Machine Learning
Exercise Sheet 3

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

12. November 2013
Submission until 19. November, 13.00 to wistuba@ismll.de

Exercise 6: Multiple Linear Regression (5 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	*The Big Lebowski*	4 Stars
2	A	*Brazil*	2 Stars
3	A	*Titanic*	5 Stars
4	B	*Brazil*	3 Stars
5	B	*The Godfather*	4 Stars
6	B	*Toy Story*	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 average absolute and the average quadratic 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 average absolute and the average quadratic error. How meaningful are the error metrics? Explain.

d)

Estimate the combined prediction for $\hat{r}_s(x) = 3.0$ und $\hat{r}_r(x) = 4.6$. What is negative? What does not fit with the parameters $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$? How can you avoid these results?

Exercise 7: R (5 Points)

Read chapters 4 and 5 of „An Introduction to R".

a)

What are „factors" in R, how are they created and how can they be used?

b)

What is the difference between an array and a vector in R? Mention three operations on arrays and matrices in R.

c)

Write a program in R that fits $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ to the data

x_1	x_2	y
1	2	3
2	3	2
4	1	7
5	5	1

You are allowed to use the built-in matrix operations and equation solvers. Submit the code.