

Final Project Guidelines (Modules)

Overview

The final project is your opportunity to work with your fellow students in applying techniques learned in class to an open problem in one of several domains.

Students will work on one of several projects proposed by various faculty members. The topic choices are:

- A. Machine Learning for Medical Diagnosis
- B. Searching for Near Earth Asteroids
- C. Predicting Citation Networks in Legal Texts
- ~~D. Scene Detection for Design [canceled]~~
- E. Applied Differential Privacy on U.S. Census Data
- F. Microbiome Dynamics in Health and Disease
- G. Visual Recognition of Images on Ancient Greek Vases
- H. Dynamics of Disease Transmission and Human Behavior

More details can be found under [Modules](#) on the public course page.

Posting on Ed is a great way to find potential partners!

Projects for 209A teams will be graded at a higher level. These projects must implement a method that was not discussed in class and so will require some outside reading. Therefore, students enrolled in 121A or 109A should only join a team with students in 209A if they are comfortable with being graded at this higher level (all 'mixed' groups will be graded at the 209 level).

Final Deliverables

1. **Code Report** - You are expected to submit the code you developed as part of the course project. The commented code should be provided in report format. This means that each group in a Jupyter notebook should explain—in a clean and concise report fashion—how they proceeded at every step and coding/methodology choices. The code report should have a structure that consists of an introduction, body and conclusion.
2. **Presentation Video** - Record a 6 minute video of your group presenting your project. This should include slides and all group members should participate. You can use Zoom to record the presentation and submit the link to the recording. Videos should be no longer than 6 minutes; content after this time limit will not be viewed. (Please, Do not speed up the video!). Submit the link to the video.

More details on the rubric outlining expectations of these deliverables to follow.

Milestones

1. **Group Creation and Project Topic Selection (due 4/3)**
Form a group of 3-4 students and submit your top 4 project topics from the list above. You may submit alone if you have no group and we will try and find one for you. But the chances of you getting one of your preferred projects are greatly increased when submitting in a group.
2. **Preliminary Exploratory Data Analysis (due 4/20)**

The purpose of this milestone is to ensure that all groups have access to their data and are able to get it into a form which can be manipulated.

Submit a 2 - 3 page revised project statement and EDA (can be created using Latex, word processing software, etc.) and an accompanying Jupyter notebook (that was used to create the visuals). Your 2 - 3 page submission should include:

- A description of the data: what type of data are you dealing with? What methods have you used to explore the data (initial explorations, data cleaning and reconciliation, etc)?
- Visualizations and captions that summarize the noteworthy findings of the EDA
- A project question based on the insights you gained through EDA.
- A baseline model or *clear* plan for its implementation

3. TF Meeting (due 5/2)

- a. Meet with your group's TF by no later than this date

4. Final Project & Presentation Video (due 11:59pm EST 5/6)

- a. Code Report (See "Final Deliverables" section above)
- b. Presentation Video

5. Peer Evaluation (due 11:59pm EST 5/9)

- a. This is a form where you will have the chance to assess your own as well as your group members' contribution to the final outcome.