CS 482/682 Machine Learning: Deep Learning

Final Project Instructions

Important Dates

- Project proposals due: Friday, Nov 6th 2020, 11:59 pm
- Final reports due: Friday, December 11th 2020, 11.59pm
- Pitch & demo session: Friday, December 18th 2020, 2.00 - 5.00 pm (Eastern)

Objective

The objective of the final project is to make use of what you have learned during this course to solve a hard deep learning problem. The final deliverables are:

- A project proposal (<1 page, double column): This proposal should later be extended to a final report
- A written report of 2 pages double column maximum
- Two presentations: A pitch presentation (1 minute, should hype the audience and focus on the significance/impact and most noteworthy achievements) and a longer presentation ideally with a demo (~4 minutes) summarizing the more technical aspects of the project. The pitch will be presented in a plenary session while the longer presentation will be delivered to a smaller audience in breakout rooms.

We provide templates for these submissions. Of note: We will deduct points for use of baseline stretches or margin modifications.

You will work in **groups of 3**. Below we define a set of projects that you can choose from. Each project proposal must include:

1. A brief problem statement - what you want to work on and your objectives
2. A brief summary of the dataset: What is available and what will you use from it
3. At least one (ideally 2-3) papers that solve the task using (un)supervised learning
4. **Important**: An outline (can be bullet point) of how you attempt to solve the problem

Custom projects

We are happy to entertain “special” projects that are not currently listed in the set of projects to choose from. In this case, however, we ask you to attend at least one office hour well before the project proposal deadline to get a written “OK” from the teaching staff (instructor or TA). This is to ensure that the project has a reasonable chance for success. Also, select a fallback project from the set of suggested projects. Be prepared to present answers to the following questions:

1. What is the problem you want to address?
2. Provide a reference to at least one document (seminal paper with GitHub repository preferred) that addresses the same or at least similar problem. If the problem is only similar, clearly outline the similarity and difference.

3. Provide a summary of the dataset you will be using. The **dataset must exist and be usable for the envisioned project with minimal work** required. Dataset curation is important but **not** a deliverable of this assignment.

Suggested theme for custom projects:

-- *Deep Learning for Good* --

Technology and artificial intelligence are shaping our everyday experiences and society. Examples that demonstrate the ubiquity of the AI revolution range from personalized news feeds over task automation to precise decision making across various domains (many of which are sensitive, such as recruitment, policing, and medicine). As educated citizens and researchers in deep learning, we should embrace the opportunity to advocate for using technology for the better. Questions that could arise when thinking about deep learning for good may include:

- How can we develop fair deep learning models?
- How can we use deep learning to solve some of the biggest societal and environmental challenges?
- How can deep learning systems interact with humans?

Hints: Here are some hints on how to find such initiatives and other resources:

- https://lmgtfy.com/?q=ai+for+good
- Google's People + AI Guidebook: https://pair.withgoogle.com/
- Federal datasets: https://catalog.data.gov/dataset
- Googles brand new dataset search: https://datasetsearch.research.google.com/
- It could also be worthwhile to think about how people's lives could be improved on campus and beyond?

Goals

The goal of the project is to demonstrate the group's understanding of the tools, tricks, and challenges when training deep neural networks. Grading will reflect the quality of approach, rigor of evaluation, and reasoning about successes and failures. The final report should concisely summarize your findings and answer the following questions:

1. What approach did you take to address this problem, and why?
2. How did you explore the space of solutions - architectures, hyperparameters, training methods etc.?
3. How did you evaluate the performance of the approach(es) you investigated?
4. What worked, what did not work, and why?

Grading

The key point is that we are not just interested in the absolute performance of what you produce, but primarily the thought process that led you there. Grading will be based on the completeness of the project, the clarity of the writeup, the level of complexity/difficulty of the approach, and your ability to justify the choices you made.

Grade Breakdown

Implementation and results - 50%

- Approach and results demonstrate understanding of the problem (includes figures, tables) → Reproducibility!: 20%
- Rigor of evaluation and reasoning / discussion: 20%
- Innovation: 10%

Quality of write-up - 30%

- Introduction is compelling, problem statement and desired outcome are clear: 10%
- Methods and evaluation protocol are clearly described (reproducibility!) and there are references to related work: 10%
- Discussion and conclusion comprise of well formulated arguments, ideally grounded in scientific literature: 10%

Presentation during pitch & demo session - 20%

Suggested Projects

Title: Robot tool segmentation

Background: The da Vinci robot is state-of-care in several routine surgeries. Automating surgery is one of the core research interests in the community. To do so, precise tool location with respect to the stereo camera is needed to enable visual-servoing of robot joints. The first step in accomplishing this is precise and reliable detection and segmentation of tools.

Goal: Train a deep neural network that accurately segments robotic tools during laparoscopic surgery.

Data: The EndoVis challenge is a great place to start.

Title: Monocular depth estimation

Background: Reconstructing 3D scenes from consumer grade cameras is receiving increasing attention. While low-cost depth-sensing cameras exist (e.g. the Kinect and Intel RealSense camera families), deep convolutional neural networks are able to estimate depth values directly from RGB images, making dedicated hardware obsolete.

Goal: The goal in this project is to develop a training scheme that enables self- or unsupervised learning of monocular depth estimation networks solely from RGB video.

Data: We recommend you start with either KITTI or CITYSCAPES.

Title: Image classification

Background: Classifying images in different categories is a standard task in computer vision and deep neural networks solve this problem on ImageNet with super-human performance when perfect annotations are provided for supervised learning.

Goal: Can you beat Andrej Karpathy's score on ImageNet without using the labels?

Data: We recommend you start with tiny ImageNet and then transition to ImageNet.

Title: Domain-unsupervised learning of generalizable models
Background: Generalizability is probably the most important characteristic of any machine learning model. Can you find a technique to train models that generalize to many different domains without knowing that different domains may exist?
Goal: Train an image classification network that naturally handles domain shift (e.g. clean to noisy images) without knowing the other domain exists. Specifically, the new domain cannot be used during training or validation, only for testing.
Data: We recommend you base your investigations of the study discussed in class that considered "Generalisation in humans and deep neural networks". Instructions on how to reproduce their data can be found on their GitHub.

Title: Semi-Supervised Reinforcement Learning (RL)
Background: Deep networks are data hungry in general, but are even more so in reinforcement learning settings, requiring orders of magnitude more training examples than supervised learning settings. Moreover, hand-crafted reward functions can lead to undesired behavior. There have been recent approaches to utilizing reward functions that are more directly a function of data alone. For example, solely curiosity-driven approaches (where an agent is rewarded for exploring new situations) can outperform competitive supervised approaches while circumventing the annotated data problem.
Goal: This project aims at exploring semi-supervised and unsupervised approaches to RL. Ideally, modifications, combinations, or new ideas can improve upon current approaches.
Data: Open AI gym has many out-of-the-box environments to play with. We recommend starting with simple environments. (https://gym.openai.com/envs/). Also, there's a RL code repository maintained by BAIR Lab: https://bair.berkeley.edu/blog/2019/09/24/rlpyt/

Title: Multiple View Project
Background: In the real world, physical processes beget multiple sensory modalities (vision, audio, etc.). In machine learning, we can utilize these multiple modalities to perform better inference. As an example, combining articulatory data and audio data representations can recover a classic linguistics chart characterizing vowel sounds (IPA vowel chart), completely unsupervised.
Goal: Utilize multiple modalities for some process via unsupervised or semi-supervised learning. Present specific insights that weren't obtainable with single-modalities on some specific dataset(s) and/or improve upon existing methods.
Data: Multimodal Machine Learning Survey (Baltrušaitis et al.) provides a good starting point for datasets and understanding the current state of multi-modality. It may be easiest to start with the split MNIST dataset.

Title: Fine-tuning BERT
Background: Recently (in 2018), Bidirectional Encoder Representations from Transformers (BERT) improved upon state-of-the-art in multiple tasks. BERT provides a general

pre-trained language architecture that can be adapted to a wide range of language tasks with minimal modifications and training. Since then, many papers have been released that apply BERT in interesting ways and improve upon benchmarks.

Goal: In this project, you will be fine-tuning BERT to new tasks or datasets that it hasn't been applied to before. Ideally, this will provide interesting insights not obtainable by other models/architectures and improve upon results.

Data: You can get started with https://mccormickml.com/2019/07/22/BERT-fine-tuning/ which will walk you through the BERT fine-tuning process in PyTorch. Though, a large portion of the project will be your responsibility in finding an appropriate scenario or task where BERT can be applied to. Because this project provides room for creativity, make sure your project proposal is strong! We want to make sure you are likely to succeed.