CS 109B / Stat 121B/ AC 209B/ CS-E 109B
Glickman, Protopapas

Final Project Guidelines

Overview

The course final project is your opportunity to work with your fellow students in exploring a topic and applying techniques you learned in class towards understanding it better.

1. **FAS students** have the chance to participate in the Modules part of the class - which starts on 4/22 - and work on projects by either choosing one of those proposed by us (D, E, G), or by the Module instructors (A, B) or preferably proposing your own (C). The project choices are:

   A = **Module 1**: A Real-Bogus Classifier

   B = **Module 2** : Microbiome

   C = Your Custom Project

   D = Cancer Diagnosis

   E = Astronomical Classification

   G = Million Song Database

2. **DCE students** will have the chance to work on projects, by either choosing one of those proposed by us, or preferably proposing their own. The project choices are:

   C = Your Custom Project

   D = Cancer Diagnosis

   E = Astronomical Classification

   G = Million Song Database

You will work in groups of 2 to 4 students. If you are looking for project partners, we suggest you post a request on Piazza. If you do not have group partners, submit Milestone #1 alone, and we will place you in a group.

Notes on mixed groups (109/209/DCE):

Students enrolled in 109b may form teams with students in 209b and/or DCE. The grading policy for projects involving a 209b student, even if the other team members are not in 209b, is

that we expect a deeper analysis and methods that go beyond the methods used in class. In other words, if any member of the team is in 209b, the project will be graded as if all members are in 209b.

Also, groups that involve at least one non-DCE student will produce **both a Final Report and a Lighting Talk**. Groups with only DCE students are responsible for submitting only a Final Report.

Final Deliverables

- **Deliverables for groups with only FAS students:**
  - **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. A good template for this deliverable is https://github.com/Harvard-IACS/2019-computefest/blob/master/Wednesday/auto_encoder/VAE_Solutions.ipynb, however please realize that we expect more detailed descriptions.
  - Lightning talk. See class Github for Ignite Talk Guidelines. You will present the talk on 5/13 or 5/14. Details to come.
- **Deliverables for groups with only DCE students:**
  - Jupyter notebook with relevant **commented** code.
  - Final Report.
- **Deliverables for mixed groups (FAS/DCE):**
  - **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. A good template for this deliverable is https://github.com/Harvard-IACS/2019-computefest/blob/master/Wednesday/auto_encoder/VAE_Solutions.ipynb, however please realize that we expect more detailed descriptions.
  - Lightning talk given by the FAS students in the group.
  - Final Report submitted by the DCE students in the group.

Milestones

1. **Milestone #1 (due Wed. April 17) : Group Creation and Project Selection** (2 points)

   - The purpose of this milestone is to make a group of 2-4 students. Please indicate which group members are FAS or DCE students. If you are a group consisting entirely of FAS students, let us know **one** of the following: (a) your choice of Module, OR (b) your choice of an existing project listed on the project GitHub, OR (c) your proposal of a custom project.

   - If your team has a custom project to propose, upload a pdf or text file containing your project description. Use the template provided in GitHub.

   Follow the instructions in the assignment. In particular, please see how to form groups when submitting on Canvas.

2. **Milestone #2 (due Sat. April 27th): Literature Review and Preliminary EDA** (15 points)

   -Project statement. Based on the project description, state a well-defined question that you'll address in the project.

   - Literature review: Please provide basic background and relevant references on your project topic. A paragraph description should suffice.

   - Preliminary EDA. It is important that at this stage you are already working with the data for your project, and should demonstrate some basic exploratory results working with your data. Explain your preliminary data exploration. You may include simple visualizations or just a verbal description.

   Follow the instructions in the assignment.

3. **Milestone #3 (FINAL due Sunday, May 12): Final Project due** (80 points)

   Submit your final project on Canvas.

   - **Deliverables for groups with at least one FAS student:**
     - **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. A good template for this deliverable is https://github.com/Harvard-IACS/2019-computefest/blob/master/W

ednesday/auto_encoder/VAE_Solutions.ipynb, however please realize that we expect more detailed descriptions.
    - A lightning talk. See class Github for instructions on how to prepare such a talk. You will present the talk on 5/13 or 5/14. Details to come.
- **Deliverables for groups with only DCE students:**
    - Jupyter notebook with relevant **commented** code.
    - Final Report.
- **Deliverables for mixed groups (FAS/DCE):**
    - **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. A good template for this deliverable is https://github.com/Harvard-IACS/2019-computefest/blob/master/Wednesday/auto_encoder/VAE_Solutions.ipynb, however please realize that we expect more detailed descriptions.
    - A lightning talk given by the FAS students in the group.
    - Final Report submitted by the DCE students in the group.

Follow the instructions in the assignment.

4. **Peer evaluation form (due Sunday, May 12)** (3 points)

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.