

CSE 8331 - Spring 2012

Advanced Topics in Data Mining

<http://michael.hahsler.net/SMU/8331>

Guidelines for the Main Project

For the final project you have a choice between a data mining project, implementing a data mining method or writing a review paper. The topic of the final project should be coordinated with your tutorial's topic.

Tasks

1. Please come to the meeting about the tutorial topic on **January 30/February 1** already with an idea about the project. Note that in Week 3 no regular class meetings are held but we will have individual meetings to discuss the project.
2. A mid-term progress report (1 page) for the project is due at the meeting in **Week 10 (March 19 and 21)**. Note that in Week 10 no regular class meetings are held but we will have individual meetings to discuss the progress with the project.
3. The final project report is due midnight on **April 22**.
4. You needs to give a short presentation (15 minutes) of the highlights of your project on **April 23, 25 or 30**.

Project types

- **Data mining project:** Identify a suitable data set, identify research questions, use and compare different approaches, write a short report (15 pages + tables, figures, etc)
- **Implementation:** Identify an algorithm which you would like to implement. Implement the algorithm (preferable with an interface to a data mining environment (R, WEKA, MOA, etc.). Test/compare your implementation with a very small data set. Write a short report (10 pages + tables, figures, etc.)
- **Review paper:** Identify a sub-area of the topic you choose for your tutorial. Write an extensive review paper where you compare results from different researchers. In this paper it is very important to critically evaluate the achievements in the field and results presented in the literature. The paper should be 30 pages + an extensive bibliography (30+ relevant references).

Grading

Your work will be graded using the following categories:

- Abstract (10%)
- Difficulty of topic (10%)
- Clear definition of problem area (10%)
- Technical detail (40%)
- Results (10%)
- References (10%)
- Presentation style and support (10%)

Distance students

Distance students need to prepare a video presentation (using e.g. CamStudio). The video is due at noon on April 30.