Bootstrapping

From a dataset with *N* observations

1. Randomly select (with replacement) *N* examples and use this set for training
   - remaining examples not selected for training are used for testing
2. Repeat this process for a specified number of folds *K*
3. The true error is estimated as the average error rate on test examples


data
train
test
X_3 X_4 X_5 X_3 X_2 ...
X_2 X_1 X_2 X_5 X_5
X_2 X_1 X_4 X_3 X_4
X_1 X_4 X_5 X_3 X_1
X_3 X_1 X_4 X_5 X_5
X_4 X_2 X_1 X_1 X_3
X_5 X_3 X_3 X_2
X_5 X_4

Bootstrapping

From a dataset with *N* observations

1. Randomly select (with replacement) *N* examples and use this set for training
   - remaining examples not selected for training are used for testing
2. Repeat this process for a specified number of folds *K*
3. The true error is estimated as the average error rate on test examples

Comparison CV and Bootstrapping