

A Function, Bayes, Meta And Tree Classifier Perspective Based On Ilpd Dataset

Prasun Chakrabarti, Manish Tiwari, Aditya Maheshwari, Tulika Chakrabarti, Sibabrata Mukhopadhyay

Abstract: Several discovered research findings entail that for the liver cancer detection, supervised machine learning approaches play a pivotal role. In this paper the classifiers performance parameters such as accuracy and precision are used for analysis purpose in the light of Function, Bayes, Meta and Tree Classifiers.

Keywords: accuracy, precision, function, Bayes, Meta, tree, classifiers

1. INTRODUCTION

Cancer is the fast-growing disease in the world. It came from the uneven life style and chemical contained food are cause of liver cancer. Machine Learning algorithms are used to prognosis of liver cancer. The algorithms are applied on the dataset for better accuracy [1-2]. These are used for early detection of cancer by which algorithms help the society to reduce the death rate in the world. The applications of artificial neural networks in liver cancer diagnosis [3-6] is a great research challenge in modern era.

2. RESEARCH METHODOLOGY

In this work, Indian Liver Patient Dataset is considered and then supervised filter has been applied. Subsequently the sampling is used on the ILPD dataset. Model is built using class data in binary representation and the accuracy, precision, recall and error of the various classifiers have been computed. The Multilayer perceptron and LibSVM classification algorithms are also used for evaluation purpose in context to each classifier.

3. RESULT ANALYSIS BASED ON INDIAN LIVER PATIENT DISORDER (ILPD) DATASET

In this experimental purpose WEKA tool is applied on the ILPD dataset and the 10 cross validation has been used.

- Aditya Maheshwari *, Project Lead (New Initiative and Research), Techno India NJR Institute of Technology Udaipur-313003, Rajasthan, India. Email: aaditya@technonjr.org
- Prasun Chakrabarti, Executive Dean (Research and International Linkage) and Institute Distinguished Senior Chair Professor in Department of Computer Science and Engineering, Techno India NJR Institute of Technology Udaipur-313003, Rajasthan, India. Email: drprasun.cse@gmail.com
- Manish Tiwari, Associate Professor Department of Computer Science and Engineering, Geetanjali Institute of Technical Studies Udaipur-313022, Rajasthan, India. Email: immanishtiwari@gmail.com
- Tulika Chakrabarti, Department of Chemistry, Sir Padampat Singhania University, Udaipur – 313601 Email: tulika.chakrabarti20@gmail.com
- Sibabrata Mukhopadhyay, Emeritus Scientist G and Former Deputy Director, Indian Institute of Chemical Biology (A Unit of CSIR, GOI). Email : sbm.chem@gmail.com

3.1 Function Classifiers

Multilayer perceptron is a part of function classifiers. Its properties are discussed previous such as auto build, batch size, hidden layer, momentum, nominal to binary filter, normalize attributes etc. 520 value find out correctly and accuracy is 89.19% and 63 cases are not correct (incorrect 10.81%), Precision of this algorithm is 92.1% .

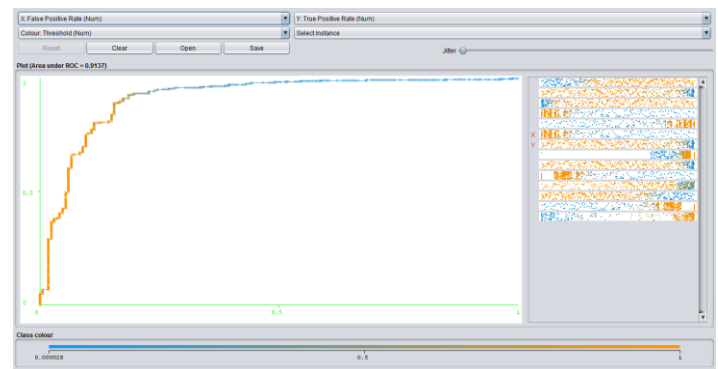


Figure 1. ROC curve for Multilayer perceptron value 0.9137.

LibSVM algorithm is a part of the function classifier. Its properties are mentioned already e.g. epsilon, type of kernel, normalized dataset, do not check capabilities, do not replace, shrinking heuristic, degree of kernel etc. 568 instances classified from total 583 and got accuracy(97.43%). Time taken by building model is 0.08 seconds. Precision obtained from this 96.5%.

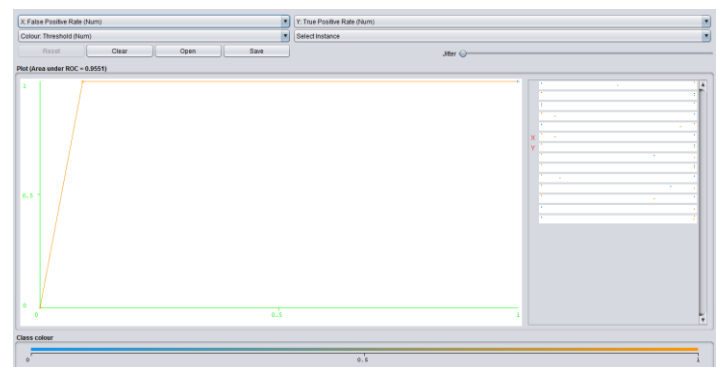


Figure 2. ROC for LibSVM value 0.9551.

3.2 Bayes Classifiers

Bayes net algorithm is a part of Bayes classifiers. It defines the property at describes earlier e.g. sample size, random seed, invert selection, no replacement simple estimator, batch size, etc. and 10 folds cross validations uses for testing purpose. Time taken to build model is 0.05 second. Using WEKA tool it gives 83.3619% accuracy and Precision as 91.0%.

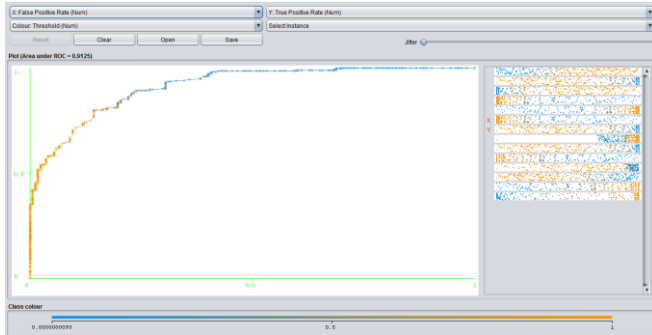


Figure 3. ROC for Bayes net value 0.9125.

Naïve Bayes algorithm is come under the category of Bayes classifier. Properties are stated earlier such as batch size, kernel estimator, supervised discretization etc. There are 342 correct instances , accuracy is 58.6621% and Precision is 97.3%.

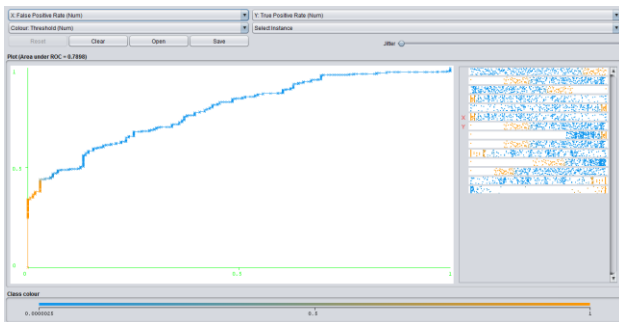


Figure 4. ROC for Naïve Bayes value is 0.7898.

3.3 Meta Classifiers

Random Committee algorithm is a part of Meta classification algorithm using the properties such as cross validations, batch size, number of iterations, number of execution slots, etc. Time taken building model is 0.05 seconds. It gave highest accuracy (98.11%) and precision obtained is 99.0%.

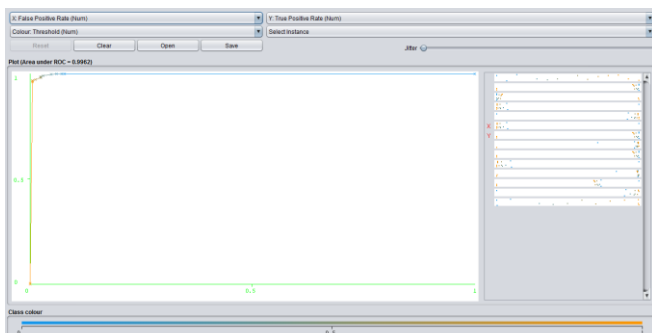


Figure 5. ROC for Random Committee value is 0.9962.

Bagging classifier is a part of Meta classifier using the properties as already mentioned e.g. do not check capabilities, bag complexity statistics, number of iterations, and so forth. It's accuracy is 95.71%. Time taken to build model is 0.08 seconds. 96.0% Precision has been achieved by this algorithm.

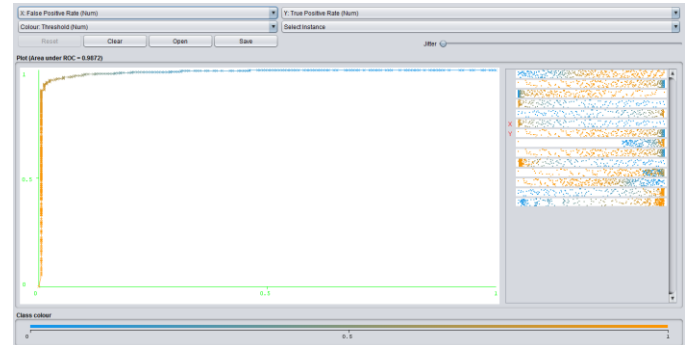


Figure 6. ROC for Bagging value is 0.9872.

Randomizable filtered algorithm's properties are seed value, batch size, random projection filter and etc. 566 instances are correct out of 583 instances. It gave 97.08% accuracy and precision as 99.3%.

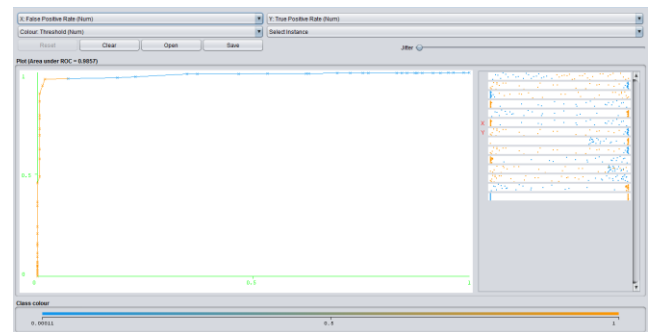


Figure 7. ROC for Randomizable filter value is 0.9857.

Random subspace algorithm is a part of meta classifier. Properties are set as described earlier e.g. seed value, cross validations, batch size, subspace size, number of iterations etc. It gave accuracy (97.77 %) and precision as 97.4%.

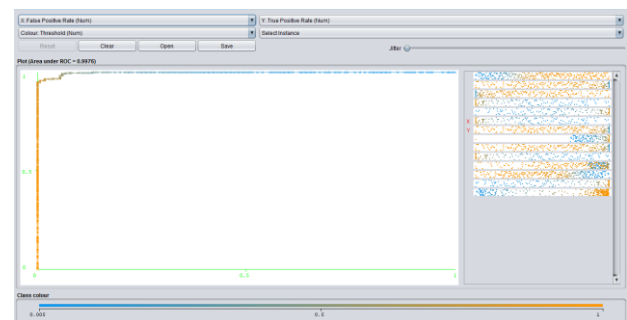


Figure 8. ROC for Random subspace value is 0.9976.

3.4. Tree Classifier

REPTree algorithm is based on tree generation classifier. Properties already defined such as batch size, initial count,

max depth, min variance, seed value etc. model taken time is 0.03 seconds. It correctly classified 549 instances of the dataset (covering 94.1681%) and incorrectly classified 34 instances (covering 5.8319%) and 94.8% precision has been obtained.

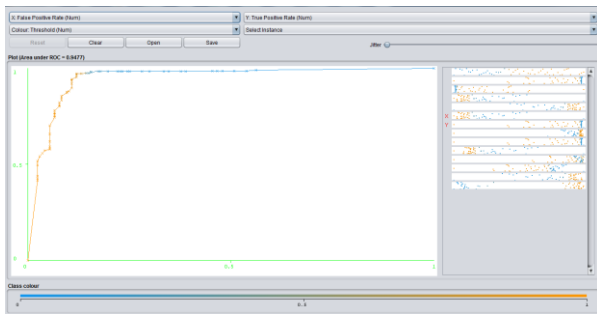


Figure 9. REPTree algorithm ROC value is 0.9477.

J48 Tree algorithm is a part of tree classifiers. The classifier setting already defined such as batch size, confidence factor, seed value etc. 0.06 seconds for build model and accuracy given by this algorithm is 94.51%.and Precision is 95.3%. Random forest algorithm is based on the tree classifier in WEKA tool. Properties mentions previously for example batch size, number of execution slot, number of iterations etc. 0.28 seconds for model. Random forest gave 97.94% accuracy and 98.8% precision.

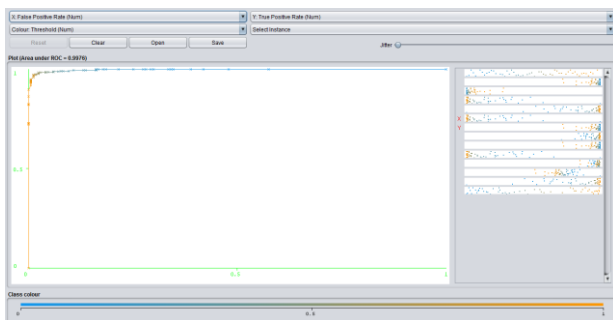


Figure 10. Random Forest ROC value is 0.9976.

LMT(logistic model trees) algorithm is a part of tree classifier, properties stated earlier such as do not make split point actual value, boosting iterations, etc. It's performance in term of accuracy is 97.08%. Its building time is 1.36 seconds for LMT and precision is 97.4% .

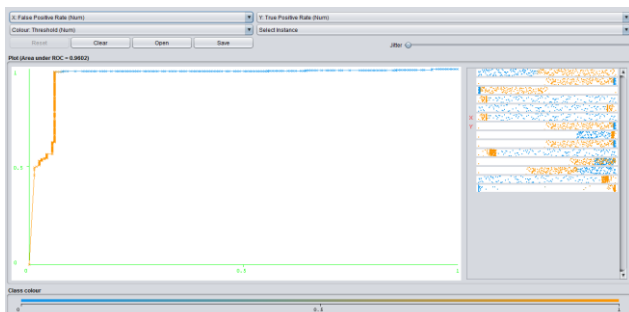


Figure 11. LMT algorithm ROC value is 0.9602

4. CONCLUSIONS

Liver Cancer is recently the incurable disease in the world. It is spreading day by day in developed as well as developing countries. Recently many Machine Learning algorithms are working on liver cancer disease. In this paper four classifiers viz. Function Classifiers, Bayes Classifiers, Meta Classifiers and Tree Classifiers are used and related performance metrics are checked based on the Accuracy, Precision, Recall. Among all the classifiers Random Committee gave better accuracy(98.11%) and it is the part of Meta Classifier.

5 REFERENCES

- [1] Tiwari M. and Chakrabarti P. , "Discovered facts related to liver cancer diagnosis – A Research Perspective", International Journal of Computer Applications, 133(12),pp.31-35, 2016.
- [2] Tiwari M., Chakrabarti P. and Chakrabarti T., "Performance Vector analysis in context to liver cancer – A Support Vector Machine Approach with a survey on the latest Perspectives of Chemistry in liver cancer treatment", International Journal of Computer Science and Information Security, 14(9),pp.1238-1242,2016.
- [3] Tiwari M., Chakrabarti P and Chakrabarti T., "Performance analysis and error evaluation towards the liver cancer diagnosis using lazy classifiers for ILPD", Communications in Computer and Information Science ,837, pp.161-168,2018.
- [4] Tiwari M., Chakrabarti P and Chakrabarti T., "Novel work of diagnosis in liver cancer using Tree classifier on liver cancer dataset (BUPA liver disorder)" , Communications in Computer and Information Science , 837, pp.155-160, 2018.
- [5] Tiwari M., Chakrabarti P. and Chakrabarti T., Liver cancer analysis using supervised machine learning classifiers . LAMBERT Academic Publishing ISBN: 978-3-330-33591-2 (1st Intl. Ed. 2019).
- [6] Kahramanli H., Allahverdi N., A system for detection of Liver Disorders based on Adaptive Neural Networks and Artificial Immune System Proceedings of the 8th WSEAS international conference on applied computer science pp 25-30, 2008.