Ensemble methods

Kristian Wichmann

October 13, 2017

Ensemble learning in machine learning is broadly speaking the practice of making a collection - an *ensemble* - of models, and combine them into a single model with more desired properties.

1 Bagging and boosting

There are two broad categories of such meta-algorithms:

1.1 Bagging

Bagging is short for *bootstrap aggregation*. This is because the ensembles are made through bootstrapping, i.e. random resampling from the training set. Properties of these meta-algorithms:

- Ensembles are build independently in *parallel*.
- Tends to reduce variance.

1.2 Boosting

Properties of these meta-algorithms:

- Ensembles are build *sequentially*, each building on the ones before.
- Tends to reduce bias.