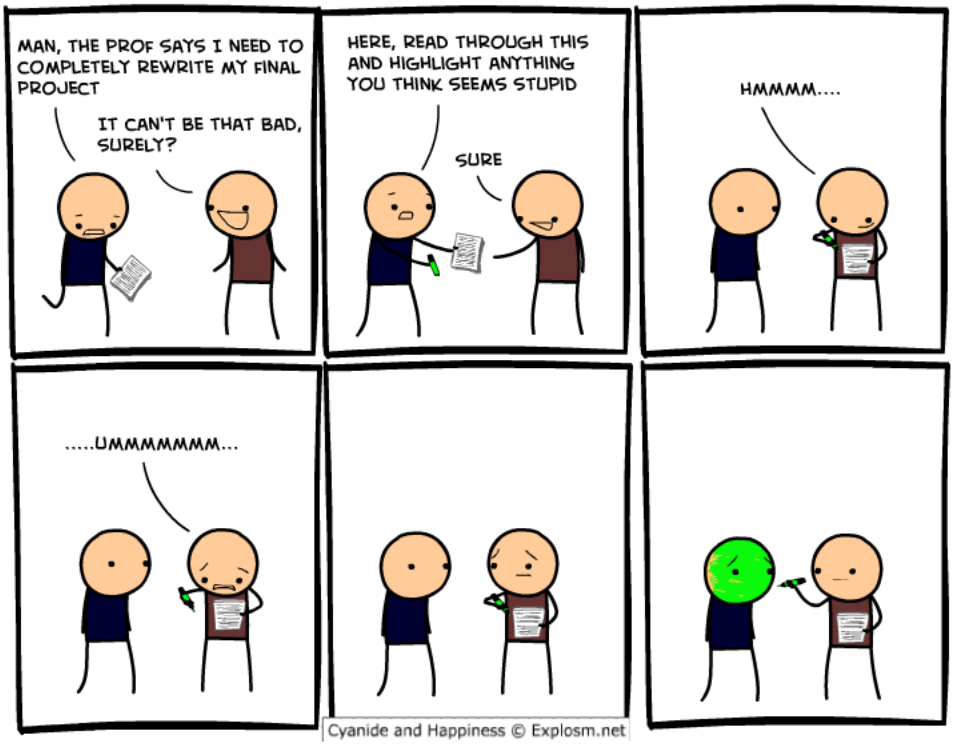
CS151 - Final Project
Spring 2013



http://www.explosm.net/comics/2083/

Overview

In this class we have looked at a number of techniques, and have examined a few in-depth in assignments. The purpose of the final project is to explore a topic we have examined (or not examined - but related to AI) 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, give a final presentation and review others' reports along the way.

The project should meet the following guidelines:

- Your project should relate to something we have or will talk about in class. I give some examples of past projects at the end of this document. Feel free to ask me if your idea is

appropriate. *Find something that your excited about and interested in since you'll be working on this for the rest of the semester!*

- You can develop a technique or an application, or both. For example, you might implement hierarchical clustering, or you might build a NB spam filter.
- You *must* evaluate the success of your approach. Even if you develop a technique, you need to find some reasonable way of determining how well your approach works. Generally speaking, the most successful evaluations will compare multiple different techniques, for example something we say in class vs. an improved variant or a baseline approach vs. an improved approach.
- 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
4/16	In-class project proposal
4/19	Project proposal write-up
4/26	Status report 1
5/6	Status report 2
5/10	Paper draft
5/14	Reviews
5/20	Final paper and presentation

Project proposal presentation [5 points]

Your first task will be to come up with your project group and figure out what you'd like to work on.
In class on Tuesday you will have 2 minutes to present the following information (*make sure you've thought about each of these things*):

- Who is on your project.
- What do you plan to do for your project:
  - Brief overview of what you plan to investigate, e.g. the application area, the technique, etc.
  - At least one "question" you hope to answer by doing your project (e.g. "What is the best initialization technique for k-means?" or "How does Theta* search compare to A* search?")
  - The data you will use.
  - How you plan to evaluate your project.

- Any other relevant information.

We will then have a few minutes of discussion from the class on the topic.

Project proposal write-up [10 points]

Submit online as .txt or .pdf by 6 pm as *proposal*

Your project proposal should be a 1-2 page write-up with clear section headings containing the following information:

- **Team:** Members of the team. I'm strongly encouraging groups of 2 or 3. If you want to work solo, please come talk to me.
- **Summary:** A one paragraph description of your project including:
  - What you hope to accomplish *and* how you will evaluate your approach.
  - What "question(s)" you plan to answer by doing this project.
  - How you will evaluate your approach. We've seen a few ways of evaluating approaches for both the NB classifier and k-means clustering, but come talk to me if you're having problems thinking about this part.
- **Resources:** 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, code provided with the book, 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.
- **Background:** Two papers in the literature (full citation) that tackle the same problem *and* a sentence or two summary of each of the papers.

Status reports [10 points each]

Submit online by 6pm on due date as *status1/status2*

Your project proposal *must* include the following (make explicit headings):

- **Members:** Names of team members
- **Summary:** A one paragraph summary of what was accomplished since last checkpoint.
- **Results:** A quantitative result:
  - Status 1: One or more numerical results. This could be some analysis of your data, a preliminary result from your system, etc.
  - Status 2: One or more results using your final evaluation metric.

- **Problems:** Any problems/issues that have arisen that might keep you from finishing your project.
- **Hours:** The number of hours each person put into the project since the last checkpoint
- **Code:** A snapshot of your current code-base. You may submit this as a link to an online repository (e.g. GitHub) or just a directory of code.

This is not meant to take you a long time, but please do spend a little bit of time putting this together.

Paper [90 points]

Draft due by 6pm online in .pdf
Final paper due at beginning of exam period both online and printed out

Your paper must be complete enough for me (or anyone else) to fully understand what you did. 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 AAAI workshop or symposium in the future.

We will use the AAAI paper format:

`http://www.aaai.org/Publications/Author/author.php`

The website has templates for both latex and word, either of which are fine. You paper will be short, specifically **no more than 3 pages** (you may have it a up to 4 pages if you have *many* images/figures). 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).

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

- 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 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.

- References: Your paper should have citations and must include *at least 3 citations*.

Reviews [10 points]

Due by 6pm as .txt submitted online

Each of you will be asked to review 1-2 papers written by other groups in your class. Your review will be similar to previous ones we've done in class, but should be easier since the papers should be shorter and should be on topics you're familiar with. For these reviews it will be important that you give very specific feedback about things that were good/bad/unclear in the paper since the authors will use your feedback to improve their final paper.

Presentation [20 points]

(All presentations will be during our exam period on 12/15)

Each group will give a short (10 minute) 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 4 weeks to accomplish the tasks, which should be plenty of time if you stay on top of things.

- Project proposals - Meets specifications above.
- Status reports - Meets specifications above. How much work was accomplished during the time period?
- Paper draft (10 points) - Do you have a reasonable draft in place for other students to review?
- Paper reviews
- Project and paper
  - The scope/difficulty of your project.
  - 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
  - Paper meets specifications above
  - The quality of your write-up, including addressing reviewers' comments
- Presentation
  - Covered content
  - Organized and well-prepared
  - Presentation style

Past Projects/Project Ideas

- Improved mancala player: examine improvements from other game playing components (e.g. end-game table or transposition table), learn weights using a machine learning technique for evaluation function
- Examine different search algorithm performance on some tasks (for example, A* vs. Theta*).
- Play with games with different characteristics: games with chance, unobservability (blind tic-tac-toe, stratego), games with betting
- Compare local search methods on an application
- SPAM identification/detection
- Compare document classification techniques (NB vs. SVMs vs. ?)
- Play with some machine learning approaches (`http://archive.ics.uci.edu/ml/` has a bunch of data sets)
- Implement and evaluate other clustering techniques: EM-clustering, spectral clustering, hierarchical clustering.
- Tackle other puzzles/games: sudoku, minesweeper, ...
- Implement a learning algorithm on neural nets.