CS457 - Final Project
Fall 2011

Overview

In this class we have looked at a number of techniques and application areas, and have examined a few in depth in assignments. The purpose of this project is for you to explore a topic we have examined (or not examined - but related to NLP) that's interesting to you in more depth as a mini research project. You will choose a topic, write a proposal, and then complete what you propose over the rest of the semester. You will also have to give status updates, write a final report and give a final presentation.

The project should meet these specifications:

- Your project should relate to something we have talked about or will talk about in class. I will give some examples in class to get you started, but I encourage you to pick something you're excited about. Feel free to ask me if your idea is appropriate.
- You *must* evaluate the success of your approach. We've looked at how to do this in many domains, but if you're having trouble thinking about evaluation, come talk to me.
- Your project should be in a pair or group of three. If you'd like to do it solo, please come talk to me.
- You should aim for a project that will take about 20 hours of coding work per team member (about 5 hours per week). This is not a lot of work, and things always take longer than you expect, so try to be conservative.

Schedule

date	description
11/15	Project proposal
11/24	Status report 1
~~12/1~~ 12/2	Status report 2
12/9	Writeup
12/16	Presentation

Project proposal [15 points]

(Submit using submit script in txt/pdf by midnight 11/15)

Your first task will be to come up with your project group and figure out what you'd like to work on. Your project proposal should include the following information:

- Members of the team. I'm *strongly* encouraging groups of 2 or 3. If you want to work solo, please come talk to me.
- A one paragraph description of your project including what you hope to accomplish *and* how you will evaluate your approach. You should think about what would be an appropriate way for evaluating your success. We've seen a few ways of evaluating approaches for different application areas, but come talk to me if you're having problems thinking about this part.
- What you plan to accomplish by status report 1 and status report 2. Try and break the project down into intermediate steps so that you can start working on it now.
- What resources you will use/need including code, data, etc. You may use any resources you can find, including code you have written for this class or other classes, data you find on the web, etc. If you would like a resource and can't find it, ask and I might be able to help you. However, you must have found *ALL* resources by the time you submit your proposal. Come talk to me (early) if you're having trouble finding appropriate data.
- One academic paper from the literature (full citation) that tackles the same problem. Some good places to find papers is the ACL anthology (`http://aclweb.org/anthology-new/`). ACL, EACL, NAACL, EMNLP and COLING are all good conferences that contain papers on a variety of NLP topics. If you're having trouble finding good papers come talk to me.

Status reports [10 points each]

(Submit using submit script in txt/pdf before class on the dates specified)

In your project proposal you will specify a list of intermediary goals/accomplishments. A status report should include the following:

1. A couple of sentences summarizing the current state of the project
2. To what extends you accomplished your proposed tasks (from your project proposal)
3. A fairly detailed list of what has been accomplished
4. The amount of time each person spent since last status report (or since beginning of the project)
5. One additional academic paper from the literature (full citation) that is related to your work
6. Any issues/problems that have arisen

Your status report should be concise, but should contain the important details. Your report should be less than a page long.

Writeup [90 points]

(Submit using submit script in PDF by 12/9)

The majority of the points of your project will be determined based on your writeup. Your writeup should follow an academic writing style that is similar to that of a research paper. I would like you to think of this as a real (potential) submission to a conference or workshop. It is unlikely that you can complete enough work to have a submittable paper in this short time (most workshop projects take several months to develop and write up, if not several years!). But if you get excited about your project, I would encourage you to continue working on it after the end of the semester and plan to submit it to a future workshop or conference.

You may writeup your work in one of two ways:

- **Option 1**: Write a research-like paper. We will use the ACL paper format:

  `http://acl2012.org/call/sub01.asp`

  The website has templates for both latex and word, either of which are fine (though if you use word, please print it to a pdf at the end).

- **Option 2:** Write a research-like web page. The web page should still be written in a research style and you should give some thought to formatting and organization. If you do a web page, please print a version to pdf and submit this. Somewhere in the pdf should be the web-page location.

Your writeup should be short (**no more than 2 pages of written text, and 4 pages overall**). Even though it is short, I expect it to be well written, well organized and present what you've done (including your results) clearly and concisely. You should include at least one table or figure displaying your results (though more may be useful) and should have at least 4 citations.

You may organize the writeup however you like, but a common approach would include the following:

- Title and authors
- Abstract: Gives a very high-level view of the problem, approach and results. An abstract is almost never more than a paragraph.
- Introduction: Describe the problem and motivate why the problem is interesting/useful.
- Algorithm description: Clearly describe your algorithm/approach including any challenges you encountered.

- Results: Describe your data, experimental setup, evaluation critrion and how well your system performed. You should spend some time discussing the results, including if anything was surprising or interesting.
- Conclusion: A brief summary of the paper including any challenges, where to next and any high-level comments you have at the end of the project.

Your writeup is **not** a report of what happened in the 4 weeks of your project. It should be a clear problem specification, followed by the approach, followed by the results.

Presentation [20 points]

(Presentations will be during our exam period on 12/16)

Each group will give a short (∼15 mins) presentation of their work during our final exam period. At a high-level, your presentation will have a similar flow to your paper. Your presentation must include the following information:

- Problem
- Motivation: Why is what you did useful?
- Approach: How did you solve the problem?
- Results: How well does it work?

Grading

The project will represent most of your work between now and the end of the semester, so don't get too concerned that there are a lot of things to do. You'll have 6-7 weeks to accomplish the tasks, which should be plenty of time if you stay on top of things.

- Project proposal (15 points) - Meets specifications above.
- Status reports (10 points each)
  - Meets specifications above
  - Work accomplished during this period
- Project and paper (90 points)
  - How creative is your project/solution?
  - How complete is your project? Did you accomplish what you set out to do?
  - How well your solution works and, more importantly, how you evaluated it
    - appropriate evaluation metric

    - appropriate set of experiments
  - Paper meets specifications above
  - The quality of your write-up
    - spelling and grammar
    - clarity and organization
    - proper use of figures, tables, graphs and equations
    - analysis of results

- Presentation (20 points)
  - Covered content
  - Organized and well-prepared
  - Presentation style
  - Quality of the slides/presentation material