

Milestone 3

Mateusz Pawlik

December 12, 2011

Exercises 9 - Data Mining Challenge: Introduction - (12.12.2011, 14:00-16:00)

The third milestone considers data mining part of the course. You are about to use data mining techniques to solve a challenge task. The challenge task we propose is described here <http://2012.msrrconf.org/challenge.php>. We assume that every group will work on the same challenge. However, there can be exceptions due to Your preferences and good motivation.

The Milestone delivery consists of a four-pages (or more) report in which You should describe the obtained results. As You learn new data mining techniques during the lecture You will update Your report with the findings after usage of those techniques. Feel free to use tools such as ELKI <http://elki.dbs.uni-lmu.de/> WEKA <http://www.cs.waikato.ac.nz/ml/weka/> or others to run data mining algorithms and compare them. Moreover, feel free to use any techniques that help You to understand the data on obtain the challenge's goal.

Warming-up Exercises

1. Go to the website <http://2012.msrrconf.org/challenge.php> and understand what is the challenge about.
2. You are free to choose other data instead the challenge data or choose some other ongoing challenge. You can find them here <http://www.kdnuggets.com/index.html?lg> here <http://www.kaggle.com/> or search for them.
3. Decide which the challenge You will participate in and discuss the choice with Mateusz Pawlik.

Milestone 3 Tasks

1. Understand the challenge's goal.
2. Download the challenge data, understand it and describe what is it about.
3. Use the already known data mining techniques to summarize, describe the data and obtain the goal.