TU München
Fakultät für Informatik
PD Dr. Rudolph Triebel
John Chiotellis, Maximilian Denninger

Machine Learning for Computer Vision

May 11, 2018
Topic: Bagging and Boosting

Exercise 1: Bootstrap Aggregation

a) What is the core idea in bagging? How does it differ from boosting?

b) Does bagging reduce the bias of the predictions, the bias or both? Why?

c) What is the out-of-bag error and why is it useful?

Exercise 2: Adaboost (Programming)

Download the file 'banknote_auth.zip' available at the course's website. The data are features of banknotes and the labels indicate whether a banknote is forged or not. The dataset is taken from `https://archive.ics.uci.edu/ml/datasets/banknote+authentication` with some duplicate entries removed. Implement the AdaBoost algorithm with decision stumps as weak classifiers.

a) To begin train on 50% of the data with 20 weak classifiers.

b) Generate a plot of the classification error with respect to the number of weak classifiers. What do you observe?

c) Add more weak classifiers. Does the error still change? What's the optimal number of weak classifiers to use?

d) Now keep the number of weak classifiers fixed and try different training/testing set sizes. How does it affect the classification accuracy?