TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 7376 - 7381

Comparison of the Feature-Based Combiner to Bagging and RSM using Artificial Data

Fuad M. Alkoot
Higher institute of telecommunication and navigation
f_alkoot@yahoo.com

Article Info
Volume 82
Page Number: 7376 - 7381
Publication Issue:
January-February 2020

Article History
Article Received: *18 May 2019*
Revised: *14 July 2019*
Accepted: *22 December 2019*
Publication: *03 February 2020*

Abstract

We experimentally compare the feature based combiner to the most common combiner methods of bagging and random subspace method. The experiments are made on different synthetic data sets to find when the FBC outperforms bagging and RSM. Results show that FBC outperforms other combiners when a small number of features exist, especially when the number of combined classifiers is low. As the number of combined classifiers increase, it underperforms other combiners. We mainly find that when number of classes increases some classes may not be well represented in the training set. This is where FBC performs worse than other methods. The problem increases at smaller number of samples. This is an obvious consequence because the possibility of class misrepresentation increases as the size of the training set decreases

1 INTRODUCTION

AI apparatuses are progressively being utilized in numerous application territories to robotize choices. Specialists looked with the undertaking of arrangement use classifiers or scientific models that can play out the errand of characterization or basic leadership, in light of a recently given information. These classifier models or specialists have a capacity to spot patterns and connections in enormous informational indexes, which makes them appropriate for some applications.

So as to improve the precision of classifiers scientists have discovered that consolidating (combining) the choices of more than one classifier would yield better outcomes over the best single classifier. The productions [8,13,11,12,14,17] are instances of early endeavors at demonstrating and utilizing combiners effectively. This was trailed by various ways to deal with numerous master combination [9,6,15,7] These range from basic blend rules which don't require any preparation [1,2,5,8,10], to complex combination This was trailed by various ways to deal with numerous master combination [9,6,15,7] These range from basic blend rules which don't require any preparation [1,2,5,8,10], to complex combination systems which can adapt to in consistent master qualities, can even powerfully adjust to information [14,16]. In [3,4] we proposed a novel structure theory for classifier mix by taking the view that the plan of individual specialists and combination can't be settled in disconnection. Every master is built as a feature of the worldwide plan of a last numerous master framework. The plan procedure includes together adding new specialists to the different master engineering and adding new highlights to every one of the specialists in the design. The underlying examinations indicated a minor profit by the new system [3,4]. Further examination in 20 demonstrated a significant bit of leeway over packing and irregular subspace strategy or choice tree timberlands. The explanation for the extraordinary exhibition of the proposed FBC technique was not distinguished. Thusly, we target finding when and why FBC outflanks different strategies. We can accomplish that utilizing hypothetical or logical techniques for evidence. In our analyses we focus on tentatively demonstrating the upside of our recently proposed combiner technique over existing strategies. We accomplish that, at this underlying stage, utilizing various sorts of manufactured information.

January - February 2020
ISSN: 0193 - 4120 Page No. 7376 - 7381

In [21] we have exhibited beginning outcomes that included presentation of stowing and RSM classifiers, however barring outcomes for the FBC. in this paper we present outcomes for the FBC combiner which enables us to see some portion of the solution to our primary inquiry; "When does FBC beats stowing and RSM?".
In the following segment we depict the combiner techniques and the classifiers utilized in our analyses. We likewise depict how the counterfeit informational indexes were made to reproduce the different conditions under scrutiny. In segment 3 outcomes are displayed trailed by the end in segment 4.

2 EXPERIMENTAL METHODOLOGY

2.1 Combiner methods:

We aim at comparing combiner methods at various conditions. Experiments are designed for different data types using k-NN and neural network classifiers. The compared combiners are bagging [6], random subspace, RSM [8], and our previously proposed Feature Selection based combiner, FSC [3, 4, 20]. Bagging predictors proposed by Breiman, is a method of generating multiple versions of a predictor or classifier, via bootstraping and then using those to get an aggregated classifier. The total number of samples in each bootstrap set is equal to those of the original training set. The second combiner 'RSM' aims at creating diverse classifiers by assigning different features to each classifier. The number of features is set at a fixed value, m, less than the total number of features. Each classifier is assigned a subset of features that are randomly selected without replacement from the full feature set. This results in classifiers having different views of the data space. We set m to equal 67 percent of the total number of available features. In comparison to 50% recommended by [8] we found better rates are achieved at 67%. The third combiner is the feature selection based combiner, FSC, proposed by Alkoot & Kittler [3,4] and it is based on the principal that the feature selection and the combiner performance are linked.

The best feature subset is selected for each classifier based on the combiner system performance instead of the individual classifier performance. We have set it to 5 maximum number of classifiers. Any feature selection method can be used to add the best feature subset such that the combiner system error rate is minimized. For each classifier under construction one feature is inserted at a time and the system performance is checked. After checking all features, the feature yielding the best system performance is permanently inserted to the classifier under construction. When the feature insertion process is completed for the maximum number of classifiers in the system the process is repeated from the first classifier until all features are used up or the system error rate is not improved by the insertion. The process continues if the addition of a new feature does not degrade the system performance, and there are an unused number of features. However, on the first run across the classifiers we add a feature even if it does not improve the system. That is, we force the insertion of the best feature to the classifiers, even if that does not improve the system. The feature selection method used is the 2-forward-1-backward method.

2.1 Classifier types:

We focus our experiments on two commonly known classifiers, k-nearest neighbor and neural network classifiers, [18]. For the nearest neighbor classifier k is set at $\sqrt{N}$, where N is the square root of the number of training samples. The distance metric used is the mahalanobis metric. The neural network classifier used here consists of three layers. The transfer function or output of the first two layers is log-sigmoid, while that of the output or third layer is purelin. The network training function used is backpropagation. The number of neurons in the hidden (second) layer is set at 5. The number of neurons at the output layer is equal to the number of classes.

2.3 Data creation method:

Combiner performance depends on many factors such as classifier type, fusion methods, number of combined classifiers and the data set. Some of the characteristics of the data set that may affect the combiner performance are number of features, number of samples, the number of classes and degree of overlap between the samples

TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 7376 - 7381

from the various classes. In this paper, we aim at finding the effect of data characteristics on combiner performance. Therefore, we create synthetic data with varying values of the aforementioned data characteristics.

We create synthetic data by adding two random numbers. One is generated from a normal distribution with zero mean and standard deviation S. The second is generated from a uniform random number generator between 0 and 1 which is multiplied by a factor n. Next, we add these two random outcomes to create a feature value for each sample. For each feature the values of S and n are changed, where S ranges between 0.1 and 1 for each feature, and n ranges between 0 and 0.9 for each S. This yields 100 combinations that constitute different distributions for 100 features. Samples of each feature are shifted so that the minimum sample value is zero.

For each of the 25 classes 500 samples at each feature are generated. However, values of each additional class above class 1 are shifted by an amount h to avoid overlap with the lower class. The value of h determines the amount of overlap with the previous class, and hence the degree of data difficulty. Therefore, h =mean of lower class + f x S; where S is the standard deviation of the normal distribution from which the samples were drawn. We set f = 2 to generate easy data. Two additional data are generated by setting f to a random number generated from a normal distribution with mean = -3 and standard deviation =1 and mean=-4 and standard deviation=1. These two versions represent more difficult data, where the data values of the second class are shifted by less amount making the overlap with the previous classes higher. Figures show the data sample distributions for various classes and features. Table 1 presents the various parameter values of the created data sets.

Table 1. Data sets parameters

Samples	100	200	500	--	--	--
Features	10	20	80	150	500	--
Classes	2	3	4	5	15	25

The combination of these parameters yields 3 x 5 x 6 = 90 possible data sets. For each data set we repeat the experiments for two classifier types; namely k-NN and neural network classifiers. Also, the experiments are repeated for various number of combined classifiers of 3, 5, 7, 9 and 11.

To compare the combiner methods we need to repeat all the experiments for each of the four combiner methods. For each combiner the number of run experiments are 900, that yield a total of 3600 classification rates for all four combiners.

Currently, we have created 54 data sets. In the near future we will create the data sets for feature values 150 and 400. For the created data sets we finished experiments on the bagging and RSM combiners. In the remaining time we will run the experiments for the remaining two combiner methods.

3 RESULTS AND DISCUSSION

In each of the figures we present results for our proposed feature based combiner, bagging and RSM combiners using the k-NN and neural network classifiers. Figures ae shown for classification rates at different sample sizes of 100, 200 and 500 and for different number of classes of 2, 3, 4, 5, 15 and 25. Each figure is for three combined classifiers and a different feature set size of 10, 20, and 80. This is repeated for kNN and neural network classifiers.

When using kNN classifiers we find that our proposed method outperforms all at the smallest feature set when combining three classifiers and the number of classes is 2 or 3. This advantage reduces gradually as number of combined classifiers increases or number of samples decreases. The advantage totally disappears at 9 combined classifiers for the smallest feature set size.

When the feature set size increases to 20, the advantage of our proposed feature based combiner disappears quicker at 7 combined classifiers. At the largest feature set size of 80 the advantage of FBC exists only when number of combined classifiers is 3, and for the largest set size of 500.

When using neural network classifiers the advantage of our method becomes less obvious. Only at the largest set size of 2 class problems, it equals the other combiners and sometimes

January - February 2020
ISSN: 0193 - 4120 Page No. 7376 - 7381

outperforms RSM. This is true for all feature set sizes and number of combined classifiers.

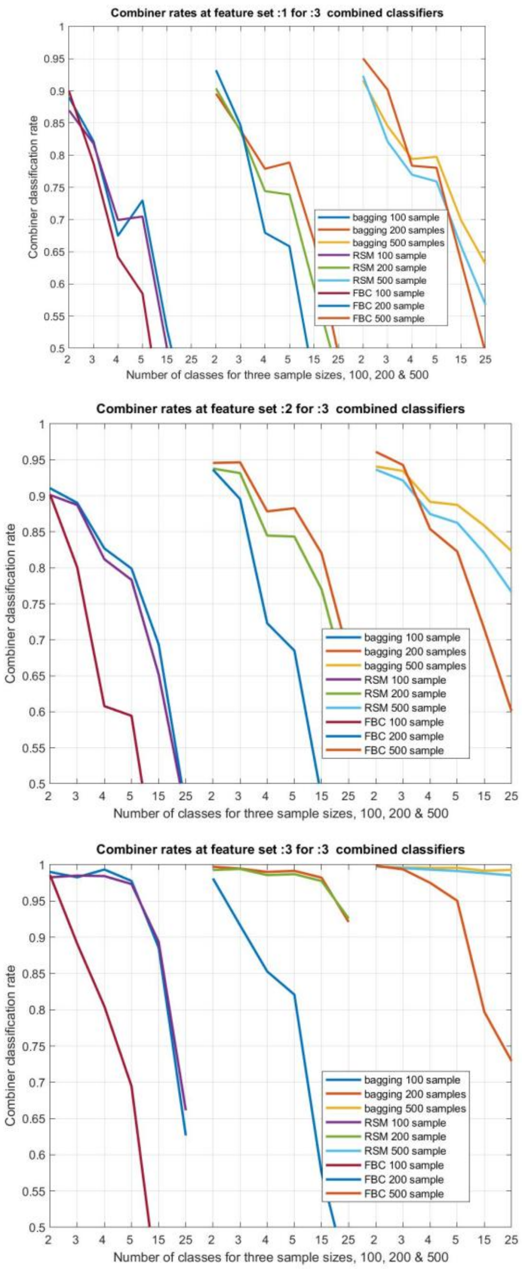

Fig.1: bagging, RSM and FBC methods when using k-NN classifiers at three different data set sizes, using 3 combined classifiers, for feature set sizes 10 , 20 and 80 features.

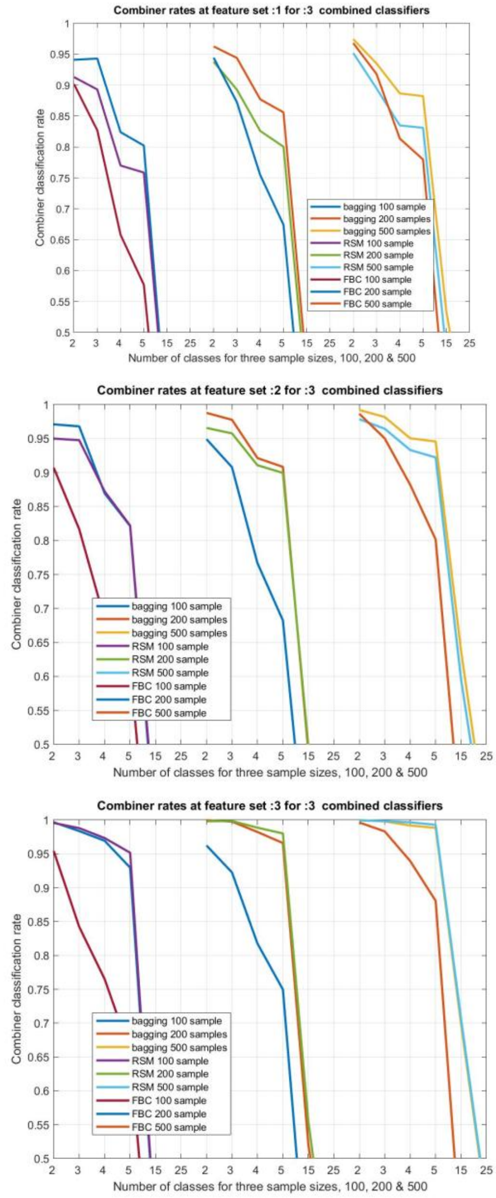

Fig.2: bagging, RSM and FBC methods when using neural network classifiers at three different data set sizes, using 3 combined classifiers, for feature set sizes 10 , 20 and 80 features.

It is obvious that our method is not suitable for high number of classes. Results show that all methods drop in performance as the number of classes increases. However, our method drops by a larger amount. At large number of classes some classes may not be well represented in the training set, especially when number of training samples is small This indicates our method is more sensitive to missing samples in some classes. All methods improve as the number of samples increases, however our method benefits most from this increase. This indicates that our method suffers

January - February 2020
ISSN: 0193 - 4120 Page No. 7376 - 7381

more from the curse of dimensionality, and also from the misrepresentation of some classes in the training set.

4 CONCLUSION

We experimentally compare the feature based combiner to the most common combiner methods of bagging and random subspace method. The experiments are made on different synthetic data sets to find when the FBC outperforms bagging and RSM.

Results show that FBC outperforms other combiners when a small number of features exist, especially when the number of combined classifiers is low. As the number of combined classifiers increase, it underperforms other combiners. We mainly find that when number of classes increases some classes may not be well represented in the training set. This is where FBC performs worse than other methods. The problem increases at smaller number of samples. We can conclude that FBC is an unstable method due to its high dependence on the training set.

Acknowledgment:

The author would like to thank the Kuwait Foundation for the Advancement of Science for financially supporting the costs of the conference participation.

REFERENCES

[1] F. M. Alkoot and J. Kittler. Directing okay-nn classifiers. Example Analysis and Applications, 5(three):326-332, 2002.
[2] F. M. Alkoot and J. Kittler. Altered object combination. Example Recognition Letters, 23(eight):957-965,2002 .
[3] F. M. Alkoot and J. Kittler. Highlight preference for a troupe of classifiers. In Proceedings of the SCI2000 amassing, pages 622- - 627, Orlando, Florida, 2000.
[4] F. M. Alkoot and J. Kittler. Various grasp framework configuration via joined component choice and chance level aggregate. In Proceedings of the Fusion 2000 amassing, volume II, pages ThC5(nine- - sixteen), Paris, France, 2000.
[5] F. M. Alkoot and J. Kittler, "Test Evaluation of Expert Fusion Strategies", Pattern Recognition Letters, vol.20, Issue eleven-13, Nov. (1999)
[6] L. Breiman. Packing signs. AI, 24:123–a hundred and forty, 1996.
[7] Fuad M. Alkoot and Josef Kittler, "Populace Bias Control for Bagging okay-NN Experts", In Proceedings Sensor Fusion: Architectures, Algorithms, and Applications V, Orlando, Florida, 17-20/April. SPIE Volume 4385, pp36. 2001.
[8] T. Ho. The abnormal subspace method for building preference backwoods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(8):832844, 1998.
[9] M I Jordan and R A Jacobs. Progressive combo of professionals and the em calculation. Neural Comp., 6:181–214, 1994.
[10] J. Kittler and F. M. Alkoot. Entirety as opposed to vote combination in severa classifier frameworks. IEEE Trans on PAMI, 25(1):a hundred and ten- - 115, 2003.
[11] J. Kittler, M. Hatef, R. Duin, and J. Matas. On becoming a member of classifiers. IEEE Trans Pattern Analysis and Machine Intelligence, 20(3):226–239, 1998.
[12] J. Kittler. Joining classifiers: A hypothetical shape. Example Analysis and Applications, 1:18–27, 1998.
[13] D. Lee and S.Srihari. Dynamic classifier blend utilizing neural systems. SPIE, 2422:26–37, 6 1995.
[14] G. Rogova. Joining the aftereffects of a few neural machine classifiers. Neural Networks, 7(five):777–781, 1994.
[15] R. E. Schapire. A concise prologue to boosting. In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, 1999.
[16] K Woods, W P Kegelmeyer, and K Bowyer. Blend of different specialists utilising neighborhood precision gauges. IEEE Trans Pattern Analysis and Machine Intelligence, 19:405–410, 1997.
[17] M. Dettling, "Bagboosting for tumor grouping with great articulation statistics," Bioinformatics, vol. 20, no. 18, pp. 3583–3593, 2004.
[18] A.R. Webb. Factual instance acknowledgment, second ed., John Wiley and kids. 2002.

January - February 2020
ISSN: 0193 - 4120 Page No. 7376 - 7381

[19] Fuad Alkoot, An Alqallaf, Investigating AI strategies for the invention of mental imbalance. Worldwide Journal of Data Mining and Bioinformatics 16 (2), 141-169.
[20] Leo Breiman, Arching classifier, Ann. Statist. Vol. 26, number 3(1998), 801-849.
[21] Fuad M. Alkoot, An Experimental Comparison of Classifier Combining Methods Using Artificial Data. The 2d International Conference on Innovation in Computer Science and Artificial Intelligence, July 26–28, 2019, London.