
CSE 146: Data Analysis (mini!) Project

— Fall 2020 —

Outline

1. You may work in groups of 2-3. You should form your group in the next week, by Nov 20.

Use slack, etc. to forms groups, and we'll give time next week for you to meet folks as well.

2. You will be given option of several datasets to work with. You may use one of these, or with approval, a dataset of your choice.

<https://drive.google.com/file/d/1Ci6aSu7TjExtHehNhObM-CXE-OfVo7Le/view?usp=sharing>

Outline

For your project, you will do some empirical analysis.

You may:

1. Compare different ML algorithms, and study the different kinds of errors
2. You may compare the fairness of different ML algorithms
3. You may look at causality, and compare different causal models
4. You may look at privacy
5. You may look at explainability and interpretability
6. Other options are possible, with approval

You may use off-the-shelf packages, and these should be documented in writeup

Outline

Deliverables

1. Project Team and name: Friday Nov 20
2. (Brief!) Project progress reports Mon Nov 23, Mon Nov 30, Mon Dec 7
3. Virtual Poster Session, **3 Minute** Project Video due Tue Dec 8, 5PM
4. **4 Page** Project Report due Fri Dec 11, 5PM
5. Both the video and report should be structured as follows:
 1. Describe problem setting
 2. Describe data
 3. Describe evaluation & results, including tools used
 4. Discuss issues, limitations, future work

