A Critical study of Adaptive Boosting and Stacking Ensemble Approaches for Liver Disease Prediction

Sivala Vishnu Murty[1] Dr. R.Kiran Kumar[2]
[1]*Aditya Institute of Technology and Management, Tekkali. Andhra Pradesh.532201 India*
[2]*Krishna University, Machilipatnam, Andhra Pradesh.521001, India*

ABSTRACT - Machine Learning algorithms are often used on liver datasets to predict the diseases automatically so as to help the doctors to predict the disease accurately and quickly. Liver disease patients are increasing day by day due to change in life style, food habits, and culture. Recent research studies have shown that traditional classification approaches do not give accurate results and often leading to over fitting models .Hence, there is a need for using modern ensemble classification approaches for liver disease prediction for more accurate prediction of liver diseases. In this backdrop, we collected Clinical data (Liver Function Test data) from north coastal districts of Andhra Pradesh and experimented with modern ensemble techniques like adaptive Boosting and stacking and evaluated their performances for liver disease prediction on the liver data sets collected north coastal districts of Andhra Pradesh, India

Keywords--- Data mining, Liver Disease, Machine Learning, boosting, stacking, Classification

I. INTRODUCTION

Liver disease is the 14th chief cause of deaths in the world and could be the 12th foremost cause of deaths in the world by 2020 Deaths due to Liver Diseases in India is increasing at an alarming rate and ranks 63rd in the world as revealed by World Health Organization (WHO) 2017. Every year about ten lakh patients are diagnosed with liver cirrhosis. The next major health concern for the country is liver. Liver disease is the tenth most regular cause of deaths in India as per the world health organization. Traditional classification methods are widely used to diagnose liver diseases but the prediction accuracies are not satisfactory as these models mostly over fit and do not generalize well. Hence, there is need for novel ensemble classification approaches for accurate and early prediction of liver diseases. In this work, we experimented adaptive boosting and stacking ensemble techniques on North Coastal Andhra Pradesh Liver patient Data sets in the process of identifying the best approach to diagnose the liver disease.

II. LITERATURE REVIEW

Maruf Pasha et.al [1] applied metal classifiers on Indian liver patients dataset (ILPD) in order to evaluate the performance of these algorithms to classify the liver patients as diseased or not. It was observed that Adaboost performed with an accuracy of 70.3259% and took 0.1 seconds for building the model where as Logitboost gave an accuracy of 70.4974% and took 0.1 seconds for construction of the model. Sanjay Kumar et.al [2] applied Adaboost with c5.0 as base classifier for liver disease prediction and their model exhibited an accuracy of 75.19%.But when c5.0 was applied gave 71.43% accuracy on its own. The dataset used contains 416 records on liver disorder and 167 records on non liver disorder patients. The dataset was collected and compiled from North Eastern Areas of India.Dutta et.al [3] proposed a voting ensemble method that considers influential factors responsible for liver disease. Their predictive model aims to enhance forecasting reports with respect to other peer intelligent model. It was observed that Adaboost Classifier gave an accuracy of 72.54%. Gradient Boost Classifier exhibited an accuracy of 74.09%. Their voting ensemble method reached an accuracy of 77.2% which is quite promising towards early liver disease prediction. Maria et.al [4] applied ensemble classification models for liver disease prediction and observed that Gradient Boosting gave 0.66 accuracy , adaptive boosting performed with 0.68 accuracy ,Xgboost gave 70%accurac,,light GBM gave 70%, accuracy and stacking exhibited an accuracy of 83%.Raheem et.al [5] applied three classification algorithms; Naïve Bayes, K-nearest neighbour and decision trees, their bagged and boosted versions, then the algorithms were combined together by ensemble methods of stacking and voting on liver diseases dataset using 10-fold cross validation. The boosted version of Naïve Bayes showed an increase in classification accuracy when compared with Naïve Bayes. The stacking and voting has a reduced root mean squared error as compared to the other algorithms, while it was observed that C4.5 decision tree algorithm gave the best classification accuracy of all the algorithms used.Abdar et.al [6] applied boosting technique as an appropriate solution to achieve the better accuracy .They constructed boosted B-C5.0, B-CHAID and B-CART methods. The 10-fold cross validation was utilized with C5.0 algorithm to improve the performance of the algorithm. It was observed that B-C5.0 method had better performance than B-CHAID and B-CART methods. The best values of metrics for B-C5.0 concerning specificity, sensitivity, precision, FPR, FNR, F1, and accuracy were 87.17, 94.16, 96.26, 12.83, 5.84, 95.20, and 92.61% respectively.Hassoon et.al [7] proposed a method which optimizes the rules released from Boosted C5.0 classification method with the Genetic Algorithm (GA), to increase the diagnosis time and accuracy. So instead of using an evolutionary algorithm for producing rules, the genetic algorithm is used for improving and reducing rules of another algorithm. It was observed that their proposed approach gave better performance and throughput in comparison with other work in the field and the accuracy improved from 81% in to 93%.Vijayarani et.al [8] used classification algorithms namely Naïve bayes and Support Vector Machine (SVM) for liver disease prediction.

Comparisons of these algorithms are done and it is based on the performance factors classification accuracy and execution time. From the experimental results, they concluded that the SVM classifier is considered as a best algorithm because of its highest classification accuracy of 79.66 % . On the other hand, while comparing the execution time, the Naïve Bayes classifier needs minimum execution time.

3.1 Adaptive Boosting

Adaboost combines multiple weak learners into a single strong learner. The weak learners in AdaBoost are decision trees with a single split, called decision stumps. When AdaBoost creates its first decision stump, all observations are weighted equally. To correct the previous error, the observations that were incorrectly classified now carry more weight than the observations that were correctly classified.

III. MEHODOLOGY

The first model is trained by randomly picking some samples from the data set ,where every sample has equal chance to participate in training. Every model is tested on all the samples and weight of the wrongly classified samples is updated so that they are picked for the training of the next model. It builds a number of models in a sequential manner. When a test sample is to be predicted, then the predictions of majority of the models are considered and that will be the final prediction.

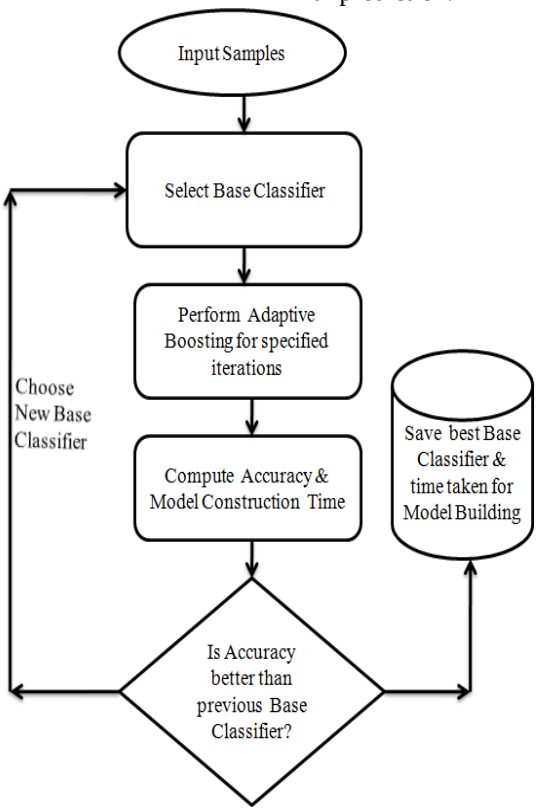

Figure 3.1 :Proposed Adabost model

3.2 Stacking Ensemble

Stacking also called as Stacked Generalization is a special prototype. The idea of stacking is to investigate a space of different models for the same problem. The idea is that you can attack a learning problem with different types of models which are competent to learn some part of the problem, but not the whole space of the problem. So, we can build several different learners and use them to build an intermediate prediction, one prediction for each learned model. Then we add a new model which learns from the intermediate predictions the same target. This final model is said to be stacked on the top of the others. Thus, you might improve your overall performance, and often you end up with a model which is better than any individual intermediate model.

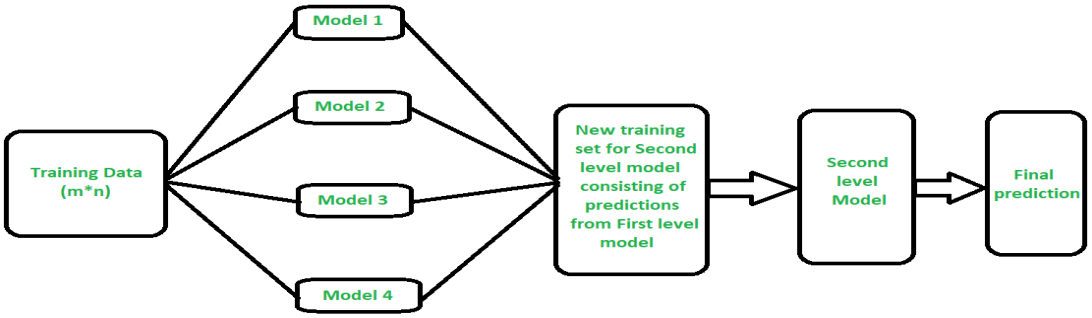

Fig 3.2 Stacking Ensemble

Algorithm

Step 1: Split the training data into K-folds like K-fold cross-validation.
Step 2: A base model is fitted on the K-1 parts and predictions are made for Kth part.
Step 3: Do for each part of the training data.
Step 4: Then the base model is fitted on the full train data set to work out its performance on the test set.
Step 5: Repeat the last 3 steps for other base models.
Step 6: Predictions from the train set are used as features for the second level model.
Step 7: Second level model is used to make a prediction on the test set.

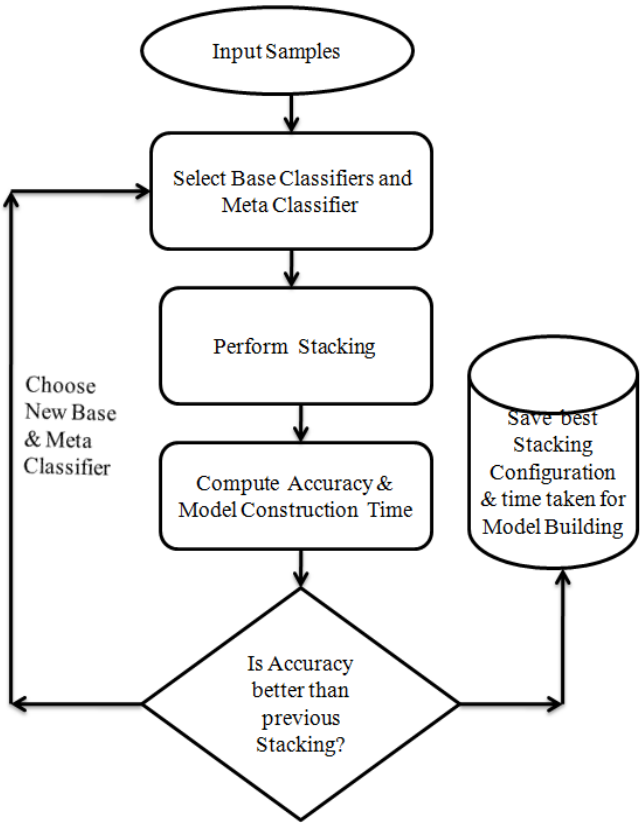

Figure 3.3 :Propsed Stacking model for liver disease prediction

IV. RESULTS AND DISCUSSIONS

This data set contains totally 882 instances; out of which 403(45.7%) instances are of class 0(non-diseased) and 479 (54.3%) instances are of class 1(diseased). Outcome is a class label to split the samples into two classes (liver patient or not).

4.1 performance of Adaboost

The following table shows the Detailed Accuracy by Class for different Base Classifiers

Table 4.1 Detailed Accuracy by Class for different Base Classifiers

	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
ADTree	0.915	0.042	0.949	0.915	0.932	NO
	0.958	0.085	0.929	0.958	0.943	YES
J48	0.939	0	1	0.939	0.969	NO
	1	0.061	0.95	1	0.974	YES
Naïve Bayes	0.976	0.442	0.656	0.976	0.784	NO
	0.558	0.024	0.964	0.558	0.707	YES
RBF Network	0.878	0.116	0.867	0.878	0.873	NO
	0.884	0.122	0.894	0.884	0.889	YES
SVM	0.683	0.137	0.812	0.683	0.742	NO
	0.863	0.317	0.759	0.863	0.808	YES

The following table 4.2 below shows the confusion matrices of Adaboost with different base classifiers.

Table 4.2 Confusion Matrix of Different Base classifiers

	Base Classifier	Class		Total	Base Classifier	Class		Total	Base Classifier	Class		Total
Class	ADTree	1	2		C4.5	1	2		Naïve Bayes	1	2	
1		75	7	82		77	5	82		80	02	82
2		4	91	95		0	95	95		42	53	95
Total		79	98	177		77	100	177		122	55	177
1	RBF Network	72	10	82	SVM	56	26	82				
2		11	84	95		13	82	95				
Total		83	94	177		69	107	177				

The following table 4.3 below shows accuracy and model construction time of Adaboost with different base classifiers.

Table 4.3 Performance e of AdaBoost with different Base Classifiers

S.No	Base Classifier	Accuracy	Time taken for model construction(Seconds)	No. of iterations
01	ADTree	93.7%	2.68	50
02	C4.5	97.1%	0.98	50
03	Naïve Bayes	75.1%	0.27	50
04	RBF Network	88.1%	1.17	50
05	SVM	77.9%	0.67	50

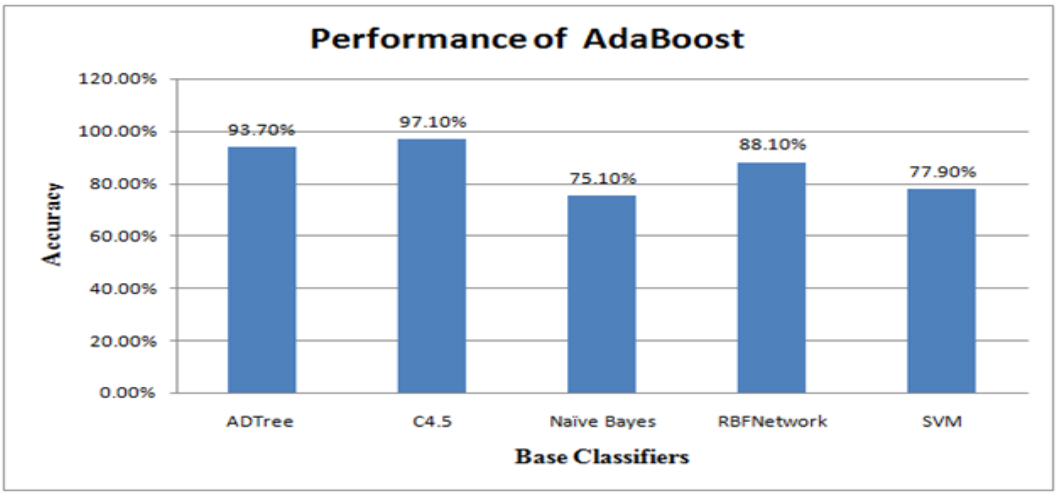

Fig 4.1 Accuracy of AdaBoost with different base classifiers

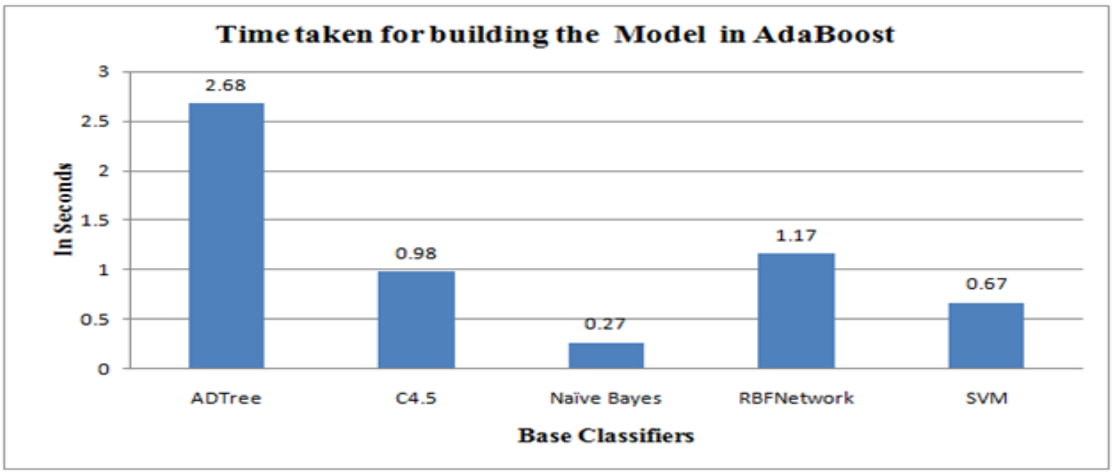

Fig 4.2 Time taken to build model by AdaBoost with different base classifiers

4.2 Performance of Stacking Ensemble

The following table shows the confusion matrix of stacking for different combinations of base and meta classifiers**.**

Table 4.4 . Confusion matrices of Stacking with different Base and Meta Classifiers

	Base & Meta Classifier	**Class**		**Total**	**Base & Meta Classifier**	**Class**		**Total**	**Base & Meta Classifier**	**Class**		**Total**
Class	ADTree(Base) C4.5(Base) Naïve Bayes(Meta)	1	2		Naïve Bayes(Base) Bayes Network(Base) SVM(Meta)	1	2		Naïve Bayes(Base) Bayes Network(Base)) RBFN(Meta)	1	2	
1		74	8	82		76	6	82		72	10	82
2		6	89	95		5	90	95		2	93	95
	Total	80	97	177	Total	81	96	177	Total	74	103	177

Table 4.5 Detailed Accuracy by Class of Stacking for different Base and Meta Classifiers

Base and Meta Classifiers	**TP Rate**	**FP Rate**	**Precision**	**Recall**	**F-Measure**	**Class**
ADTree(Base) C4.5(Base) Naïve Bayes(Meta**)**	0.902	0.063	0.925	0.902	0.914	NO
	0.937	0.098	0.918	0.937	0.927	YES
Naïve Bayes(Base) Bayes Network(Base) SVM(Meta)	0.927	0.053	0.938	0.927	0.933	NO
	0.947	0.073	0.938	0.947	0.942	YES
Naïve Bayes(Base) Bayes Network(Base)) RBFN(Meta)	0.878	0.021	0.973	0.878	0.923	NO
	0.979	0.122	0.903	0.979	0.939	YES

The below table shows the accuracy and model building times of stacking with different combinations of base and meta classifiers.

Table 4.6 Performance of stacking with different Base and Meta Classifiers

Base Classifiers	**Meta Classifiers**	**Accuracy (%)**	**Time Taken to Build The Model(In Seconds)**
ADTree C4.5	Naïve Bayes	92.09	1.19
Naïve Bayes Bayes Network	SVM	93.78	0.17
Naïve Bayes Bayes Network	RBFN	93.22	0.19

V. CONCLUSION

After a thorough examination of the results, it is observed that the C4.5 algorithm is found to be the best when used as base classifier in adaptive boosting as it performed with 97.10 % accuracy and it took 0.98 seconds for construction of the model. The stacking ensemble, performed with 93.78% accuracy, when naïve bayes and Bayes network are used as base learners and SVM as Meta classifier and took only 0.17 seconds for model construction. Hence we conclude that adaptive boosting gives more accurate results compared to stacking ensemble for liver disease prediction although stacking ensemble techniques take lesser model construction time compared to adaptive boosting.

VI. ACKNOWLEDGEMENTS

We whole heartedly thank Dr. T Venu Gopal M.D., Liver Disease Expert, Amrutha Hospitals, Srikakulam, and Andhra Pradesh, India for helping us to label the samples into diseased and healthy. We are also thanking AITAM Engineering College, Tekkali, management and JNTUK, Kakinada for providing necessary facilities for this work.

VII. REFERENCES

[1]. Pasha, M., & Fatima, M. (2017). Comparative Analysis of Meta Learning Algorithms for Liver Disease Detection. *JSW*, *12*(12), 923-933.

[2]. Kumar, S., & Katyal, S. (2018, July). Effective analysis and diagnosis of liver disorder by data mining. In *2018 international conference on inventive research in computing applications (ICIRCA)* (pp. 1047-1051). IEEE.

[3]. Dutta, S., & Bandyopadhyay, S. K. (2020). Automatic Detection of Liver Disease Using Voting Ensemble Method.

[4]. Kuzhippallil, M. A., Joseph, C., & Kannan, A. (2020, March). Comparative Analysis of Machine Learning Techniques for Indian Liver Disease Patients. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 778-782). IEEE.

[5]. RAHEEM, I. O., BABATUNDE, A. O., ARO, T. O., & MUSA, A. Y. (2017). EFFECT OF ENSEMBLES ON SELECTED DATA MINING TECHNIQUES IN CLASSIFICATION OF LIVER DISEASES. *University of Pitesti Scientific Bulletin Series: Electronics and Computer Science*, *17*(2), 27-36.

[6]. Abdar, M., Yen, N. Y., & Hung, J. C. S. (2018). Improving the diagnosis of liver disease using multilayer perceptron neural network and boosted decision trees. *Journal of Medical and Biological Engineering*, *38*(6), 953-965.

[7]. Hassoon, M., Kouhi, M. S., Zomorodi-Moghadam, M., & Abdar, M. (2017, September). Rule optimization of boosted c5. 0 classification using genetic algorithm for liver disease prediction. In *2017 international conference on computer and applications (icca)* (pp. 299-305). IEEE.

[8]. Vijayarani, S., & Dhayanand, S. (2015). Liver disease prediction using SVM and Naïve Bayes algorithms. *International Journal of Science, Engineering and Technology Research (IJSETR)*, *4*(4), 816-820.