Chair for Algorithms and Data Structures Prof. Dr. Hannah Bast Elmar Haußmann

Information Retrieval WS 2015/2016

http://ad-wiki.informatik.uni-freiburg.de/teaching



Exercise Sheet 12

Submit until Tuesday, February 2 at 2:00pm

Don't forget the (online) evaluation, it's important!

Exercise 1 (10 points)

Please carefully fill out the (online) evaluation form for this course. You should have received an email (on Tuesday, January 25) with a link to the evaluation form. The evaluation is anonymous. Just confirm in your *experiences.txt* for this exercise sheet that you have completed the evaluation. Please be honest and concrete with your feedback. We particularly value the comments in the free-text fields.

Exercise 2 (5 points)

At the beginning of Lecture 12, we discussed a variety of possible improvements of the basic Perceptron algorithm, and you are free to come up with other improvements.

Implement two versions of the Perceptron algorithm: the baseline version (10 iterations, no other refinements) and an improved version (using any combination of the suggested or your own improvements). Try to make your improved version as good as possible. You can use (part of) the *training* set to evaluate how well your current version works, but you must not use the test set at this stage (that would be overfitting and considered cheating in a real competition).

Eventually, settle on a version and evaluate it. Your program should take the name of the training set and of the three test sets provided on the Wiki. For each of the three test sets, it should output the precision of the baseline version and of your improved version.

Exercise 3 (5 points)

Determine the p-value of each of these improvements according to the (two-sided version of the) t-Test, as explained in the lecture. For each pair of compared algorithms, consider as null hypothesis that the precision of the two algorithms is the same. Add your results to the table on the Wiki, following the examples already there. Briefly discuss your results in your experiences.txt.

Add your code to a new sub-directory *sheet-12* of your folder in the course SVN. Make sure that *compile*, *test*, and *checkstyle* run through without errors on Jenkins. As usual, commit a text file *experiences.txt* where you briefly describe your experiences with this sheet and the lecture.