Ludwig-Maximilians-Universitaet Muenchen
Institute for Informatics
Prof. Dr. Volker Tresp
Gregor Jossé

26.06.2014

Machine Learning and Data Mining

Summer 2014

Exercise Sheet 7

Presentation of Solutions to the Exercise Sheet on the 23.06.2014

Aufgabe 7-1 Model Comparison

Compare the models of regression and basis functions. Let the prediction for a data point $\mathbf{x}_i \in \mathbb{R}$ be given as:

$$f(\mathbf{x}_i, \mathbf{w}) = \sum_{j=1}^{M_\Phi} w_j \phi_j(\mathbf{x}_i)$$

Employ the PLS-solution $\hat{\mathbf{w}} = \left(\Phi^T\Phi + \lambda I\right)^{-1} \Phi^T \mathbf{y}$ mit $\Phi_{i,j} = \phi_j(\mathbf{x}_i) = \mathbf{x}_i^{j-1}$. The following dataset $\mathbf{X}, \mathbf{y}$ of size $N = 10$ with variance $\sigma^2 = 0.25$ is given:

$\mathbf{X}$	0.3	0.4	0.8	1.5	1.8	3.6	4	4.3	4.6	5
$\mathbf{y}$	7	4.7	0.6	−1.1	−0.3	4.6	5.5	5.7	3.1	−0.3

We want to find the optimal model with basis functions $M_\Phi \in \{1, \ldots, 6\}$. Employ the mean squared error (MSE) as loss-function.

a) Find the best model using cross-validation (5 and 10 times). Do the pairwise tests introduced in the lecture support the decision of the MSE? What influence does the λ-parameter have?

b) Which result do the frequentistic (C_p statistic and AIC) and the bayesian approach produce?

c) Which influence does the data size N have, if you were to simulate a comparable data set with $N = \{100, 1000\}$?

Aufgabe 7-2 Comparison: Next Neighbor Estimator vs Perceptron

Compare the Next Neighbor Estimator to the Perceptron. How could the two classification methods be visualized?