Data Mining Lab Course SS 2014

Introduction

L. Richter     J.M. Cejuela

Department of Computer Science
Technische Universität München

Wednesday, Apr 9th
Outline

Timeline

Organization

Data Set Presentation – Task1
Organization

- schedule
- meeting structure
- task for next week
- remaining for today
## Schedule

**Table: Intended Schedule**

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 9th</td>
<td>Intro</td>
<td>May 28th</td>
<td>Descriptive Mining V</td>
</tr>
<tr>
<td>Apr 16th</td>
<td>Data Set Presentation</td>
<td>Jun 4th</td>
<td>Predictive Mining I</td>
</tr>
<tr>
<td>Apr 23th</td>
<td>Data Set Selection</td>
<td>Jun 11th</td>
<td>Predictive Mining II</td>
</tr>
<tr>
<td>Apr 30th</td>
<td>Descriptive Mining I</td>
<td>Jun 18th</td>
<td>Predictive Mining III</td>
</tr>
<tr>
<td>May 7th</td>
<td>Descriptive Mining II</td>
<td>Jun 25th</td>
<td>Predictive Mining IV</td>
</tr>
<tr>
<td>May 14th</td>
<td>Descriptive Mining III</td>
<td>Jul 2nd</td>
<td>Predictive Mining V</td>
</tr>
<tr>
<td>May 21st</td>
<td>Descriptive Mining IV</td>
<td>Jul 9th</td>
<td>Final Presentation</td>
</tr>
</tbody>
</table>
Slot Layout

- each lab meeting is ~90 minutes
- we plan to give mini-talks for selected topics
- 6 out of 7 groups present every week, round robin for the non-presenters
- each slot: 5 min presentation, 10 min discussion
Documentation and Communication

- your work is accompanied by a wiki-based lab journal
  - http://i12r-studfilesrv.informatik.tu-muenchen.de/dmlab2014
- the wiki is the place to document experimental set-ups, results, figures and decisions
- keep your entries up-to-date
- we provide a forum:
  - http://i12r-studfilesrv.informatik.tu-muenchen.de/forum/
  - for extended offline discussions which do not fit in the meeting slot
  - this is for discussion with the supervisors and among the students
Weekly Preparations

▶ prepare your group’s presentation with a clear and elaborate report in the wiki
▶ for the meeting prepare a compressed digest for a 5 min presentation
▶ have your wiki report ready until Tuesday morning to give us time to read
Data Set Presentation

- due to April 16th
- present your data set in 3 min and answer the questions:
  - **Why?** (do you find it interesting?)
  - **What?** (is described by the data set, size, contents, type of data, source of the data set . . .)
  - **Goals?** (give some ideas for prediction tasks)
  - put a elaborate description in the wiki
Remaining

- compile list of participations for the logins
- answer questions