

Data Mining: Practical Assignment #3 (pre-sheet)

Due on Thu & Fri, May 08-09 2014, 10:15am-13:15 & 14:15am-17:15

Pre-Task 1

This is a precursor 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 (08.05.14) and Friday (09.05.14) inform yourself or, respectively, recapitulate the following topics:

1. Data representation for classification, especially have a look at confusion matrices. In addition, you should be familiar with the quantities you can derive from them, e.g. False Positive Rates (FPR) and their computation
2. Association rules, especially the apriori-algorithm
3. Decision Trees: Entropy, Information Gain and their computation

Recommended literature:

1. Han J. & Kamber, M. Data mining: Concepts and techniques. Elsevier/Morgan Kaufmann, Amsterdam, 2006.
2. Kantardzic, M.: Data mining : concepts, models, methods, and algorithms. Wiley, NY, 2011.

You are further encouraged to seek appropriate 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.