

Machine Learning

Exercise Sheet 1

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

October 30th, 2013
Submission until November 6th, 13.00 to wistuba@ismll.de

Exercise 1: Weka (5 Points)

Install Weka: <http://www.cs.waikato.ac.nz/ml/weka/>.

Open Weka and have a closer look to the iris data set by plotting the data (CTRL + P). The data set is found in your weka folder in the folder "data". Reproduce the plots you have seen in the lecture.

- Which classes do exist?
- Which features are given?

Now classify the data using SMO with a 10-fold cross-validation.

- What is cross-validation? Why is this useful?
- What is your accuracy? (Correctly Classified Instances / All Instances)
- What is your accuracy per class? Were this results expectable?

Try further classifiers and report its accuracies.

Exercise 2: R (5 Points)

Install R 2.10.0: <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.