LITERATURE SURVEY

Statistical Machine Learning Approaches to Liver Disease Prediction

S. NO	NAME OF PAPER	JOURNAL PUBLISHED	FEATURES	TECHNIQUES USED	OBSERVATION
1	Liver Disease Prediction System using Machine Learning Techniques, Rakshith D B ,Mrigank Srivastava,Ashwani Kumar,Gururaj S P	IJERT	age, gender, total Bilirubin, direct Bilirubin, total proteins, albumin, A/G ratio, SGPT, SGOT and Alkphos	SVM,ANN ,KNN,Naive Bayes	Details including age, gender, total Bilirubin, direct Bilirubin, total proteins, albumin, A/G ratio, SGPT, SGOT and Alkphos are taken. Values of last eight parameters mentioned here, can be known by blood test report of the user. After taking these inputs from the user, the system compares the data input with the training dataset of most accurate model and then predicts the result accordingly as risk or no risk.SVM gives the best accuracy of all(100%)
2	A Prediction Model of Detecting Liver Diseases in Patients using Logistic Regression of Machine Learning PSM Keerthana,	INTERNATIONAL CONFERENCE ON INNOVATIVE COMPUTING AND COMMUNICATION (ICICC-2020)	It consists 10 variables of which 1 is a dependent variable, and the remaining 9 are independent variables used for predicting	Logistic Regression Algorithm	Logistic Regression comes under the supervised machine learning algorithms. This is used in binary classification. Logistic Regression algorithm takes a data set line and then calculates the probability for

	Nimish Phalinkar, Riya Mehere, Koppula Bhanu Prakash Reddy, Nidhi Lal		whether the person is affected by liver disease or not.		classifying the sample among the two classes with respect to the threshold value. The final accuracy score obtained in this model is 0.859649.
3	Prediction of Liver Disease Using Machine Learning Algorithm and Genetic Algorithm B.Poonguzharselvi , Mohammad Mahaboob Ali Ashraf , Vadlamani V S S Subhash , S.Karunakaran	Annals of R.S.C.B		Adam optimizer, Adagrad, RMSProp,SGD, Backpropagation, Genetic algorithm	From the comparison of various algorithms, we can clearly see that Random Forest performs the best followed by Voting Classifier and Adaboost among the machine learning models with accuracies of 84%, 84 %,79% respectively. Neural Nets gave a validation accuracy of 74 percent. The Genetic Algorithm improved performance of Adaboost by 3 percent, decision trees by 4 percent and SVCby 19 percent.
4	Machine Learning Techniques in Analysis and Prediction of Liver Disease Dr. Dattatreya P Mankame , Harshitha R , Navya N C , Nitin Ravichander	IJIRT	Considering Indian Liver Dataset, we can see a very high linear relationship between Total and Direct Bilirubin and by considering	Logistic Regression, Decision Tree, KNN, Random Forest Classifier, SVM	This solution gives a comprehensive analysis of “Indian Liver Patient Records' ' dataset with Liver patient and Not Liver patient as classification is performed and this relies upon various machine learning algorithms which provides high accuracy and consumes very less

			this linear relationship, Direct Bilirubin can be opted to be dropped, But by as per medical analysis Direct Bilirubin constitutes to almost 10% of the Total Bilirubin and this 10% may prove crucial in obtaining higher accuracy for the model, thus none of the features are removed.		time for entire processing. The process includes data analysis, data pre-processing which includes filling of missing values with mean, label encoding, identifying duplicate value, outlier detection and resampling to improve the performance. Accuracy is effectively utilized to analyze the performance of various classification algorithms. Thus, we can conclude that SVM classifier proved its worthiness in prediction of liver patients by achieving high accuracy amongst the other classifiers.
5	A Comparative Study On Liver Disease Prediction Using Supervised Machine Learning Algorithms A.K.M Sazzadur Rahman, F. M. Javed Mehedi Shamrat, Zarrin Tasnim,	IJSTR	Age, Gender , Total Bilirubin, Direct Bilirubin ,Alkphos,Alkaline Phosphotase , Alamine Aminotransferase , Asparatate Aminotransferase ,Total Proteins	Logistic Regression, Decision Tree, KNN, Random Forest Classifier, SVM	In this experiment, we considered different analyses to examine the six-machine learning classifier for the classification of liver disease dataset. In terms of accuracy, LR achieved the highest accuracy of 75% and NB achieved the worst performance 53%. With respect to precision, LR

	Joy Roy, Syed Akhter Hossain		,Albumin ,Albumin and Globulin Ratio ,Selector field		achieved the highest score 91% and NB performs worst 36%. When considering the sensitivity, SVM achieved the highest value 88% and KNN obtained the worst 76%. Logistics Regression was also the best performer in terms of f1 measure 83% and NB obtained the worst performance 53%. When considering specificity DT achieved the highest value 48% and LR the lowest 47%. According to compare these measurement criteria LR classification technique is more effective than the other classifiers for predicting chronic liver disease.