Machine Learning
Exercise Sheet 3

Prof. Dr. Dr. Lars Schmidt-Thieme, Osman Akcatepe
Information Systems and Machine Learning Lab (ISMLL)
University of Hildesheim

14. November 2011
Deadline: 21. November

Exercise 1: Multiple Linear Regression (5 Points)

In one website, DVD ratings are collected, and new DVDs are recommended to their users with these collected ratings. And the following ratings (1 star is the worst, 5 star is the best) are given to two users among these users:

Index	User	Film	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 prediction methods would fit to following predictions with respect to other collected ratings:

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)

Calculate the mean absolute and mean squared error for each method comparing with actual ratings.

b)

A model for the combination of first both method is

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

Calculate the estimations $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ with help of methods presented in the lecture. Use Gauss Elimination Method for the solution of occurring linear equation systems. Write the equation of solving systems explicitly. Write the intermediate steps (rounded to two decimal place) in matrix written way.

Remarks:

- You can carry the matrix multiplications and the row operations out with the computer assistance (e.g. in R).
- You can check your results out with a solver for linear equation systems, e.g. of `solve()` function in R.

c)

Calculate the residual sum of squares, the mean absolute and the mean squared error for the combined methods for the known data. How meaningful are the so calculated error measures? Give reasons!

d)

Calculate the combined prediction for $\hat{r}_s(x) = 3.0$ und $\hat{r}_r(x) = 4.6$. Which negative situation do you notice? What does not fit with the parameters $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$? How could such results be avoided?

Exercise 2: Variable Selection (3 Punkte)

a)

Why for the linear regression models of amount of coefficients is not a good measure for the factor of a variable?

b)

What is the main difference between the Akaike Information Criterion (AIC) and the Bayes-Schwarz Information Criterion (BIC)?

c)

Calculate a linear regression model for the data from exercise 1, which combined all the three prediction methods (presentation of the intermediate steps is not necessary). Calculate AIC and BIC for this model and the model from exercise 1. Which model is to be preferred according to these criteria?

Exercise 3: R (2 Punkte)

Read the capital 4 and 5 from „An Introduction to R“.

a)

What are „factores“ in R, how are they produced and how could they be applied?

b)

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