Bagging Stable Classifiers for Data Streams

Jan N. van Rijn [1] **Geoffrey Holmes** [2] **Bernhard Pfahringer** [2] **Joaquin Vanschoren** [3]

[1]Leiden Institute of Advanced Computer Science, Leiden University, Leiden, The Netherlands

[2]Department of Computer Science, University of Waikato, Hamilton, New Zealand

[3]Department of Mathematics and Computer Science, TU Eindhoven, Eindhoven, The Netherlands

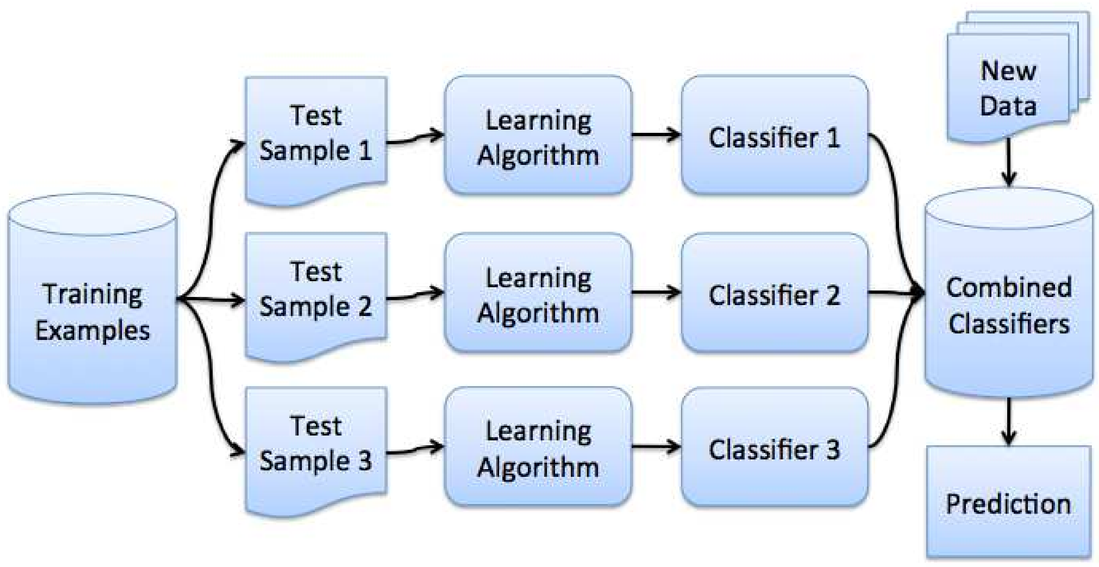
Introduction

Ensembles of classifiers are among the strongest classifiers in most data mining applications. Bagging ensembles exploit the instability of base-classifiers by training them on different bootstrap replicates. It has been shown that Bagging instable classifiers, such as decision trees, yield generally good results, whereas bagging stable classifiers, such as Naive Bayes, makes little difference. However, recent work suggests that this assumption does not apply to the Data Stream Setting.

OpenML and Data Streams

- Networked science [3]: broadcasting data (and questions) fosters spontaneous discoveries.
- Full support for Data Stream Experiments.
- Process an example at a time, and inspect it only once.
- Use a limited amount of memory and time.
- MOA plugin [1] available, for sharing Algorithms and Experiments.

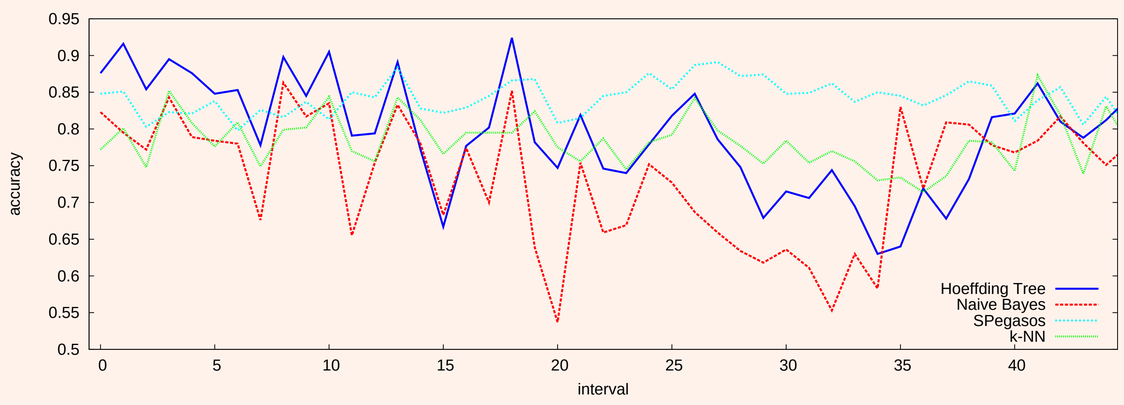
Bagging



Bagging [4] is a technique that trains various models on different samples of the data and combines the predictions.

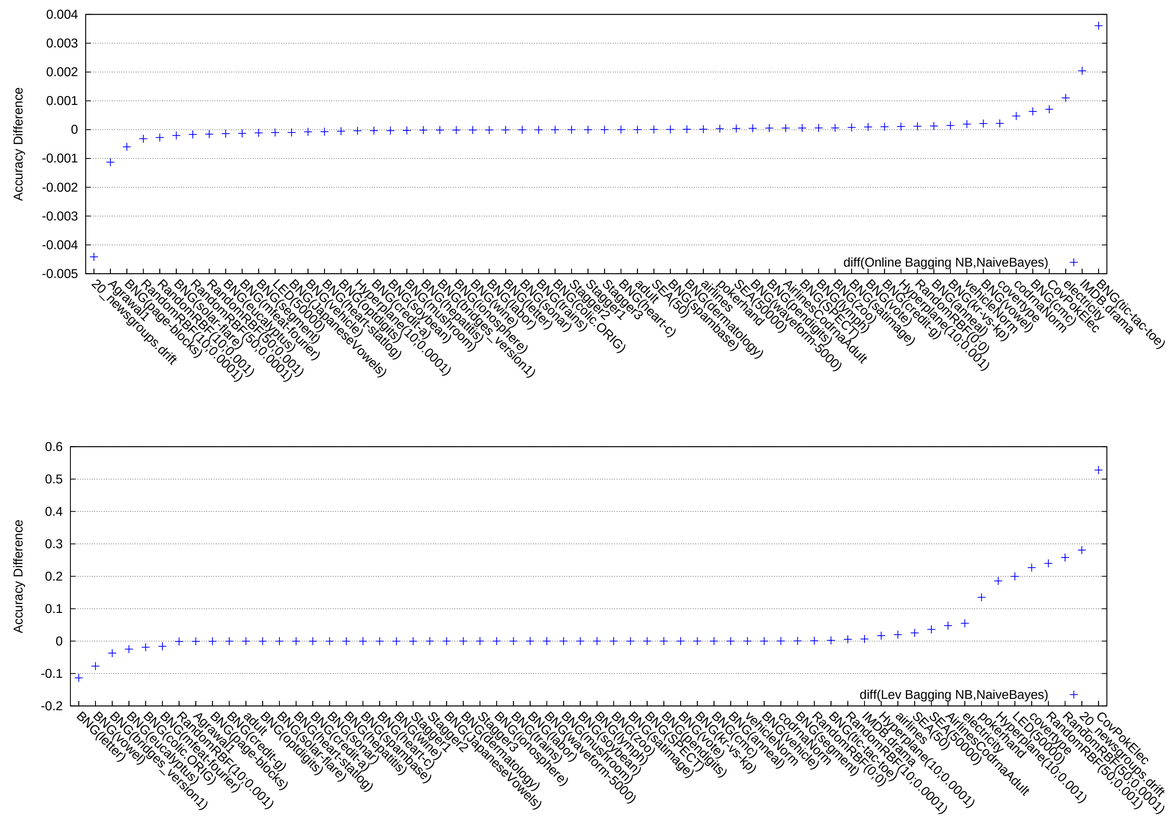
Algorithm Selection on Data Streams

Continuation of the Meta-Learning Experiment presented in [2].



Experimental Results

The experiment contains two classifiers (Naive Bayes and k-NN, with $k = 10$) and two Bagging Schema's (Online Bagging and Leveraging Bagging). The images show the results of both schema's on Naive Bayes.



Statistical Tests

Two statistical test were performed:

- T-Test found no significant differences.
- Wilcoxon Signed-Ranks test found significant difference in many cases.

Table 1: Wilcoxon Signed-Ranks Test results, 95% confidence.

Classifier	Online Bag.	Lev. Bag.
Naive Bayes	no	yes
k-NN	yes	yes

Table 2: T-Test results, 95% confidence.

Classifier	Online Bag.	Lev. Bag.
Naive Bayes	no	no
k-NN	no	no

Conclusions

A possible explanation can be found in the fact that the Wilcoxon Signed-Ranks Test bases its conclusion on the signs of a classifier; it only considers whether one schema was better, equal or worse on a given data stream. The T-Test bases its conclusion on actual scores. The fact that the Wilcoxon test found statistical evidence that bagging actually improves the performance of stable classifiers in the data stream setting, but the T-Test not, leads to the belief that improvements can be obtained, but these are very limited. More research is required to give a decisive answer to the question whether Bagging Stable Classifiers works on Data Streams.

References

[1] A. Bifet, G. Holmes, R. Kirkby, B. Pfahringer. MOA: Massive Online Analysis. J. Mach. Learn. Res. 11, pages 1601–1604, 2010.

[2] J. N. van Rijn, G. Holmes, B. Pfahringer, and J. Vanschoren. Algorithm Selection on Data Streams. Discovery Science, 17th International Conference, pages 325–336, 2014.

[3] J. Vanschoren, J.N. van Rijn, B. Bischl and L. Torgo. OpenML: networked science in machine learning. ACM SIGKDD Explorations Newsletter 15, pages 49–60, 2014

[4] L. Breiman. Bagging Predictors. Machine learning 24(2), pages 123–140, 1996.

Technische Universiteit Eindhoven University of Technology