

# Machine Learning

## Exercise Sheet 1

Prof. Dr. Dr. Lars Schmidt-Thieme, Martin Wistuba  
Information Systems and Machine Learning Lab  
University of Hildesheim

October 21th, 2014  
Submission until October 28th, 13.00 to wistuba@ismll.de

### Exercise 1: R-Tutorial 1 (4 Points)

Install R: <http://www.r-project.org/>.

Download „An Introduction to R“ <http://cran.r-project.org/doc/manuals/R-intro.pdf>. Read Chapter 1 and reproduce the sample session in Appendix A (page 82). Create a screenshot during the session and add it to your solution.

### Exercise 2: R-Tutorial 2 (4 Points)

a) Read Chapter 2 and 3 of „An Introduction to R“. In R there are different possibilities to create vectors. Give an example for three of them. Tell in three sentences something about objects and classes in R.

b) Create a linear regression model for the in R integrated data set `cars`. We want to find a predictor for the variable `cars$dist` given `cars$speed`. Get the coefficients, plot the data and add a regression line. Add the used R code to your solution.

Hint: The needed commands can be found in Appendix A.

### Exercise 3: Linear Regression (6 Points)

a) Given are the data instances of the example from the lecture (gas consumption):

$$\mathcal{D} = \{(2, 6), (6, 5), (8, 4.5)\}$$

Estimate the target  $\hat{y}$  for  $x = 10$  using the method of least squares. The true value is  $y = 2$ . Estimate the error. Interpret the result. Create a plot of all distances and show for each data point the least square error.

b) In the lecture was proven for the simple linear regression that

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

minimizes the residual sums of squares (RSS).

Reconsider the proof and provide intermediate steps for the partial derivative.

Setting the derivative to zero is a necessary criterion for the existence of an extremum. Justify that the given solution is a global minimum.

## Exercise 4: Multiple Linear Regression (6 Points)

A website collects DVD ratings and then uses them to recommend users a DVD. Given are the ratings of two among all users (1 Star is the worst rating, 5 the best):

Index	User	DVD	Rating
1	A	<i>The Big Lebowski</i>	4 Stars
2	A	<i>Brazil</i>	2 Stars
3	A	<i>Titanic</i>	5 Stars
4	B	<i>Brazil</i>	3 Stars
5	B	<i>The Godfather</i>	4 Stars
6	B	<i>Toy Story</i>	4 Stars

Three different recommenders which use all ratings would make following predictions:

Index	$\hat{r}_s$	$\hat{r}_r$	$\hat{r}_k$
1	3.7	3.8	3.9
2	2.4	2.5	2.3
3	2.2	3.0	4.1
4	3.2	3.1	2.9
5	4.7	4.4	4.2
6	4.1	3.9	4.2

a) Estimate for every recommender the mean absolute and the mean square error in comparison to the true ratings.

b) A model for combining the first two recommenders is

$$r(x) = \beta_0 + \beta_1 \cdot \hat{r}_s(x) + \beta_2 \cdot \hat{r}_r(x) + \epsilon$$

Estimate the predictions  $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$  using the method introduced in the lecture. Use the Gaussian Elimination to solve the linear equation system. Write down the equations of the system explicitly. Show the intermediate steps (rounded to two positions after the decimal point) in matrix notation.

Hints:

- You can use any software (e.g. R) to execute matrix multiplication or line operations.
- You can check your result with a solver for linear equation systems, e.g. `solve()`-function in R.

c) Estimate for the combined method the RSS, the mean absolute and the mean squared error. How meaningful are the error metrics? Explain.