CS 441/541: Machine Learning

Fall 2020

A. Rhodes

**Group Final Programming Project**: In small groups (3-5 people, no smaller please), you will pursue an ML-related research project (with a reasonable breadth and depth) in an intermediate to advanced topic that extends beyond the bounds of the core topics formally covered in lecture. Together, you and your group members will submit a short write-up of your project, including, importantly, code and cogent data results. **During finals week** (over two sessions) **your group will give a short 10-15 minute presentation summarizing your work**.

I want to keep the topic choice flexible so that you can pursue an area of interest, but here are some suggested areas: Computer vision (classification, etc.), medical applications (e.g. disease detection), NLP, computational creativity (e.g. music, art), clustering problems, deep learning, agent-based search, etc..

In selecting a project you should, of course, do something that will be fun for you. It's also important, however, that your project involve a significant ML component and that it be of the right scope. In order to help ensure that that will be true, you should talk to me and/or Li-Yun about what you hope to do. Some of you have already done so, and I very much look forward to seeing how your projects turn out.

After I informally approve your project idea, **you should prepare a written proposal, approximately half of a page in length**. Turn in one proposal per group. The proposal should include the names of all group members. It should begin with a high-level description of the project. In this section you should introduce the project and your overall goals. You should also discuss the relevance of ML to the project. You might include other information as well, such as what you find motivating or compelling about the project, for example.

Next the proposal should expand on the description of what you plan to do. When in doubt, provide more detail rather than less. In particular, be sure to itemize the various components that you will need to implement, and tell me which programming language you will use.

Be sure to provide references for any sources (including software) that you are using. Also remember to read and cite research articles. Any ML techniques or algorithms you implement will have their foundation in ML research, either current or past.

*__Please turn-in your proposal in class on Monday, 2/24 – it won't be graded__.* This assignment is mostly for your benefit to ensure that you're on track with the project.

*Project Presentations and Demos (Finals Week, TBD)*

We will schedule presentation sessions at the time of official final, as well as an additional time during finals week TBD.

The presentations and demos are intended to be formal presentations of your work. They should be well-organized with appropriate visuals (slides, work on the board, as appropriate). Be sure to build time into your schedule to prepare the presentation. Demonstrate your respect for them by presenting a thoughtful and well-planned talk and demo. **The presentation will be between 10-15 minutes** (no longer, due to time constraints).

**Your final deliverable will be a paper describing your project**. In the paper you will discuss your project in detail, situate it in the context of others' work, and evaluate your progress as a team and as individuals. **The paper should be 5-7 pages long (this length requirement should include results)**.

If your project involves building a system, you should clearly state your goals for the system, give an overview of its architecture, and describe in some detail the algorithms selected to implement the ML modules of the system. You should also cite relevant research and other articles to situate your project in the broader context of work in this general area. In assessing your system, you will need to clearly articulate exactly what you are evaluating, how you have gone about performing evaluation, and how it stacks up. For instance, say you've implemented a system that "reads" long articles and generates one-paragraph summaries of them. What inputs can your system handle? What does it produce? Have you built in any special domain knowledge that restricts the topic of the articles? Does your system identify the most relevant information? Does it organize it well in terms of the flow of language? Is it grammatically correct? What are the limitations of your system?

If your project includes an empirical comparison of algorithms, you should clearly specify what you are comparing, what you are measuring, the methodology you developed for the empirical investigation, and any data sets or existing code you used as part of the work. You should discuss -- critically and analytically -- the results of the investigation.

Finding Relevant Research Articles and Resources (not required – but encouraged)

Here are a few recommended resources for research articles and project ideas: ArXiv ,Google Scholar, and Kaggle. If you are having difficulty finding a topic for your project, I also recommend implementing/extending any of the relevant programming assignments from the exercises of our text.

**You will turn-in your written project materials the day that you present your work**.