Data Mining Lab Course WS 2014/15

Introduction

L. Richter    M. Hecht

Department of Computer Science
Technische Universität München

Wednesday, Oct 8th
Outline

Timeline

Organization

Data Set Presentation – Task1
Organization

- schedule
- meeting structure
- task for next week
- remaining for today
Table : Intended Schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct 8th</td>
<td>Intro</td>
<td>Dec 3rd</td>
<td>Predictive Mining I</td>
</tr>
<tr>
<td>Oct 15th</td>
<td>Data Set Presentation</td>
<td>Dec 10th</td>
<td>Predictive Mining II</td>
</tr>
<tr>
<td>Oct 22nd</td>
<td>Data Set Selection</td>
<td>Dec 17th</td>
<td>Predictive Mining III</td>
</tr>
<tr>
<td>Oct 29th</td>
<td>Descriptive Mining I</td>
<td>Jan 7th</td>
<td>Predictive Mining IV</td>
</tr>
<tr>
<td>Nov 5th</td>
<td>Descriptive Mining II</td>
<td>Jan 14th</td>
<td>Predictive Mining V</td>
</tr>
<tr>
<td>Nov 12th</td>
<td>Descriptive Mining III</td>
<td>Jan 21st</td>
<td>Final Presentation</td>
</tr>
<tr>
<td>Nov 19th</td>
<td>Descriptive Mining IV</td>
<td>Jan 28th</td>
<td>Final Presentation</td>
</tr>
<tr>
<td>Nov 26th</td>
<td>Descriptive Mining V</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Organization

Slot Layout

- each lab meeting is ~90 minutes
- we plan to give mini-talks for selected topics
- group formation depending on the number of participants
- each slot: 5-10 min presentation, 10 min discussion
Documentation and Communication

- your work is accompanied by a wiki-based lab journal
- This will be published soon
- the wiki is the place to document experimental set-ups, results, figures and decisions
- keep your entries up-to-date
- we can provide a forum if there is demand
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 5-10 min presentation
- have your wiki report ready until Tuesday morning to give us time to read
Grading Criteria I

▶ We award grades for groups based on: The whole semester performance: the complete wiki entries and all presentations.
▶ Presentation criteria: focus to things that matter, communicate you message in an easy and understandable way, it is important, that the audience can follow.
▶ Wiki criteria: volume, conciseness, clearness
Grading Criteria II

- The volume should reflect the amount of time and work you spend on the topic.
- The conciseness refers to the level of details and the precision.
- The clearness: Is the red line of the work visible, is it easy to understand.

Therefore the wiki entries should have meaningful structure and cover different levels of abstraction. You may reference white papers, manuals, tutorials and further reading. We record weekly notes for the final grading. Small derivations in a group can occur. If there are objections or group internal issues you can request an individual grading. In such a case please contact us as early as possible!
Data Set Presentation

- due to Oct 15th
- 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 more elaborate description in the wiki
Remaining

- compile list of participations for the logins
- answer questions