hospital, Zhejiang University China, College of medicine

Observation : Machine learning techniques were introduced to evaluate the optimal predictive clinical model of NAFLD. Bayesian network accuracy 83%.

Conclusion : The Bayesian network has better performance than other algorithms.

4) Diagnosis of liver diseases using machine learning

Authors : Sumedh Sontakke, Jay Lohokare, Reshul Dani

Year : 2017

ML algorithms used : SVM & Backpropagation

Dataset used : (UCI) Machine Learning Repository

Observation : This paper aims to improve diagnosis of liver diseases by exploring 2 methods of identification-patient parameters and genome expression. SVM(accuracy 71%) & Backpropagation(accuracy 73.2%)

Conclusion : More accuracy result in Back propagation

5) A Survey on machine learning techniques for the diagnosis of liver disease

Author : Golmei Shaheamlung , Harshpreet Kaur , Mandeep Kaur

Year : 2020

Observation : ML algorithms used in this paper are Supervised, Unsupervised and Semi- Supervised, Reinforcement Learning for diagnosis of liver disease such as SVM, KNN, K-Mean clustering, neural network, Decision tree etc and give difference accuracy, precision, sensitivity.

Conclusion : Machine learning algorithm such as Decision tree , J48 and ANN provide better accuracy on detection and prediction of liver disease.

6) Prediction of Liver Fibrosis stages by Machine Learning model: A Decision Tree Approach

Author : Heba Ayeldeen

Year : 2015

ML algorithms used : Decision Tree

Software used : MYSQL 5.2 Community Server 5.5.27 for storing data.

Observation : This paper uses the decision tree classifier to predict the liver fibrosis. The interpretation of liver fibrosis stages is mainly classified into five different stages based on ANOVO:

- F0 =no fibrosis
- F1 =portal fibrosis without septa
- F2 =fibrosis with few septa
- F3 =fibrosis with numerous septa
- F4 =cirrhosis

Ten variables are used for prediction further divided into two classes Routine function test and Serum Tests.

Conclusion : Results obtained on decision tree classifier accuracy is 93.7%.

7) Liver Classification Using Modified Rotation Forest

Author : Bendi Venkata Ramana , M.Surendra Prasad Babu

Year : 2012

Dataset : UCI and India

ML algorithms used : Modified Rotation Forest

Software used : Weka© Data Mining - open source machine learning software

Observation : Classification technique contains combination of selected classification algorithm and feature selection technique. The category of classification algorithms are Tree based, Statistical based, Neural Networks based, Rule based and Lazy learners.

J48 and simple cart classification algorithms - Tree based algorithms

Naïve bayes and Bayes net classification algorithms - Statistical based algorithms.

MLP and SMO classification algorithms - Multi layer perception based algorithms.

IBK and KStar classification algorithms - Lazy learners.

PART and Zero classification algorithms - Rule based algorithms .

The accuracy of a classifier is the percentage of the test set tuples that are correctly classified by the classifier.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

Conclusion : The results shows that multi layer perception classification algorithm with random subset gives highest accuracy that is 74.7826 for the UCI liver data set nearest neighbor with CFS gives highest accuracy that is 73.0703 for the INDIA liver data set.

8) Prediction of different types of liver diseases using rule based classification model

Author : Kumar, Yugal

Year : 2013

ML algorithms used: DT , SVM , NB , ANN

Observation : They have developed a Rule Base Classification Model (RBCM) to predict different types of liver diseases

Support Vector Machine (SVM), Rule Induction (RI), Decision Tree (DT), Naive Bayes (NB) and Artificial Neural Network (ANN) data mining techniques with K-cross fold technique are used. The proposed model with decision tree (DT) technique provides the better result among all techniques (RI, SVM, ANN and NB) with all parameters (Accuracy 98.46%, Sensitivity 95.7%, Specificity 95.28% and Kappa 0.983)

SVM exhibits poor performance (Accuracy 82.33%, Sensitivity 68.03%, Specificity 91.28% and Kappa 0.801). It is also found that the best performance of the model without rules (RI, Accuracy 82.68%, Sensitivity 86.34%, Specificity 90.51% and Kappa 0.619) is almost similar to the worst performance of the rule based classification model (SVM, Accuracy 82.33%, Sensitivity 68.03%, Specificity 91.28% and Kappa 0.801 and the accuracy of chi square test is 76.67%.

Conclusion: Rule based classification with DT Algorithm has better accuracy.

Team :

- Ashok Kumar R – 2019503008
- Dharshan R – 2019503515
- Saktheeswaran K – 2019503041
- Santhosh S - 2019503044