Intro Master's Seminar + Data Analysis Project Data Science

Torsten Möller

Outline

- The idea
- Requirements
- Expectations
- Timeline
- How to find a topic

- These slides are also here:
  http://vda.univie.ac.at/Teaching/P1_P2_Masterseminar/

The Idea

There different stages in your Master / for your scientific contributions:

- 1st semester: 6 ECTS Doing Data Science
- 3rd semester: 4 ECTS Research seminar
- 3rd semester: 12 ECTS Data Science Project
- 4th semester: 26 ECTS Master thesis + 2 ECTS Masterseminar

We offer help

- Doing Data Science: getting to appreciate the breadth and diversity of applications
- Research Project: help in finding a topic
- Research Seminar: solidifying the scientific approach
- Masterseminar: solidifying + communicating the topic

The idea

- Several opportunities throughout the master's program to get to know different research directions.
- Try one out in the Projects course
- Solidify in your master's seminar

However

- if previous projects didn't work, there is a chance to switch topics for the thesis still

The idea: DS Project

- Finally your own research project
- Find something that excites you
- Find someone whose research excites you
- Use your strengths (math, stats, programming, etc.)

The idea — Masterseminar

- Here you are writing an expose / survey paper on the topic you have chosen. This will become your first draft / literature review chapter for your thesis.
- You will give 1 presentation—Before writing: a so-called “pre-paper talk”
- Then you will write another chapter of your thesis

universität wien
Faculty of Computer Science

Formal Requirements: Master's Seminar

- You need to have successfully completed the following courses:
  - Doing Data Science
  - Ethical and Legal Issues
  - Project
  - Research Seminar
  - CORE:
    - Intro to ML
    - Statistics for DS
    - Mathematics for DS
    - Optimisation Methods for DS
    - Mining Massive Data
    - Visual and Exploratory Data Analysis

Formal Requirements: DS Project

- You need to have successfully completed the following courses:
  - 4 courses out of CORE:
    - Intro to ML
    - Statistics for DS
    - Mathematics for DS
    - Optimisation Methods for DS
    - Mining Massive Data
    - Visual and Exploratory Data Analysis

Expectations

- DS Project: find and complete an exciting research project
- Master Seminar: you are supposed to present your thesis topic to your peers to get early feedback and to become aware of related work / what others are doing
- Thesis: you are supposed to tackle the state-of-the-art in a well-defined research topic

Expectations

- It's work, i.e. studying is a full-time job!
  - 4 ECTS (Research Seminar) = 100h of your time or 7h/week
  - 12 ECTS (DS Project) = 300h of your time or 20h/week
  - 2 ECTS (Masterseminar) = 50h of your time or 3.5h/week
  - 26 ECTS (Thesis) = 650h of your time in a semester

Expectations

- Project: find topic
  - best before the start of the semester (but not necessary)
  - latest by deadline for dropping the course
- Masterseminar: you should already have a topic and supervisor for your master thesis!
- meet at least 4 times during the semester with your supervisor (ideally every week!)
  1. in the beginning to clarify the topic
  2. after 4 weeks to clarify progress and milestones
  3. one month before end of semester to clarify progress and expectations
  4. end of the semester: to present your results

Expectations

- You need to attend (at least) 2 DS research talks during the course of the semester
- Why?
  - Get to know some great and inspiring researchers
  - Find out how to give inspiring talks
  - Expand your horizon
- Talks valid are once offered by the research network data science; see https://datascience.univie.ac.at/events/

Upcoming talks

- Gramoz Goranci, 11.3, 2-3pm
- Franz König, 8.4., 2-3pm
- Krzysztof Janowicz, 6.5., 2-3pm
- Yoshua Bengio, 7.5. 5-6pm
- Rebecca Kahn, 3.6., 2-3pm
- Stephan Günnemann, 7.6. (time tba)
- Also — feel free to join the Mailinglist https://datascience.univie.ac.at/about-us/mailinglist

Grading: DS Project

- This is individual with your supervisor
- There are no further meetings on the projects with the supervisor
- However, 10% of the grade is based on attending (at least) 2 DS research talks

Grading: Master's Seminar

- 30% of the grade: quality of the survey paper / thesis proposal
- 30% of the grade: quality of the pre-paper talk
- 30% of the grade: quality second chapter
- 10% — for attending (at least) 2 DS research talks
- In order to pass the course you need to achieve at least half of the points for the paper and the presentation, each.

Plagiarism

You will need to write your report / submission in your own words. When referring to the contents of other papers, e.g., regarding the considered problem settings or findings, you need to clearly mark this by adding a reference and, if appropriate, quotes. If you fail to do so, this would be plagiarism and will result in an “X”.

In the case of an existing survey on your topic, your paper should be substantially different. Please consult your supervisor to agree on the focus of your survey.

Timeline Masterseminar

- **Mar 15th** (deadline for dropping the course): confirm a topic and supervisor, enter into Moodle
- **Mar 15th**: if you have no topic, either drop the course or email me, and I will assign you a topic
- **Mar 15th**: if you didn't drop the course NOR emailed me about a topic, it is too late to assign you one; I will drop you from the course
- meet with your supervisor at least twice between agreeing on a topic and presenting your final result
- (MS only) ~**Apr 15th** — presentation day for pre-paper talk
- (MS only) **May 19th** (23:55) — submission of your survey/expose (Moodle)
- (MS only) **Jun 23rd** (23:55) — submission of the second chapter (Moodle)
- (deadlines are strict, no extension is possible)

How to find a topic

General remarks

- you want to enjoy it! — what was the most fun subject thus far?
- take advantage of your strength (programming, math, design, …)
- search for it early (you don't want one assigned)
- talk to potential supervisors!

Which supervisors?

- In principle, everyone from the research network: https://datascience.univie.ac.at/about-us/
- Also, everyone who you had in your CORE classes is eligible
- Here are a few more specific labs

universität wien
Faculty of Computer Science

Data Mining + Machine Learning

- http://dm.cs.univie.ac.at/teaching/open-topics/
- possible supervisors:
  - Christian Böhm (christian.boehm@univie.ac.at)
  - Nils Kriege (nils.kriege@univie.ac.at)
  - Claudia Plant (claudia.plant@univie.ac.at)
  - Benjamin Roth (benjamin.roth@univie.ac.at)
  - Sebastian Tschiatschek (sebastian.tschiatschek@univie.ac.at)
  - Yllka Velaj (yllka.velaj@univie.ac.at)

universität wien
Faculty of Computer Science

Visualization and Data Analysis

- http://vda.cs.univie.ac.at/teaching/open-topics/
- Main contact:
  - Torsten Möller (torsten.moeller@univie.ac.at)
  - Laura Koesten (laura.koesten@univie.ac.at)

Nina Rastinger

- In the context of “Wien's 'Merkwürdigkeiten' - Deep Mapping frühneuzeitlicher Reisehandbücher”
- https://www.oeaw.ac.at/acdh/projects/viennas-memorable-sites-deep-mapping-early-modern-travel-guides
- NinaClaudia.Rastinger@oeaw.ac.at