ECE 1786:
Creative Applications of Natural Language Processing

Lecture 10

November 22, 2022

Agenda

Last Day:

- How GPT-3 responds well to human intent
- Using GPT-3 as a processing pipeline
- The GPT-3 API – fine-tuning & embedding

Work-in-Flight: The Project!

Today:

- Course logistics and Deliverables
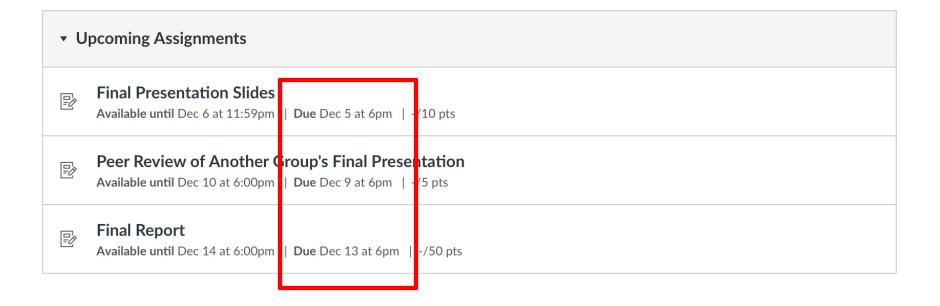- Open time to discuss/consult on projects

Project Timeline and Deliverables

Date	Item
04-Oct-22	Project Discussion in Class
11-Oct-22	Team Forming Deadline
27-Oct-22	Approval-in-Principle of Project Topic complete
31-Oct-22	Project Proposal Document Due
31-Oct-22	Project Proposal Slides Due
01-Nov-22	In-Class Proposal Presentations + Extra Class in Evening
21-Nov	Progress Report Due
05-Dec	Final Presentation Slides Due
06-Dec-22	Final Presentations - **Extra Class in Evening**
09-Dec-22	Peer Review Due
13-Dec-22	Final Report Due

See Quercus for Details Presented Today

▼ Upcoming Assignments

Final Presentation Slides
Available until Dec 6 at 11:59pm | **Due** Dec 5 at 6pm | -/10 pts

Peer Review of Another Group's Final Presentation
Available until Dec 10 at 6:00pm | **Due** Dec 9 at 6pm | -/5 pts

Final Report
Available until Dec 14 at 6:00pm | **Due** Dec 13 at 6pm | -/50 pts

Progress Report

Progress Report

- All reports received in good order last night
    - A quick glance at them indicates some good progress!
- I will review them as quickly as possible, and provide feedback to be helpful with your projects
- Will send them back to as they are graded, not wait till all done

- I also checked the Github repositories, and most had something pushed to them
    - It was expected that you would show activity there for this report
    - For the final report it is a requirement that use the repo, and it should show continuous activity, not just one or two pushes.

Final Presentations

Slides Due December 5 at 6pm

Presentations On December 6, 2022

Note: Extra Hours for Final Presentation

- There are extra course hours for the Final Presentations
  - Similar to the proposal
  - **Two** sessions on December 6th
  - 10am -12 noon and 6:00pm-9:00pm in GB 221
- You will present during one of these sessions
- You will be do a peer review in the other session
  - Also want everyone to see all presentations
- Be sure you're available 6:00pm-9:00pm December 6
  - Let me know of any hard constraints that prevent by Nov 30
  - Will provide dinner again

Final Presentations

- **Maximum 6 Minutes**
  - More time than usual!
- **Must Be Self-Contained**
  - Meaning: assume audience has **no** prior knowledge of project
  - the presentation must stand alone
- Who is the audience?
  - This class, TAs, Instructor
  - **And** if you publish the video: Your future employer, Your friends, parents

Video Recordings

- **We will record the final presentations**
  - Will be edited into separate videos, with slides inserted
  - You will have the choice as to whether recordings will be posted publicly or not.
  - The entire group must agree to posting
  - (as well as the final report and the software repository)

Required Content of Final Presentation

1. **Goal & Motivation**: What & Why

   – Third time, keep improving!

2. **Data and Data Processing**:

   – What data did you use or collect?

   – What does it look like?

   – What data processing did you do and why?

   – Was their any unusual issues with the data?

3. **Model and Software**:

   – Describe the model(s) did you built and tested

   – Were there any issues in training?

   – Describe any other software needed.

Required Content of Final Presentation

4. Quantitative and Qualitative Results:

  – How are you measuring success and comparing performance?

  – Discuss and interpret your results.

5. Discussion and Learnings:

  – Do your results make sense intuitively, or were the results surprising in some way?

6. Reflect: What would you do differently in a similar project, based on your experience in this project?

Peer Review of Final Presentation

Peer Review

Like Proposal, you'll be assigned to a group during the opposite period (morning/evening) of your presentation:

Answer these questions:

1. State the goal of the project in your own words.
2. Say what was the most interesting & why.
3. Give one suggestion to improve their final report
4. Provide feedback on the quality of the oral presentation

- 300 words, due **Friday December 9 at 6pm**, 20% penalty if late

Final Report

Due December 13, 2022

Final Report

- A summary of the project

- Describes what you have done and why, what your results are, and an interpretation of your results.

- Word limit 2000 words

- 1% penalty for every word in excess of the 2000 limit.

- You must count the words in your document, compute the penalty, and put it on the front page. If count is missing, a grade will be deducted

Submission of Report and Software

- Submit final report as a group on Quercus to the 'Final Report' assignment
  - due **Tuesday December 13th**, at 6pm.
- Commit your final source code to your course Master GitHub repository by that same time. There should be no commits to the Master branch after the deadline.
- Usual 1 hour grace period/20% penalty up to 24 hours late, 100% penalty after that

Permissions

On a separate page, at the end of report:

- each member of the group must indicate if they will grant or deny permission to post the project on public website
- Both members of the group must say yes for each permission to be granted. (i.e. everyone has a veto)
- You can also wait to see the video before deciding
- Table not included in the word count

Team Member	Post Video?	Post Final Report?	Post Source Code?
Person 1	Yes/No/Wait till see	Yes/No	Yes/No
Person 2	Yes/No/Wait till see	Yes/No	Yes/No

Final Report Sections

1. ## Introduction (2 points)
    - A brief description of the goal and motivation of the project.This should include why the goal is interesting or important.

2. ## Illustration / Figure (2 points)
    - A figure or a diagram that illustrates the overall model or idea of your project. The idea is to make your report more accessible, especially to readers who are starting by skimming your work. You are graded on the design and illustrative power.

3. ## Background & Related Work (2 points)
    - A description of 1-2 related works in the field, to provide reader a sense of what has already been done in this area, e.g. papers or existing products/software that do a related thing.

Document Sections, cont'd

4. Data and Data Processing (4 points)

   - Describe the data that you collected, how you collected it, and how you cleaned it. Show some statistics and examples of your data.

5. Architecture and Software (4 points)

   - A description of the final neural network model architecture. Do not describe all the intermediate models that you tried. Instead, present the model (or models) whose quantitative results you will show. These should be your most interesting models. Describe any other software you created for the project.

6. Baseline Model (2 points)

   - If you chose to make a baseline, describe it. If you didn't have a baseline, this grade will be included in the results section, where you'll need to describe how you measured success carefully.

Document Sections, cont'd

7. Quantitative Results (4 points)
   - A description of the quantitative measures of your result and why it is appropriate. What measurements can you use to illustrate how your model and software performs?

8. Qualitative Results (4 points)
   - Include some sample inputs and outputs of your model/software that illustrate when the model is working well and when it is not. The qualitative results should also put your quantitative results into context (e.g. Why did your model perform well? Is there a type of input that the model does not do well on?)

9. Discussion and Learnings (4 points)
   - Discuss your results. Do you think your model is performing well? Why or why not? What is unusual, surprising, or interesting about your results? What did you learn, and what would you do differently when starting another, similar project?

Document Sections, cont'd

10. Individual Contributions (10 points)
    – Each partner should write a description of what their contribution to the project. This must contain specific details such as:
      • collected dataset X
      • hand labelled X data samples
      • was responsible for the training of the baseline software
      • wrote the gradio implementation of the user-facing side of the project
    – just a few examples, there are many other possible activities!

■ The project Github repository should show commit activity that aligns with the work stated.

Two Other Aspects of Final Report Grade

For your information (does not require anything to be written in report)

Project Difficulty/Quality (4 points)

■ A measure of how "difficult" the project is, and how well your model performs given the difficulty of your problem. If your problem is more difficult than what one might expect, you should clearly articulate why in the body of your report.

■ There are a variety of ways to increase your project complexity, even after the fact. For example:

  – Try different models,

  – Explore and expand your results.

  – Explore and expand your data

  – Use data augmentation techniques, and discuss results

  – Explore how to make your model smaller/faster

Structure, Grammar & Mechanics (8 points)

- The document should be easy to understand, be grammatically correct and well-written.

Questions?

Project Consultations