Data Mining: Practical Assignment #2 (pre-sheet)

Due on Thu & Fri, April 24-25 2014, 10:15am-13:15 & 14:15am-17:15

Pre-Task 1

This is a precusor for the upcoming practical courses, so that you can prepare the topics with the according literature at home. The assignment sheets will be handed out in the practical course to be solved there. The sheets will be uploaded in the Data Mining Commsy-room after the Friday group.

For the Data Mining Practical Course on Thursday (24.04.14) and Friday (25.04.14) inform yourself or, respectively, recapitulate the following topics:

1. Chi Square Statistical Tests

2. Correlation Analysis

3. Principal Component Analysis (PCA)

Further, you will make your first steps in Matlab programming.

Recommended literature:

1. Kantardzic, M. Data Mining: Concepts, Models, Methods, and Algorithms (2nd Ed.). Wiley-IEEE Press, 2011.

2. Han J. & Kamber, M. Data mining: Concepts and techniques. Elsevier/Morgan Kaufmann, Amsterdam, 2006.

3. Marsland, S. Machine Learning - An Algorithmic Perspective. CRC Press, 2009.

4. Rojas, R. Neural networks : a systematic introduction. Springer, 1996

All books are available in the Informatics library. In addition, Rojas' book is available in German, also the individual chapters are publicy available in pdf-format (page.mi.fu-berlin.de/rojas/neural/neuron.pdf).

You are further encouraged to seek appropiate literature. If you find any books interesting and suitable for the Data Mining lecture/practical course apart from the suggested, please let us and other students know.