

CS674 Natural Language Understanding

Project Information

1 Overview

As part of the coursework for CS674, there is an independent research project that will give everyone a chance to look at some specific topic in NLP in more detail. The project can be a programming project or a survey paper. In both cases, the project should be focused and should include a careful evaluation.

- **Programming projects** should be more than just the implementation of a standard NLP algorithm on some text (e.g. n-gram modeling using the WSJ). They should instead focus on analyzing the effects of various factors/components of the implemented algorithm (e.g. size of n, effect of smoothing, size of training data) on performance; or otherwise perform some sort of empirical analysis of the task under investigation.
- **Surveys** should do more than describe a set of papers on a particular topic. They might, for example propose a unifying framework for the papers, propose a new approach or formalism for the topic in question that incorporates the advantages of the methods discussed in each of the papers, or explain how each paper might be extended to cover some additional aspect of the phenomenon under study.

The project grades are based on four items:

1. **Proposal.** The project proposal should be 2-3 pages in length. It must include: (1) a short description of the topic/problem in which you're interested; (2) a preliminary list of related references; (3) a description of the general approach(es) you plan to investigate to solve the problem; (4) an evaluation plan; and (5) the schedule that you'll follow. Survey proposals should include the questions that you will focus on as you read each of the papers.
2. **Preliminary literature survey.** This should be about 5 pages in length, plus the bibliography. It should include descriptions of some of the main papers on the topic that you are investigating.
3. **In-class presentation.** These will be 15 minutes in length and should include a description of the problem investigated, the approach that you used to tackle the problem, and summarize your results on the problem to date.
4. **Final report.** This should be about 10 pages in length (15 for surveys) plus the bibliography and any relevant appendices. A thorough discussion of related work as well as qualitative and/or quantitative analyses of the results are expected.

Due dates are posted on the course web page.