cs294-5: Statistical Natural Language Processing
Final Project Guidelines

Final Projects: Final projects will entail original investigation into any area of statistical natural language processing, defined very broadly. That means that machine learning over text, HCI, language-vision interfaces, and so on, are fair topics, in addition to the core NLP topics.

Scope: As a broad target, final projects should involve approximately as much work as two homework assignments. For groups of more than one person, the total work should scale roughly linearly with the group size, and be distributed roughly equally. An ambitious, well-done project from a group of two or more (or shared between two or more classes) should be on the order of a conference paper in depth of experimentation.

Grading and Milestones: The milestones will be:

Oct 12th	Abstracts due
Nov 2nd	Proposals due
Nov 23rd	Progress reports due
Dec 5th	Preliminary results due
Dec 5,7th	In-class presentations
Dec 20th	Final reports due

The abstract is just a short paragraph telling me who's in your group, describing the problem you've chosen, sketching the general approach you intend to take and the kinds of data you're going to need. If you haven't already spoken to me about project ideas, you should make an appointment and do so before this point (also, please feel free to use the newsgroup to form groups and bounce around ideas). The abstract mainly serves to give me a chance to help you get resources you may need, and to make sure you've got a plausible direction in mind.

The proposal is a one page description of what exactly you plan to do, designed to convince me that you've got a research plan and that you've started on the project (or at least starting thinking seriously about it). When you submit your proposals, you should have your groups and topics completely firmed up.

The progress report is a statement of what you've accomplished, early numbers, problems, and so on. It can be as short or long as necessary, under a page is fine if all's going well. At this stage, your project should no longer be vaporware.

The preliminary reports should be 1-2 pages convincing me that your basic implementation is complete, you've got some solid results and baselines, and all that you've got left is extensions, comparisons, data analysis, and so on.

On the last week of class, each group will give short presentations on what they've accomplished (5-10 minutes each). The final write-up should be on the order of 8 pages describing your approach, results, data analysis, and so on. Submission of the milestones is required, but you will only receive a grade at the end, based on your write-ups and presentation, so don't stress about the formatting or polish of the intermediate submissions. Under normal circumstances, all group members will receive the same grade for the final project. Late days will not apply to the final reports (since I do have to get your grades in to the university).