Project Details

This project provides an opportunity for you to: (1) gain experience implementing Machine Learning models and (2) try Machine Learning on problems that interest your team. Teams will write a proposal to solve or implement a challenging machine learning problem. Once the proposal is approved the team will begin to implement and research a solution to this problem or idea. Team sizes can only be 2 people unless the class has an odd number then one group of 3 can exist. This project will require you to turn in a final report pdf and a link to a repo on github.

Github Repo (40 points)

Students will turn in a link to a repository. The repository must have the code that you used to solve your problem and write the paper. Here are some things required for the repo

- Well documented README.md
    - List of python libs used and how to install them (requirements.txt)
    - Guide on how to run the project (only include a guide on how to run the main portion, you do not have to document how to run all the experiments you ran in order to write the paper)
- Documents folder with any images or diagrams used in the report
- Code files that were used to run experiments and solve the problem

Final Report Template

Your final write-up is required to be between 4-7 pages and structured like a paper from a computer vision conference (CVPR, ECCV, ICCV, etc.). Use this template https://docs.google.com/document/d/1778QnUi2KloO0HbC-T4dTYy8ei9KJ7HKY5akYbyil4U/edit?usp=sharing

Rubric (60 points)

I am not looking to see if you succeeded or failed at accomplishing what you set out to do. It's ok if your results are not "good". What matters is that you gain an in-depth understanding of Machine Learning as it relates to your project, and can clearly communicate that understanding through your analysis. Note that you must justify your design, implementation, and experimentation decisions using your knowledge and data. You should make claims about why you think the results turned out the way they did and perform specific experimentation (or gather relevant data) to justify your claims.

A former DARPA director, George H. Heilmeier, came up with a list of questions for evaluating research projects. I've adapted that list for our rubric.

Introduction / Background / Motivation:

- (5 points) What did you try to do? What problem did you try to solve? Articulate your objectives using absolutely no jargon.

- (5 points) How is it done today, and what are the limits of current practice?
- (5 points) Who cares? If you are successful, what difference will it make?
- (5 points) What data did you use? Provide details about your data, specifically choose the most important aspects of your data mentioned here: Datasheets for Datasets (https://arxiv.org/abs/1803.09010). Note that you do not have to choose all of them, just the most relevant.

Approach:

- (10 points) What did you do exactly? How did you solve the problem? Why did you think it would be successful? Is anything new in your approach?
- (5 points) What problems did you anticipate? What problems did you encounter? Did the very first thing you tried work?

Experiments and Results:

- (10 points) How did you measure success? What experiments were used? What were the results, both quantitative and qualitative? Did you succeed? Did you fail? Why? Justify your reasons with arguments supported by evidence and data. Make sure to mention any code repositories and/or resources that you used!

Additional: 15 additional points will be distributed based on:

- (5 points) Appropriate use of figures / tables / visualizations. Are the ideas presented with appropriate illustrations? Are the results presented clearly; are the important differences illustrated?
- (5 points) Overall clarity. Is the manuscript self-contained? Can a peer who has also taken Machine Learning understands all of the points addressed above? Is sufficient detail provided?
- (5 points) Finally, points will be distributed based on your understanding of how your project relates to Machine Learning. Here are some questions to think about:
    - What was the structure of your problem? How did the structure of your model reflect the structure of your problem?
    - What parts of your model had learned parameters (e.g., convolution layers) and what parts did not (e.g., post-processing classifier probabilities into decisions)?
    - What representations of input and output did the neural network expect? How was the data pre/post-processed?
    - What was the loss function?
    - Did the model overfit? How well did the approach generalize?
    - What hyperparameters did the model have? How were they chosen? How did they affect performance? What optimizer was used?
    - What Machine Learning framework did you use?
    - What existing code or models did you start with and how did these starting points help?

Note that at least some of these questions and others should be relevant to your project and should be addressed in the PDF. You do not need to address all of them in full detail. Some may

be irrelevant to your project and others may be standard and thus require only a brief mention. For example, it is sufficient to simply mention the cross-entropy loss was used and not provide a full description of what that is. Generally, provide enough detail such that someone with an appropriate background (in both Machine Learning and your domain of choice) could replicate the main parts of your project somewhat accurately.

References

Your report must include citations for relevant work including papers that inspired you or that you reimplemented, etc. The papers you cite in this section should be referenced within the paper itself. Note that this section does not count towards your page limit, so please be comprehensive.

Presentation

After the due date students must sign up for a timeslot to present the project to the instructor. During this time you will explain what you were trying to do and talk about your successes or failures. This will be where the instructor will ask questions and critique your project.