### Introduction

### The bagging algorithm is also known as bootstrap aggregation. This algorithm represents a solution to DTs high-variance estimators. Variance reduction is done by combine multiple learners all fitted on separate bootstrapped samples and average their predictions.

*Proof:*

Consider independent and identically distributed (iid) observations

Variance of each is .

– mean of the observations

The variance of the mean of the observations is the following:

Therefore applying the mean is a method for reducing variance.

### Input parameters

### D: training tuples dataset

K: number of classifiers combined

S: classification learning algorithm

M: ensemble classifier (output of the algorithm)

### Pseudocode

### 

### Bagging in algorithmic trading

There is generally one database in quantitative finance so there is one set of training data. Machine learning algorithms may need multiple independent training sets. Bootstrapping can provide solutions to this problem by generating multiple training sets. Bagging algorithm generates hundreds of thousands of deeply-grown trees across bootstrapped samples of training data. The results are combined in order to reduce variance.

Bagging algorithm improves forecasting accuracy in algorithmic trading, the main disadvantage being the reduced interpretability of data. However prediction accuracy is much more important in algo trading than research interpretability.

Another advantage of bagging is that there is no overfitting by increasing the number of bootstrap samples.