International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

A Survey on Properties of Adaptive Boosting with Different Classifiers

Sumayya P[1], Shahad P[2]

[1]MEA Engineering College, State Highway 39, Nellikkunn-Vengoor, Perinthalmanna, Malappuram

[2]Assistant Professor, MEA Engineering College, State Highway 39, Nellikkunn-Vengoor, Perinthalmanna, Malappuram

Abstract: ***Adaboost or Adaptive boosting is a recently using boosting technique in machine learning. Generally boosting is used to improve the performance of various learning algorithms. It is mainly used in conjunction with different classifiers. This paper takes a survey based on the properties shown by adaboost with different base classifiers like, support vector machine, naive bayes, decision trees and decision stump. Boosting of these classifiers with adaboost shows difference in properties such as accuracy, error rate, performance etc. Adaptive boosting is widely used in various fields of data mining and image recognition .Now a days it is very helpful in the prediction and diagnosis of particular disease patterns with maximum degree of accuracy by converting a weak learner into a strong learner.***

Keywords: Adaboost, AdaBoostSVM, RBFSVM, Naive Bayes, Decision trees, Decision stump

1. Introduction

Boosting is a technique in ensemble learning that converts weak learners into strong learners. Adaboost is a type of boosting, introduced in 1995 by Freund and Schapire. It runs in polynomial time [1]. And requirement of more parameters is insignificant. Its name indicates its adaptive nature. It works by selecting a base algorithm. The output of the selected algorithm is taken as the final output of the boosted classifier that must be combined a weighted sum.

Adaboost works in the training phase. First step is assigning equal weights to the given training examples. Then select the base algorithm and apply it on the data. If prediction is incorrect using the first learner, then it gives higher weight to observation which has been predicted incorrectly and also decreases the weight of correctly predicted samples. Continue the iteration n times and update weights until the maximum accuracy obtained.

We can use adaboost with different base classifiers and each can show variations in properties. A combination of Adaboost with an appropriate base classifier lead to successful prediction in different fields such as medical data mining, educational data mining, image recognition etc. Different base classifiers used here to explain with Adaboost are SVM, Naive bayes, Decision trees and Decision stump.

2. Working of Adaboost with Base Classifiers

2.1 Support Vector Machine

Support vector machines include in supervised machine learning category. It usually separates a hyperplane in a high dimensional space and is a discriminative classifier. Hyperplane gives the lowest distance to the training examples. It can be used for data analysis of classification and regression. The main goal of the margin is to maximize the number of support vectors and thereby increasing the prediction accuracy [2]. In addition, SVM uses kernels for non linear classification. Figure 1 shows a separated optimal hyperplane that maximizes the margin of training examples.

Adaboost with SVM as weak learner is called AdaBoostSVM.SVM uses suitable kernel functions to compute weights, threshold values, location of centers etc. In most cases, RBF is used as kernel function. The performance of classification depends on its parameters. The Gaussian width σ and the regularization parameter C are the parameters used for RBFSVM. Although RBFSVM cannot learn well when a very low value of C is used, its performance largely depends on the σ value if a roughly suitable Cisgiven[3].

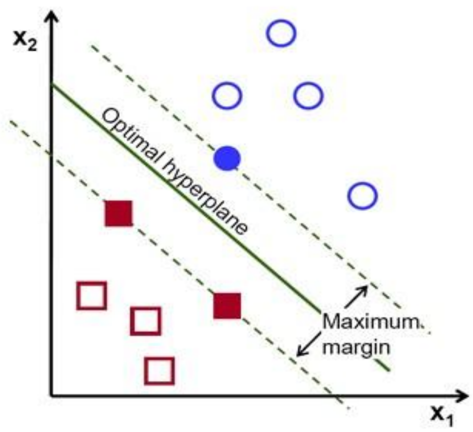

Figure 1: Separated Optimal Hyperplane

But increasing or decreasing the value of σ creates complex classification problem. So the AdaBoostSVM uses a set of moderately accurate RBFSVM that adjusting the σ values adaptively rather than using a fixed σ value. This gives better classification performance and also finds solution for unbalanced classification problems. Figure 2 shows the improvement of AdaBoostSVM over SVM on unbalanced problems when using UCI repository 'splice' dataset as UCI benchmark.

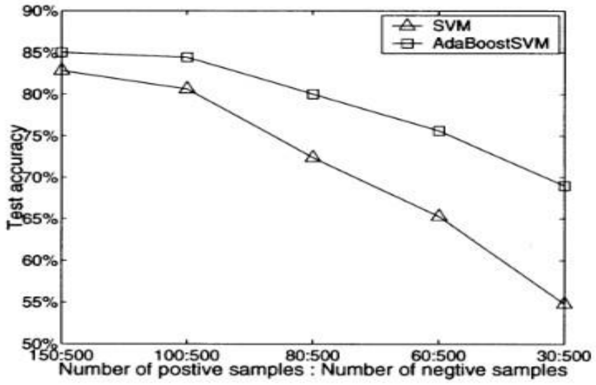

Figure 2: Compare SVM and AdaBoostSVM on unbalanced problem

The performance of AdaBoostSVM over SVM increases with the decreasing ratio of number of positive samples to the negative samples. The AdaBoostSVM has advantages of easier model selection and better generalization performance. It also provides a possible way to handle the over-fitting problem in Adaboost. In the case of unbalanced problems, AdaBoostSVM shows better performance when comparing with support vector machine.

2.2 Naive Bayes

The working of Naive Bayes classifiers are based on Bayes Theorem. They are very efficient and probabilistic classifiers and includes in supervised machine learning category. The algorithm is assigned to a class label which is learned by conditional probability and from training data and all attributes is independent with respect to the class label. Bayesian rule is applied to compute the probability of class label and the highest posterior probability class is calculated [5].

W_t

The boosting of naive bayes with Adaboost depends on the stability of the classifier. Studies show that boosting does not work well for naive Bayesian classification because of the stability of naive bayes with strong bias [6]. However by reducing bias and increasing its variance makes the classifier unstable and helps in successful boosting. Introducing tree structures into naive bayes can reduce bias and increases variance by using maximum depth of trees.

2.3 Decision trees and Decision stump

Decision trees are graph like structures with decisions and their possible effects and contain root node, internal nodes and leaf nodes. The root node and terminal node includes different test conditions to separate different attributes [7].

Applying adaboost algorithm to decision trees are very effective than other classifiers with adaboost because of its non linearity. And also they are possibly fast to train.

Decision stump is a category of decision tree consisting only one level. It uses the most prominent attribute for prediction [2]. Figure 3 shows working of decision stump.

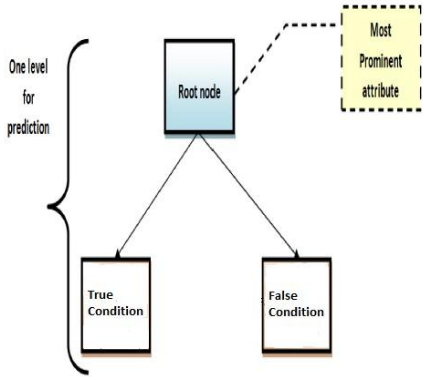

Figure 3: Working of Decision stump.

Adaboost with decision stump as weak learner is widely used now a days because of its great success in boosting. Boosting is started by applying equal weights to all points and find decision stump with lowest weighted training error according to the initial weights. Finding best decision stump or best attribute is important in boosting. So for selecting best attribute, consider splitting on each feature and pick the lowest weighted error. The coefficient for the particular classifier is computed by the formula

$$W_t = \frac{1}{2}\ln\left(\frac{1-weighted\ error\,(f_t)}{weighted\ error\,(f_t)}\right) \quad (1)$$

Where f_t is selected decision stump or attribute. Then recomputed and normalize the weights and final model is predicted by the formula

$$y = sign\left(\sum_{t=1}^{T} W_t\ f_t(x)\right)$$

Where is the coefficient.

[4] Giorgio Valentini and Thomas G. Dietterich, "Bias-variance analysis of support vector machines for the development of svm-based ensemble methods:' Journal of Machtine Learning Research, vol. 5, pp. 725-775,

Boosting decision stumps improves predictive performance as well as speed. Different studies show that Adaboost- decision stump classifier provides highest accuracy and low value of error rate [2]. Many face detection algorithms like Viola-Jones and many disease predictions employ Adaboost- decision stump combination.

3. Discussion

Most fields in data mining and image recognition tried to utilize the benefits with boosting and many of them achieved successful results. Introduction of Adaboost is a milestone in ensemble learning because of its adaptive nature.

Boosting of different base classifiers with Adaboost shows variations in properties. Boosting of support vector machine shows better generalization performance and ways to handle the overfitting problem by selecting proper parameters. It also shows that a noticeable increase in accuracy by the use of moderately accurate SVM learners. In the case of Naive bayes , boosting does not work well because of its stability. However introduction of tree structures into naive Bayesian classifier makes it unstable and helps in successful boosting. Non linear structure of decision trees is very effective in boosting. Also it solves overfitting problem as SVM do. Performance and speed are more than SVM and Naive Bayes. The one level decision tree, Decision stump shows high accuracy and low error rate because of the use of best single attribute with adaptive boosting.

4. Conclusion

This survey concludes that SVM, Naive Bayes, Decision trees and Decision stumps shows variations in their properties with Adaboost. Comparing these classifiers, SVM, decision trees and decision stump improves performance with adaboost while Naive Bayes does not show any improvement. Decision stump provides higher accuracy and low error value in prediction than the other classifiers.

References

[1] Daumé III, Hal. "A course in Machine Learning." chapter 5 (2012): 69.
[2] V. V. Vijayan and C. Anjali, "Prediction and diagnosis of diabetes mellitus — A machine learning approach,"
[3] 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, 2015, pp. 122-127. doi: 10.1109/RAICS.2015.7488400
[4] Xuchun Li, Lei Wang and E. Sung, "A study of AdaBoost with SVM based weak learners," Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005, pp. 196-201 vol. 1. doi:10.1109/IJCNN.2005.1555829Jul 2004.
[5] Qiong Wang, George Garrity, M, James Tiedje, M and James Cole R, "Naïve Bayesian Classifier for Rapid Assignment of RNA Sequences into the New Bacteria Taxonomy", Applied and Environmental Microbiology, vol 73 pp- 5261 b- 5267.
[6] Ting, Kai Ming, and Zijian Zheng. "A study of adaboost with naive bayesian classifiers: Weakness and improvement." Computational Intelligence 19.2 (2003):186-200.
[7] http://www.saedsayad.com/decision_tree.htm
[8] Robert E. Schapire and Yoram Singer, "Improved boosting algorithms using confidence-rated predictions," Machine Learning, vol. 37, no. 3, pp. 297-336, Dec 1999.
[9] Thomas G. Dietterich, "An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization," Machine Learning, vol. 40, no. 2, pp. 139-157, Aug 2000.
[10] Yubo Wang, "Real Time Facial Expression Recognition with Adaboost", " Proceedings of the 17th International Conference on patter Recognision",Vol –3, pp-926-929,2004

Author Profile

Sumayya P. received the B.Tech. degrees in Computer Science and Engineering from MEA Engineering college Perinthalmanna, Malappuram During 2009-2013 . Right now she is pursuing her M. Tech degree in computer science at MEA Engineering College, Kerala from 2015 to 2017. Her research interests lie in Data Mining field.

Shahad P, Assistant Professor, MEA Engineering College, Perinthalmanna. Received the B.Tech degree in Computer Science & Engineering from Maharaja Prithvi Engineening college Coimbatore. M.Tech degree in Computer Science from College of Enginnering, at MEA Engineering College Perinthalmanna. His research interests include Data Mining.