Statistical NLP - Project Proposal

Proposal Due Monday November 4 - at 11:59pm by email to aparikh@cs.nyu.edu and urvish@nyu.edu

1 Your Final Project

A considerable part of your grade for this course is the final project. This is your chance to explore a topic of your interest and you can choose either a more application-driven or method-driven project. Feel free to ask me if you have questions if your project would be appropriate.

The final project will have 3 deliverables:

- A project proposal due onNovember 4 (3-4 pages, maximum of 4 pages, excluding references)
- A final report due Monday December 16th at 11:59pm (approximately 8 pages, excluding references)
- A final presentation during the final exam time on Dec 17th

All reports should be in LaTeX with the NeurIPS template that can be found here:

```
https://nips.cc/Conferences/2019/PaperInformation/StyleFiles
```

Other points:

- You are encouraged to work in groups (preferably of size 2, but up to 3 is also acceptable) on the final project. If you plan on working alone, please justify why this is necessary. You can learn a lot from your teammates and I therefore discourage individual projects. On the other hand, when the team gets too big, it is difficult to know who did what and one can more easily hide; I therefore prefer small teams.

  Please note that the larger your group, the more total work you will be expected to do. However there is also a minimum of work expected regardless of the size of your team so if you are working by yourself you will likely end up doing more work than each person in a team of two.
- In general, try to stick to problems with existing data sets. Collecting a data set specifically for the class project will likely set you back too much.
- For the course project, you are allowed to use whatever programming language / external resources you want, providing you cite them appropriately.

2 Proposal Write-Up

I am looking for a 3-4-page write-up in LaTeX (maximum 4 pages, excluding references) using the NeurIPS template:

https://nips.cc/Conferences/2019/PaperInformation/StyleFiles

You should discuss the following points:

- Problem Setting
- Data Set(s)
- Evaluation Metric
- Survey of related work / baselines
- Your Proposed Approach

3 Good conferences to find related work

If your project falls under a traditional NLP problem, I recommend these conferences as good starting points when searching for related work:

- ACL (Annual Meeting of the Association for Computational Linguistics)
- EMNLP (Conference on Empirical Methods in Natural Language Processing)
- NAACL (Annual Conference of the North American Chapter of the Association for Computational Linguistics)
- ICML (International Conference of Machine Learning)
- NeurIPS (Neural Information Processing Systems)

There are obviously many good papers not in these conferences. Moreover, if you have a niche application area you may see that papers in other venues may be a better fit. Feel free to consult with me if you are doubtful of the quality of the related work you have found.

4 Next Steps

You are welcome and encouraged to start working on the project early. This is your opportunity to apply what you have learned in the class and have fun.