Data Mining Lab Course SS 2014

Organization and Contents

L. Richter    J.M. Cejuela

Department of Computer Science
Technische Universität München

Premeeting, Wednesday, Feb 5th
Outline

Contents

Organization
Mission statement: At the end of the course you have gone through the whole path of data mining, from dataset preparation up to meaningful predictions.
Steps

This includes o.a. the following steps:

- Dataset search & Description
- Feature Construction & Selection
- Prediction & Evaluations
Our tools

- Use scripting languages/scripts for data processing
- Use Weka and RapidMiner for various processing steps:
  - discretization
  - feature selection
  - learning algorithm implementations
What we provide:

We provide infrastructure and support:

- A wiki for the documentation of your work and communication with peers and supervisors
- Computation power: various multi-core machines for the analyses
- Mini-talks to selected topics from theory
- Feedback to your questions and discussion of your findings
What we expect:

- Curiosity
- Creativity to come up with own solutions
- Readiness to retrieve (some) necessary background knowledge autonomously
Organization

- All of you are asked to suggest interesting datasets
- We select 6-7 datasets
- There is one team of 3-4 people for one dataset
- You populate the wiki every week
- You get substantial feedback every week in the plenary meetings