

T61-5060 Algorithmic Methods of Data mining, Autumn 2012,

Groupwork 2

Submission deadline: Fri 14.12 at 23:59. Presentation sessions: Thu 13.12 16:15-18, Fri 14.12 10:15-12

1 Choosing a group work assignment

To start the group work, please send the following information to the course email address (see below for information):

- members of the group,
- type of project, and
- a topic description of a few sentences.

Please do this by Monday 3.12 lecture time at the latest. It is possible to alter the topic and type of project on the fly if the first topic does not seem to work.

2 Arranging to groups

Like in group work 1 you have a choice to self-organize into a group of 3-5 people or keep your group from group work 1. This time we will not assign groups or topics on behalf of you, unless you explicitly ask us to!

3 Types of project

We have three types of projects:

1. Implementation of a data mining algorithm and testing on a dataset. Submission: working implementation + presentation slides.
2. Comparison of existing data mining algorithms. Submission: presentation slides+written report of 4 pages or more.
3. Literature review of data mining algorithms of specific type. Submission: presentation slides+ written report of 8-12 pages.

Each type of project will be presented in a 15 minute presentation in the presentation sessions on 13-14.12.

4 Topics

1. Frequent itemset mining

Resources: <http://fimi.ua.ac.be/>

2. Wikipedia link mining

Data: <http://haselgrove.id.au/wikipedia.htm>

3. Molecular Graph mining

Tools: <http://hms.liacs.nl/graphs.html>

PTE Challenge: <http://www.cs.ox.ac.uk/activities/machlearn/PTE/>

4. Learning to Rank

Resources: <http://research.microsoft.com/en-us/um/beijing/projects/letor/>

5. Collaborative filtering

Resources: <http://ict.ewi.tudelft.nl/jun/CollaborativeFiltering.html>

6. Pick a problem and mine!

Resources: <http://www.inf.ed.ac.uk/teaching/courses/dme/html/datasets0405.html>

7. Propose your own topic!