Final Project

Theory of Machine Learning
CU Boulder

Spring 2021

1 Description

The final project is a theoretical investigation of a problem, setting, or algorithm in the sphere of ML theory. Projects may be either:

- **Research.** Students will identify an original, unsolved problem, and attempt to solve it. Examples can be extensions of existing algorithms to more difficult settings, or identifying a new machine-learning setting.
- **Expository.** Students will identify an advanced topic with relatively few or difficult resources, and synthesize these into a helpful and readable form. Examples can be advanced algorithms or more sophisticated proofs; or a learning-theory setting not covered in class.

Resources used to accomplish this may include textbooks, research papers, etc.

2 Logistics and Assessment

Project groups may consist of 1, 2, or 3 people.

- By **Wed, March 17**, the group will submit a project proposal on Gradescope, consisting of one to two paragraphs.
- By **Wed. March 31**, the group will submit a project update on Gradescope of about one page.
- By **Wed, April 14**, the group will submit a project update on Gradescope of about one page.
- By **Tue/Thu, April 27/29**, the group will give a short presentation (10-20min) to the class on their problem and findings.
- By **Sat, May 1, 10:00pm Mountain time**, the group will submit a project report of 5-10 pages on Gradescope.

Remember that the final report is just a condensed summary and writeup or product. It will not be able to capture everything you've explored in your project; that's okay.

3 Example Topics

Groups are encouraged to propose their own topic, especially one relevant to their outside interests.

- PAC learning topics
  - Structural risk minimization (in more detail than class)
  - Optimal sample complexity bounds for PAC learning
  - "Compression" based sample complexity bounds (information theory connections); minimum description length
  - "Fast rates", when we can generalize with error $O(1/n)$ rather than $O(\sqrt{1/n})$.
  - Pseudo-dimension, fat-shattering dimension, etc.
- Topics in theory of neural networks
  - Theory of GANs (generative adversarial networks)
  - Explaining how networks generalize
- Online learning topics
  - Connections of regret to correlated equilibrium in games
  - Stackelberg regret and learning (can find papers)
  - Convex optimization and two-player zero-sum game connections
  - Follow the Perturbed Leader, Mirror Descent, and other advanced online learning algs.
- Reinforcement learning
- Advanced bandits problems, such as contextual bandits
- Gaussian processes and kernels