STAT 518/539 Course overview

Marina Meilă

Department of Statistics
University of Washington

April 2, 2015
**Expectations**

- understand the paper
  - discuss with me if some proofs should be skipped
- implement the algorithm(s) or methods in the paper
  - you have to implement the paper’s methods yourself
  - OK to use existing code for methods you will compare with
  - OK (& encouraged) to use libraries for peripheral tasks (preprocessing, postprocessing, visualisation)
- reproduce the experiments in the paper
  - discuss with me which experiments, to what extent
- design new experiments (e.g. to demonstrate additional properties, to compare with other methods)

- recommended: try a new programming language (python, C/C++) if you are serious about big data and machine learning
Help at the start

- Mini-lecture(s) on topics of common interest (e.g. models for rank data)
- Lecture notes (some) on networks models courses/stat538/winter15
- (Some) support learning new programming languages
- Help, resources and feedback on writing papers, giving presentations
<table>
<thead>
<tr>
<th>Week</th>
<th>Write</th>
<th>Do</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Summary</td>
<td>Implement data generating process</td>
</tr>
<tr>
<td>3</td>
<td>Approach</td>
<td>Implementation, testing, basic experiments</td>
</tr>
<tr>
<td>4</td>
<td>Algorithm description and 2nd pass over summary</td>
<td>Full implementation, plan experiments</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>Implement variations, methods for comparison</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>Gather data</td>
</tr>
<tr>
<td>7</td>
<td>Experimental results (first pass)</td>
<td>Empirical evaluation</td>
</tr>
<tr>
<td>8</td>
<td>Comparison with previous work</td>
<td>Improvements, more evaluation</td>
</tr>
<tr>
<td>9</td>
<td>Areas of application</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Significance, criticism</td>
<td></td>
</tr>
<tr>
<td>Exam</td>
<td>Experimental results (refined)</td>
<td></td>
</tr>
<tr>
<td>Break</td>
<td>General revision</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prepare presentation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PRELIM</td>
<td></td>
</tr>
</tbody>
</table>
The first assignment

- A naming convention:
  - "paper" = the research article you have chosen
  - "report" = the paper you are going to write about this research article

- The first assignment is to write a paper summary
- This is the first draft of your report’s introduction
- Write it in \LaTeX. Use Vladimir Minin’s template from the Resources
- Due Tuesday April 9 (submit on Catalyst)
How to write the paper summary

Read the whole paper then ponder the answers to the following questions:

1. What is the statistical problem that the paper is trying to solve? Challenges.
2. What is the method/approach of the paper? Only general idea.
3. What are the scientific problems that can be solved by the new method?
4. What is the significance of the paper? How does it advance the state of the art?

Your summary should be organized around these questions (not necessarily in the given order). Total pages: approximately 2–3