

# Business Analytics

## Exercise Sheet 5

Martin Wistuba (wistuba@ismll.de)  
Information Systems and Machine Learning Lab (ISMLL)  
Universität Hildesheim

3 June 2014  
Submission until 10 June 2014 23:59

### Exercise 14: Nearest Neighbor Regression (5 points)

- Describe the Nearest Neighbor Regression for  $k$  neighbors ( $k$ -NN). (1 point)
- Consider the *Fundamental Growth Rate in EPS by Industry*<sup>1</sup> dataset accessible via the course web site as two files. The training instances are located in *fgr\_TRAIN.csv* and the testing instances in file *fgr\_TEST.csv*. The predictor variables are Number of Firms", ROE", Retention Ratio and the target variable is the Fundamental Growth". Implement the  $k$ -NN algorithm in R and predict the fundamental growth of the testing instances. Measure the prediction accuracy via the MSE. (Report your source code with brief comments.) (2 point)
- Experiment for various values of  $k \in 1, \dots, 40$ . Plot the results in a line chart with X-axis being  $k$  and Y-axis the MSE over the test set. Which value of  $k$  yields the best accuracy? Comment shortly on the sensitivity of the results on  $k$ . (2 point)

### Exercise 15: Naive Bayes Classification (5 points)

- Concisely rephrase the learning and inference algorithm of the Naive Bayes Classifier. (1 point)
- Consider the Labor dataset (file *labor.pdf*) describing predictor attributes such as Cost of Living Adjustment", Pension", "Vacation", "Bereavement Assistance and Contribution to Health Plan". Based on those data, the aim is to predict whether a job is good or bad (its Class"). Compute all the elements of the learning algorithm, including the probabilities of each predictor attributes' value conditional to the class. (3 point)
- Apply the inference algorithm to compute the probability that a job is good if cost-of-living-adjustment=tc, pension=empl\_contr, vacation=generous, bereavement-assistance=no and contribution-to-health-plan=half? (1 point)

### Submission

- Electronically to wistuba@ismll.de. Text submitted as pdf, code submitted as source files. No archives.

---

<sup>1</sup> [http://pages.stern.nyu.edu/~adamodar/New\\_Home\\_Page/datafile/fundgr.html](http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/fundgr.html)