DS 4420: Project Requirements

Spring 2020

In place of a final exam, you are to complete a final project for this course that implements a non-trivial machine learning method. You may work either alone or in pairs (no groups larger than two). In the latter case, the project scope and requisite effort is expected to be commensurate (i.e., appropriate for two people). Here are some example projects for inspiration (this is intended to be *illustrative*, absolutely not exhaustive!):

- Apply a model to a new dataset/task in which you are interested. In this case, much of the work may go into curating data and critically analyzing results.
- Re-implement a model from a recent ML, vision or *ACL conference.
- Perform an extensive empirical evaluation of a particular model or architecture.

You are expected to consult with me (and/or Sarthak Jain, our TA) on the project to ensure appropriate and realistic scope.

Proposal (3/13)

You (+/- your team member) are to submit a a brief (< 1 page; probably a few paragraphs or half a page). This will describe the following: (1) General project idea; (2) Dataset(s) to be used; (3) Metrics for evaluation; (4) Hypotheses concerning outcomes; (5) Anticipated difficulties/obstacles.

In-cass pitch (3/17)

You (+/- your team member) will pitch your project proposal in class. This will afford chance for me, your classmates, and the TA to provide feedback. Timing and other details will be announced closer to the date (this will depend in part on the number of teams). By this date you should have assembled your data and have some initial exploratory analyses.

Presentation (4/7 or 4/9)

You (+/- your team member) will present your final

Project Write-Up

By 4/14 (midnight) you are to turn in a final write-up (as a PDF file) describing your project and accompanying results. This should roughly be structured as follows: (1) Abstract, (2) Introduction/motivation, (3) Experimental setup; (4) Results/discussion; (5) Conclusions/future work. There is no hard page minimum; the document should be sufficiently long to communicate your work.

Code

All code is to be submitted a single zipped repository. I encourage you to also post your code to GitHub, but this is not strictly required. A Jupyter notebook that produces the main results must be included.

Grading

Projects will graded under the following general set of criteria.

- Project ambition/creativity (20%)
- Execution/implementation (20%)
- Final write-up (40%)
- In-class project presentation (20%)

Execution and implementation will be assessed both indirectly (via the artifacts produced and reported upon in the final report) and directly (via the code).