

CS 505 Final Project - Fall 2016

Example sources of data.

A website with all the publicly available datasets from the Government under the Open Data Policy.

<http://www.data.gov/>

Public data from cities of Boston and New York (including 311 service requests and more).

<https://data.cityofboston.gov/>

<https://data.cityofnewyork.us/>

Other interesting datasets:

<http://aws.amazon.com/datasets/>

<https://github.com/caesar0301/awesome-public-datasets>

The following link contains links to other dataset sites (including the ones that I have above)

<http://www.kdnuggets.com/datasets/index.html>

Lots of interesting social network data is at <https://snap.stanford.edu/data/>.

1. Project proposal

Template:

- 1) Discuss the dataset and their nature. How you got them, how they have been collected, what are their main characteristics, if you need to pre-process them, etc.
- 2) Expected analysis on the dataset (this of course is something that may change during the project.) What kind of techniques you plan to use.
- 3) Application. Which application is associated with the dataset and how it will benefit from your analysis.
- 4) Expected results (that also can change after you do the actual analysis).

Data collection:

If you plan to scrape your data, provide some results of data collection.

2. Mid-progress report

Template:

Title

Your Name(s)**Abstract**

3-5 sentences. Describe the main idea of the project and the expected results

Introduction

3-4 paragraphs. Describe the main application/motivation, the dataset, your approach and the experimental results at a high level (details in the next sections).

Technique (that will be) used in this project (e.g., what clustering, classification, or regression method and why)

2-4 paragraphs. Describe the main ideas of the techniques that you (will) use in the project.

Datasets and experiments.

3-5 paragraphs (or more). Describe datasets and experiments.

Results and Discussion

1 small paragraph for now with initial results if any or what you expect.
To be completed with the final report.

Conclusion

1 small paragraph.
To be completed with the final report.