“No Free Lunch” Theorem: No single algorithm wins all the time

Ensemble learning: select a collection (ensemble) of classifiers and combine their predictions.

* A set of classifiers with similar training performances may have different generalization performances.
* Combining outputs of several classifiers reduces the risk of selecting a poorly performing classifier
* if each classifier makes different errors, then their strategic combination can reduce the total error.

1. Bagging: resample training data
2. Boosting: Reweight training data
3. RandomSubSpace : random subsets of features

1. Bagging:Bootstrap Aggregation

* Horizontal splitting of DS
* Sample with replacement: multiple training dataset:on average contain 63.2% of the unique training examples, the rest are replicates.
* Train M classifiers on M training set samples
* M classifiers output are evaluated for same test dataset

Various output deciding mechanisms

1. Majority voting
2. Weighed avg, Sum, Mean, Median, minimum ,... etc

Boosting: Adaboost

* Horizontal splitting of DS
* Sample without replacement
* guarantee performance improvements on fitting training data for a weak classifier
* Start with uniform weighting to all samples in the training dataset.
* During each step of learning perform validation on the given training dataset.
  + Calculate misclassification error for each sample
  + Increase weights of the examples which are not correctly classified by the weak classifier
  + Decrease weights of the examples which are correctly classified by the weak classifier
* Do the process repeatedly until all the training samples classified correctly.
* Multiple classifier outputs are evaluated for same test dataset

Various output deciding mechanisms

1. Majority voting
2. Weighed avg, Sum, Mean, Median, minimum ,... etc

Random subspace:

* Vertical splitting of DS
* Get multiple training sample with all the records with randomly chosen attributes.
* Multiple classifier outputs are evaluated for same test dataset
* Various output deciding mechanisms

1. Majority voting
2. Weighed avg, Sum, Mean, Median, minimum ,... etc

Ensemble: Same training DS with different base classifiers

1. Performance of poor base classifier can be upgraded by combining that with other best performing classifier.
2. But the output is depending more on best performing classifier.

Solution: Better to use same classifier trained with different training dataset and enhance its performance by combining them. (Different training DS with same base classifiers)

Final model is built with all the samples in the training DS. We can expect accurate results on new test samples.