

Final Project Instructions

Important dates

- Project Proposal Presentation: April 9th
- Preliminary report: April 17th
- Demo/Poster session: May 6th 9am-12pm
- Final report: May 10th

Objectives

The objective of the final project is to make use of what you have learned during this course to solve a real-world task or address an open question/hard problem in the information retrieval literature. You will work in groups of 4 and will have the freedom to choose a project to work on. If you don't have ideas or particular interests you can pick a project from a list of default projects that will be posted on Piazza. After deciding what you want to work on write a short project proposal including:

1. A brief problem statement - what you want to work on and your objectives
2. A brief summary of the dataset: What is available and what will you use from it
3. At least one (ideally 2-3) papers that solve the task
4. An outline (can be bullet point) of your proposed approach
5. Expected results, outcomes or insights

The goal of the project is to demonstrate the group's understanding of the concepts, tools, methodologies and challenges of building data collection, retrieval and analysis systems. Grading will reflect the quality of approach, rigor of evaluation, and reasoning about successes and failures. The final report should concisely summarize your findings and answer the following questions:

1. What approach did you take to address this problem, and why?
2. How did you explore the space of solutions (e.g. different weighting schemes for retrieval, classification models and features, etc...) ?
3. How did you evaluate the performance of the approach(es) you investigated?
4. What worked, what did not work, and why?

Deliverables

- Project proposal presentation (< 3 mins, **+/- 5 slides**) including
 - goals of the project
 - proposed approach (including architectures to be explored, baselines)
 - datasets and evaluation
 - expected results, outcomes, insights
- **Preliminary report** (< 1 page, double column) including:
 - Project proposal
 - Initial results of the baselines
 - Timeline of key milestones
 - The preliminary report should later be extended to a final report
- **Final report** of 2 pages double column maximum
- **Poster** pitch presentation ideally with demo (3 - 5 minutes) summarizing the project.

Templates for these deliverables will be posted on Piazza. Deliverables should be submitted to Gradescope.

Presentations

The project proposal and demo/poster presentations will be held online via zoom. Given that we have people in different timezones, I understand that not everyone will be able to participate. There won't be a perfect solution for this but according to the "where art thou?" polling we did after spring break, most of you are not far from Baltimore so we expect that at least one team member will be close enough to the official timezone to be able to attend. If no one from your group can attend, please contact us. These sessions will be recorded and posted online so if you cannot attend you can always watch them later.

Grading

The key point is that we are not just interested in the absolute performance of what you produce, but primarily the thought process that led you there. Grading will be based on the completeness of the project, the clarity of the writeup, the level of complexity/difficulty of the approach, and your ability to justify the choices you made.

Grade Breakdown

Implementation and results - 50%

- Approach and results demonstrate understanding of the problem (includes figures, tables) 20%

- Rigor of evaluation and reasoning / discussion: 20%
- Innovation: 10%

Quality of write-up - 30%

- Introduction is compelling, problem statement and desired outcome are clear: 10%
- Methods and evaluation protocol are clearly described and there are references to related work: 10%
- Discussion and conclusion comprise of well formulated arguments, ideally grounded in scientific literature: 10%

Presentation during pitch & demo session - 20%

- Discussion is clear, and reflects the *entire group's* understanding of the project: 10%
- Presentation is neat and thorough: 5%
- Questions are answered in a manner that demonstrates mastery of the material: 5%