CS224n Final Project Proposal

The project proposal is only for groups doing the choose-your-own final project. If you are doing the default final project, you just have to fill out the Google Form with your team members, but no submission to Gradescope is required! Please submit one proposal per team (any member of your team can submit) to Gradescope by **11:59pm on Februrary 8th**. You can **not** use late days for the project proposal. Remember to also fill out the the Google Form telling us who your team is. Your project proposal can be short (a single page). It should have the following headings:

Team: List the members of your team (names and SUIDs)

Mentor: If someone has already agreed to be your mentor, list them. Otherwise, we will assign a mentor to your team.

Problem Description (1-2 sentences): What is the problem that you will be investigating?

Data (1-3 sentences): What data will you use? If you are collecting new datasets, how do you plan to collect it?

Methodology/Algorithm (2-4 sentences): What method or algorithm are you proposing? If there are existing implementations, will you use them and how? How do you plan to improve or modify such implementations?

Related Work (3+ prior works): Which papers will you read to inform your understanding of the problem, and the appropriate methodology to tackle it?

Evaluation Plan (2-3 sentences): How will you evaluate your results? Qualitatively, what kind of results do you expect (e.g. plots or figures)? Quantitatively, what kind of analysis will you use to evaluate and/or compare your results (e.g. what performance metrics or statistical tests)?

Minimal Requirements: Does your project meet the following requirements?

- Dataset with at least 10,000 labeled examples (more will be needed for some tasks like machine translation and 100-way classification).

- Dataset can be completely collected by the project milestone due date.

- Task is feasible: either prior work on the dataset exists or a human can get good accuracy on it.

- You have identified an automatic (i.e., can be computed by a computer) evaluation metric for the task.

- Using NLP is required to get good performance on the task. For example, predicting stock prices from twitter data would not fit this requirement because the strongest signal would be the time series of previous prices, and you'd probably be better off improving that part of the model than working on the NLP.

If not, justify why your project will still be feasible.