

CS 272: Statistical NLP: Winter 2020

Project Proposal

Sameer Singh

<https://canvas.eee.uci.edu/courses/22668>

As the second submission for your course project, you will be submitting a project *proposal*, a more detailed version of your pitch that incorporates any feedback provided by me. Each project proposal (one per group) will be a **two page** PDF write-up, in the ACL format, uploaded to Canvas by **February 13, 2020**, containing the following sections.

What to Submit?

The project proposal should contain everything that was included in the pitch, with more details where necessary. Such details include additional content such as more citations, details of the evaluation plan, description of the proposed approach, etc. In particular, it should address any feedback that I provided on your pitch.

For individual sections, I foresee the following changes:

- **Title:** Some of you gave just the name of your team as the title. Use a more descriptive title, akin to papers you have read, that sufficiently summarizes your project, including the part that makes it unique. For example, “Sentiment Analysis” is a bad title, but “Exploring Recurrent Structures for Analyzing Sentiments on Twitter” is pretty good.
- **Team Details:** Include the names and UCINetIDs as author information (**no student IDs please**). If there have been any changes in the team, highlight it (e.g. in red).
- **Problem Setup and Motivation:** I expect you to present a more concrete problem setup, and its relevance/importance to the field at large. I also expect you to show a better understanding of the existing work through description and citations to the related work (you can include a separate related work section if you deem necessary). The references section is not included in the 2-page maximum (so go wild!).
- **Proposed Approach** Give more details about the approach that you are proposing, including what makes it different from existing approaches to your problem, if any. Mention the potential issues you are likely to face, i.e. scalability, lack of labeled data, etc. and how you plan to address them. For replication, describe the particulars of the approach you are proposing to implement, and what resources you have available to you (data/code/compute), and what challenges you anticipate (“data too large”, “team members not familiar with RNNs”, etc.). If you have any preliminary results (yay!), include them in this section, otherwise include other progress indicators, e.g. detailed analysis of the datasets you have.
- **Evaluation Plan** Give details of the specific dataset/plan you have identified for evaluation. Describe the metrics you will use to test the success of your project at the end of the quarter. Given that this is the most important aspect of the project, you should have *definitely* decided both of these by now. For replication projects, propose some analysis that is not included in the original paper, and describe what insights it might provide.

Additionally, include the following section that was not required in the pitch.

- **Plan of Work** In brief, provide a rough outline of the major milestones of the project, and a rough estimate for when you estimate you will achieve it. In particular, include what you expect to have achieved by the mid-quarter *status* report (due in ~3 weeks, March 8). If you have already made concrete progress on the project, other than what has been mentioned in the above sections, mention so in this section.

I am happy to answer questions on Campuswire, or meet in person, if that will help.