CS159 - Final Project
Spring 2023

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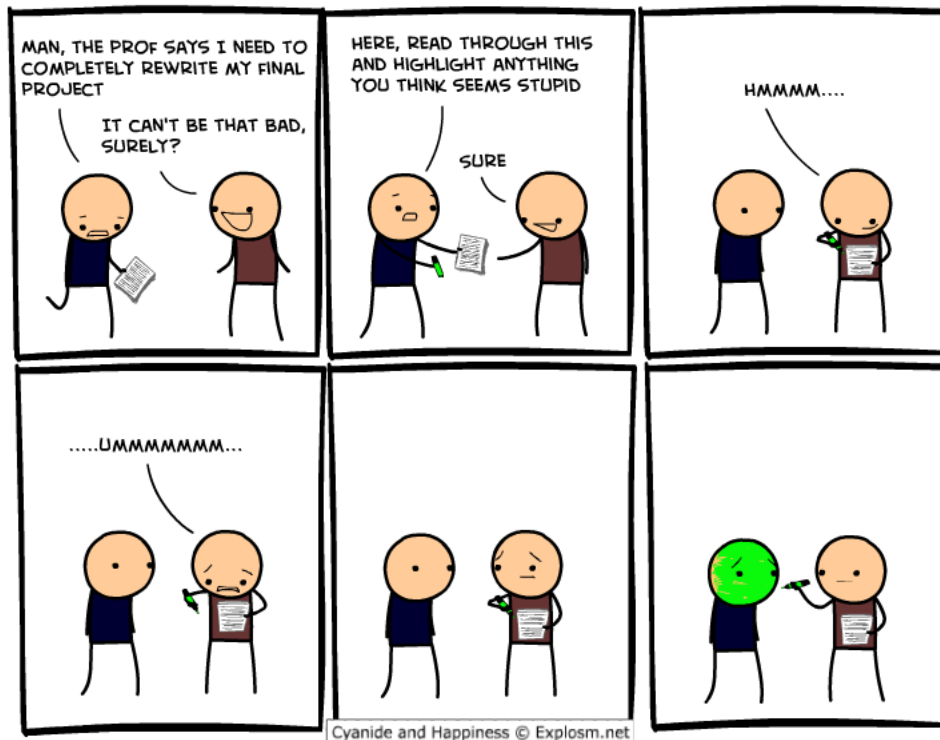
Overview

In this class we have looked at a number of NLP applications and techniques, and have examined a few in-depth in assignments. The purpose of the final project is to explore in more depth a topic we have examined (or not examined - but related to NLP) that is interesting to you.

Read through this entire handout to make sure you understand what is expected of you throughout this project. For example, as part of your status report and final project report, I'm asking that you log all the hours that you work.

The project should meet the following guidelines:

- Your project should relate to something involving NLP. I gave you a few ideas in class, but I encourage you to be creative. Feel free to ask me if your idea is appropriate. *Find something that you're excited about and interested in since you'll be working on for the last part of the semester!*

- Your project *must* include a solid experimental evaluation. This should involve picking an evaluation metric that is appropriate for your problem domain, tuning any parameters that the model may have appropriately (e.g. smoothing parameters) and using appropriate data for evaluation.

- Your project must have an ethical consideration to it. Either the project needs to be related to ethics or you need to have some discussion of the ethical considerations of your project in the writeup and presentation.

- Your project should be in a group of 2-4. If you'd like to do it solo, please come talk to me.

You may code in whatever language you would like and may (and are encourage to) use any external resources you would like including both code and data.

Schedule

date	time	description
4/17	in-class	Project proposal presentation
4/19	11:59pm	Project proposal write-up
4/26	11:59pm	Status report
5/1	11:59pm	Paper draft
5/3	in-class	Presentation
5/3	11:59pm	Final paper and code

Project proposal presentation [3 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 you will have **one** minute 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
    – What experiments you plan on running
    – The data you will use

- Any other relevant information

We will then have 1-2 minutes of discussion from the class to help you finalize any ideas.

Grading

Meets specifications above.

Project proposal write-up [10 points]

Your project proposal should be a 1 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 plan do to for the project. Be as specific as possible!
  - What experiments you will run and what metrics you will use for evaluation.

- **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.

Status report [10 points]

Your status report *must* include the following (make explicit headings):

- **Members:** Names of team members

- **Summary:** A set of bullet points enumerating what has been accomplished so far.

- **Results:** One or more results. This could be some analysis of a data set, a preliminary result from your system, etc.

- **Ethics:** How ethics will be incorporated into the project? The project needs to either be related to ethics or you'll need to have an ethics discussion in the paper/presentation. Here, you just need a sentence or two showing that you're thinking about it.

- **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 so far, including when each person worked (date/time).

- **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 effort putting this together.

Grading

Meets specifications above. How much work was accomplished during the time period? This is your work for the next 2.5 weeks and I expect you to be putting in regular time on the project. **Don't procrastinate!**.

Paper draft [20 points]

At this point, you should be done coding and should have run all of the experiments you plan to run.

You must submit an initial draft of the paper (see the paper requirements below). For the draft, *you must have completed the "Results" section*, which should show at least one graph or table summarizing your results along with a very brief experimental setup (of course, you can include more than this). Additionally, if your project isn't directly related to ethics, the "Ethics" part of the paper must have a complete draft, i.e., a discussion of the ethical implications of your project.

Grading

I expect you to be done with your experiments at this point. You will be graded on a) that you have provided the completed results for your task and b) that your results are finalized. I will be comparing your draft results to your final results and they should be roughly the same.

Paper and code [80 points]

Code: You may submit this as a link to an online repository (e.g. GitHub) or just a directory of code.

Paper: You should submit a short paper meets the following specifications:

– Be *at most* 2 pages long, not including references and the appendix on time spent.

– Follows the ACL formatting guidelines[1]

---

[1]See the "Styles and Formatting" section for `https://2023.aclweb.org/calls/style_and_formatting/`.

– Includes an appendix with two things for each member in the group: 1) hours worked and 2) 2-3 sentences describing what contributions that member made to the project.

– The paper should include the following:

  - An introduction. *Briefly* describe the technique/problem/application that you investigated.

  - Your experimental setup. What data did you use? How did you setup your experiment? What did you use for evaluation? We have talked about evaluation a lot in this class, so an important component of this project will be that you have setup a proper experiment with a proper evaluation metric.

  - Your results. These should be stated concisely and should include supporting tables, graphs and figures. You must have at least one graph/table summarizing your results. Think about whether a graph or a table (or both) would better show your results and use the appropriate one.

  - Ethical considerations. If you're project is not directly related to ethics, you need to have an discussion of the ethical considerations of the project.

  - Conclusions. Summarize your findings.

Grading

- The scope/difficulty of your project.

- How creative is your project/experiment?

- How complete is your project? Did you accomplish what you set out to do?

- Paper meets specifications above.

- Quality of evaluation (not how well did your system do in an absolute sense, but did you setup the evaluation properly).

- The quality of your write-up: spelling/grammar issues, how well are the results presented, how good is your analysis, did you use a proper table/figure, etc.

Presentation [10 points]

Each group will give a short presentation (~7 minute–I'll give the exact timing once I see how many groups there are) of their work during the last day of class. Your presentation must include a visual aide (e.g. slides). You may either use your own laptop, or, if you want to use my laptop, e-mail me your slides by 10am the day of the presentation

Your presentation must include the following information:

- What you investigated and the problem setup

- Your approach/technique

- Your experimental setup and evaluation

- Your results

- Ethical considerations (as appropriate)

Grading

- Covered content described above

- Organized and well-prepared

- Used slides/visual aides appropriately

- Presentation style

http://www.xkcd.com/1443/