Data Mining Lab

Lothar Richter & Juan Miguel Cejuela

Apr. 15th

Institut für Informatik I12
Outline

- Timeline
- Organization
- Data Set Presentation – Task 1
Organization

- Schedule
- Meeting Structure
- Task for the next week
<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 15th</td>
<td>Kick-off</td>
<td>Jun 3rd</td>
<td>Predictive Mining I</td>
</tr>
<tr>
<td>Apr 22th</td>
<td>Data Set Presentation</td>
<td>Jun 10th</td>
<td>Predictive Mining II</td>
</tr>
<tr>
<td>Apr 29th</td>
<td>Data Set Selection</td>
<td>Jun 17th</td>
<td>Predictive Mining III</td>
</tr>
<tr>
<td>May 6th</td>
<td>Descriptive Mining I</td>
<td>Jun 24th</td>
<td>Predictive Mining IV</td>
</tr>
<tr>
<td>May 13th</td>
<td>Descriptive Mining II</td>
<td>Jul 1st</td>
<td>Predictive Mining V</td>
</tr>
<tr>
<td>May 20th</td>
<td>Descriptive Mining III</td>
<td>Jul 8th</td>
<td>Final Presentation</td>
</tr>
<tr>
<td>May 27th</td>
<td>Descriptive Mining IV</td>
<td>Jul 15th</td>
<td>Final Presentation</td>
</tr>
</tbody>
</table>
Meeting Structure

- each lab meeting is ~90 minutes
- we plan to give mini-talks for selected topics
- group formation on the final 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 a demand

Data Mining Lab SS 15
Weekly Presentations

- prepare your group’s presentation with a clear and elaborate report in the wiki
- prepare a compressed digest for a 5-10 minute presentation in the weekly meeting
- have your wiki report ready until Tuesday morning to give us time to read
Grading Criteria

- We award grades for groups based on: The whole semester performance: the complete wiki entries and all presentations
- Presentation criteria: focus on things that matter, communicate your 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, it is 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!
Task of the week

- due to Apr. 22\textsuperscript{nd}
- 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 later on