

ANLY-580 Course Project

Format: Groups of (≤ 3)

Due date: Dec 08

Grade: 250 pts (25%)

Project description

The purpose of the ANLY-580 final project is to give you experience working on a (quasi) open ended problem in NLP. There are two paths: (1) select from one of a predefined set of topics, and (2) choose your own custom project. The grading criteria will largely be the same for each of these paths, however, the success criteria for custom topics that are more ambitious in nature will be adjusted as needed. The project has two major milestones, first is selecting a group and choosing a topic; this included defining the scope, goals, and deliverables for your project and getting approval from the instructor and TAs. The second milestone is the final submission.

Project options

1. Default projects

- Question answering system using model pretrained on SQuAD
- Keyboard autocorrect / autocomplete
- Twitter hate speech detection
- Text summarization system
- News headline classification
- Topic modeling of a specific text corpus / research field
- Research paper retrieval system
- Product review sentiment analyzer
- A task-oriented chatbot

\ 2. Custom project:

This option allows you to choose your own project topic. Students are free to explore their own project idea provided that is properly scoped with clear deliverables (see note below). For those students conducting graduate research, it is ok for there to be overlap between your research topic and this project, but the work must be something that you haven't explored yet.

The scope and deliverables need to be established and agreed upon with the instructor by the project selection deadline. Office hours, emails, and slack are all acceptable forms of communication. If communicating over slack or email, be sure to include all team members, TAs, and the instructor in the thread!

Deadlines

Item	Due Date
Group selection	Oct 14
Project proposal	Nov 01
Presentation / Demo	Dec 08

Deliverables

Item	% Grade
Group selection	5%
Project proposal	20%
Presentation / Demo	75%

Rubric

1. Group selection: This is the easiest 50 pts you'll earn in this course. There are two deliverables:
 - a. Submit group selection via email/slack to the instructor AND TAs no later than 11:59PM EST on Oct 14th.

- b. Create a separate dedicated github repository and invite the instructor + TAs as collaborators. Please place all project material here, and make sure to keep the repository private.
2. Project proposal: The project proposal will be a written outline that details the following items (this needs to be discussed with the instructor during office hours prior to submission):
- a. High level description of what you plan to do, and why you think it is interesting.
 - b. A detailed description of the methods you will use to achieve this:
 - What modeling approach do you intend to use?
 - What data do you intend to use?
 - How will your system be evaluated and what are the evaluation criteria?
 - Are there any special computational/hardware considerations?
 - What are the biggest unknowns that might dictate the success or failure of this project?
 - c. How will the results of your work be presented? Will this be a live demo, a written report, a slide deck + oral presentation? Any of these are acceptable! Demos can be given along with reports/presentations.
3. Final submission: The final submission will be graded on a number of axes. For a project that is modestly ambitious in it's end goal, heavy emphasis will be placed on the quality of the demo/presentation/report, whether or not the system is thoroughly evaluated, and how close you come to achieving the deliverables stated in your proposal. Sometimes things don't always go as planned, and you should have this expectation from the start. When this happens, make sure to demonstrate that you understand the underlying issue(s); this will be considered when we assign grades. For more ambitious projects, heavier focus will be placed on individual results. For example, if your final project is to reimplement a popular transformer architecture from scratch, or tweak some existing modeling approach in unique way, then success in any one area (for example just getting a model to train) will be considered much more heavily. The intention here is to allow everyone to get real world experience working on a real problem within their comfort zone.

Note: If you are having a hard time coming up with your own project idea, we recommend you choose one of the default projects (or something closely related) and we can provide scaffolding to get you started. Also, if you need assistance with finding a partner/group, please contact one of the TAs or the instructor ASAP.