Machine Learning SS2013

Ulrike von Luxburg
Assignment 01

Arne Schröder, Falk Oswald, Angel Bakardzhiev

April 14, 2013

Matlab Implementation

First, we introduce and briefly describe our M files, included in the attached zip file.

- **knnClassifySingle.m** - function, that uses k-nearest neighbours method to predict label of single datum
- **knnClassify.m** - function, that uses k-nearest neighbours method to predict labels
- **evaluateK.m** - evaluates knnClassify for different k-values and returns the minimal k
- **loss01.m** - Gets as input a prediction calculated by the knnClasifiy and correct labels y. The function returns the average error (empirical risk with respect to the 0-1 loss) for this prediction.
- **drawNumber.m** - visualize a number using *imagesc*
- **doExercise1.m** - loads all training and test data for exercise 1, calls knnClassify and plots the result
- **doExercise2.m** - loads all training and test data for exercise 2, calls knnClassify and plots the result
- **Assignment01.m** - the main script, calls doExercise1 and doExercise2 with different parameters

Questions

Exercise 1

7. Plot the training and the test errors. Do results change between different runs? Why?

Yes, the results change between different runs. The reason is, that we use random training and test data. For each run the data is different, so we get different results.

9. More training examples. How does the performance of kNN classifier change?

The performance of the classifier is the same like before for the test data, increases however approximately by factor 10 for the training data.

10. Unbalanced classes. More training examples. How does the performance of kNN classifier change?

The performance of the classifier increases approximately by factor 1/3 for the test data and by factor 40 for the test data.