**Boosting Algorithm for Machine Learning and Pattern Recognition**

*Technology #cu12127*

Lead Inventor: [Prof. Tony Jebara, Ph.D.](mailto:jebara@cs.columbia.edu)

A boosting algorithm for machine learning and pattern recognition that improves on existing algorithms (e.g., Adaboost) by attempting to minimize both the error and the variance in classification problems.

In machine learning, classification algorithms use classification functions to search for patterns and/or images within data. The better the classification functions, the more accurately the algorithm can identify patterns. Boosting is a technique whereby poor classification functions (“weak learners”) are combined to form a single function with a much higher accuracy (a “strong learner”). This is typically done by determining how accurate a number of learners are against a series of examples and adding them together with a weight that improves the overall learner’s ability to identify previously inaccurately-identified examples. This technology, the EBBoost algorithm, improves upon previous boosting algorithms by weighting the weaker learners not only according to their error, but according to their variance as well. Weighting weaker learners by their error and variance improves classification accuracy with little minimal increase in complexity compared to algorithms like Adaboost.

The technology’s performance was compared to that of the standard boosting algorithm, AdaBoost, and was found to outperform AdaBoost on each and every one of more than 20 datasets. The cost for using EBBoost is also a fraction of the cost for using other boosting algorithms.

Applications:
- Image processing and recognition
- MRI automated diagnosis; e.g. analyzing scans for irregularities such as tumors, fractures, blood clots
- Computer vision; automated driving, identifying obstacles, avoiding collisions
- Security technologies; identifying explosive material chemical signatures from spectroscopy data, identifying weapons on mm-Wave images
- Character and handwriting recognition
- Speech recognition
- Systems biology and drug discovery
Advantages:
– Better classification accuracy for any boosting technology
– May be readily integrated into current machine learning algorithms

Licensing Status: Available for licensing and sponsored research support

Related Publications:

Inventors

Tony Jebara