Binary Classification: Is Boosting stronger than Bagging?

DISCLAIMER: Summarized by AI

Problem they are trying to solve / Purpose of method

The authors aim to evaluate and compare two popular ensemble methods—**Boosting** and **Bagging**—in the context of **binary classification**. The key questions they seek to answer are:

- Under what conditions does **Boosting outperform Bagging**?
- Are there scenarios where **Bagging is more effective** than Boosting?
- How do these methods handle **bias and variance**?

Why the method is introduced / needed:

Boosting and Bagging are both designed to improve the accuracy of weak classifiers by combining multiple models. However, they differ significantly in how they achieve this. The paper addresses:

- The **lack of clarity** on which method is better under which circumstances.
- The need to **empirically and theoretically** understand the behavior of these methods, especially when applied to **decision tree classifiers** like CART and C4.5.

How does it differ from other methods?

Boosting and Bagging both aim to improve prediction accuracy but follow different approaches:

Bagging:

- Focuses on **reducing variance** by training base learners on **random resamples** of the data (bootstrapping).
- All models are **trained independently**.
- Final prediction is made by **majority voting** (classification).

Boosting:

- Aims to **reduce both bias and variance**, especially effective for high-bias models.
- Learners are **trained sequentially**, with each new learner focusing on **examples misclassified** by previous ones.
- Final model is a **weighted vote** of all learners.

Unique Aspects Highlighted:

- Boosting may lead to **overfitting resistance**, even when training error is low.
- Bagging improves **unstable learners** more significantly (e.g., CART), while Boosting can **also improve stable learners** like Naive Bayes .

How the method works

Overview:

- The paper examines **Boosting (AdaBoost and Arcing)** and **Bagging** using empirical tests on 23 datasets.
- The base classifiers tested include **CART**, **C4.5**, and **Naive Bayes**.
- The focus is on comparing their **classification error**, **bias**, and **variance**.

Boosting (AdaBoost/Arcing):

1. Train a weak learner on the full dataset.
2. Increase weight on misclassified instances.
3. Train next learner on the **weighted data**.
4. Repeat for a set number of iterations.
5. Final decision is a **weighted vote** of all learners.

Bagging:

1. Generate multiple bootstrap samples from the dataset.
2. Train a learner on each sample independently.
3. Final decision is made via **majority voting**.

Findings:

- Boosting tends to reduce both bias and variance in many cases.
- Bagging is more consistent in reducing **variance**, especially with high-variance learners like CART.
- Boosting can outperform Bagging, especially when the base learner has **high bias**, but is **more sensitive to noise**.