Ideation Phase

Literature Survey and Information Gathering

DATE	1TH SEPTEMBER
TEAM ID	PNT2022TMID24928
PROJECT NAME	Statistical Machine Learning Approaches to Liver Disease Prediction
MAXIMUM MARK	

LITERATURE SURVEY

Bendi et al. [1] authors used two different input dataset and evaluate that the AP datasets has better than UCLA dataset for all the different selected algorithms. Based on performance on their classification KNN, Backward propagation and SVM are giving better results. The AP data set is better than UCLA for the entire selected algorithm. And found out Naïve Bayes, C4.5, KNN, Backward propagation and SVM has 95.07, 96.27, 96.93, 97.47, & 97.07% accuracy respectively.

Bendi et al. [2] proposed a paper based on Modified Rotation Forest, used two dataset as an input UCI liver dataset and Indian liver dataset. And results show that MLP algorithm with random subset gives better accuracy of 94.78% for UCI dataset than CFS achieved accuracy of 73.07% for Indian liver dataset

Yugal Kuma & G. Sahoo [3] proposed a paper based on different classification technique and used north east area of Andhra Pradesh (India) liver dataset. And the results shows that Decision tree(DT) algorithm has better than other algorithm and provide accuracy of 98.46%.

S.Dhamodharan [4] proposed a paper based on two classification technique naïve Bayes and FT tree and used WEKA (Waikato Environment for Knowledge and Analysis) dataset. Naïve Bayes is 75.54% accuracy and FT Tree is 72.6624% accuracy and concluded Naïve Bayes gas better algorithm compare to other algorithms

Heba Ayeldeen et al. [5] propose a paper for prediction of liver fibrosis stages using decision tree technique and used Cario university data set and result shows that decision tree classifier accuracy is 93.7%.

D.Sindhuja & R. Jemina Priyadarsini [6] survey a paper for classification of liver disease. In this survey different classification techniques of data mining are study and used dataset of dataset of AP liver has better than Dataset of UCLA, and concluded C4.5 achieved better results than other algorithms.

Somaya Hashem et al. [8] presented a paper for diagnosis of liver disease. In this paper they used two algorithms, SVM & Backpropagation and used UCI machine repository dataset. And concluded SVM has accuracy 71% better result than Backpropagation accuracy 73.2%.

Han Ma et al. [9] in this paper 11 different classification are evaluated and Demonstrated in China Zhejiang University, College of medicine and concluded Bayesian network accuracy of 83%, specificity 83%, sensitivity of 0.878 and F-measure of 0.655

Joel Jacob et al. [10] proposed a paper to diagnosis of liver disease by using three different algorithms, Logistic regression, K-NN, SVM, and ANN and used Indian Liver Patient Dataset comprised of 10 different attributes of 583 patients. And concluded Logistic regression, K-NN, SVM,& ANN has 73.23, 72.05, 75.04 & 92.8% accuracy respectively.

Sivakumar D et al. [11] proposed a paper for prediction of chronic liver disease by using two different techniques K-means and C4.5. UCI repository.

Mehtaj Banu H [12] in this paper authors study different machine learning technique, Supervised, unsupervised & reinforcement and also analysis UCI dataset database and concluded that KNN and SVM improved better performance and exactness of liver disease prediction.

Comparison table of various machine learning technique used to detect liver disease based on performance.

Methods	Accuracy	Specification	Sensitivity
Decision Tree	98	95	95
Bayesian Network	83	87	67
SVM	71		
Naive Bayes	55		
KNN	70		