Large-Scale Machine Learning

Instructor: Dimitris Papailiopoulos, Assistant Professor, ECE Department
Email: dimitris@papail.io
Course website: papail.io/901
Class meets: Tuesday and Thursday, 11:00am-12:15pm in EH3534
Student hours: Please email the instructor to schedule a time that is mutually convenient.

Course Outline:

This course will explore the mathematical foundations of a rapidly evolving new field: large-scale machine learning and optimization. We will focus on recent texts in machine learning, optimization, and randomized algorithms, with the goal to understand the tradeoffs that are driving algorithmic design in this new discipline. These trade-offs will revolve around statistical accuracy, scalability, algorithmic complexity, and implementation.

Topics include:

1. Optimization and Learning
- Stochastic Methods for Convex and Nonconvex Settings
- Overfitting, Generalization, and Algorithmic Stability
- Expressive Power of Neural Nets, Hardness, and Recent Results

2. Large-scale Learning and Systems
- System Tradeoffs and Platforms
- Synchronous and Asynchronous Distributed Optimization
- Stragglers and Adversarial Attacks during Distributed Learning

Student Evaluation

- **Semester Project:** 60%
    Milestone 1: mid semester project proposal, and presentation (20%).
    Milestone 2: end-semester report, and poster presentation (40%).
- **Homeworks:** 20%
- **Scribe Notes:** 10%
- **Student Participation:** 10%

Semester project: Groups of 2-3 students will work on an open problem that is relevant to the course. The *first component* of the project will consist of a written (at most 3-page) project proposal that will be due approximately by the middle of the semester, followed by a slides presentation. During the presentation (expected to last approximately 15 minutes) each group will receive feedback from the instructor and the rest of the class. The *second component* of the project will be an end-of-semester report, where the members of each group will present the progress of their project, their potential innovations, and their concluding remarks. All projects will be accompanied by a poster. All posters will be presented during a poster day (at a date to be determined), where other faculty and students will be invited to attend.

Homeworks: These will involve exercises relevant to the covered material and paper reviews. The frequency of homeworks is expected to be one every around 3 weeks.

Scribing: All students are required to scribe notes for at least one lecture. The scribes will be due one week after their corresponding lecture. Depending on the size of the class, up to B students will be selected per lecture, so that B = #lectures/#enrolled_students/.

Student Participation: The students are expected to actively participate in the course with questions, suggestions and feedback, and are expected to be present during classes.

Key References

[1] ***"Convex Optimization: Algorithms and Complexity"***
by Sébastien Bubeck
Source: http://research.microsoft.com/en-us/um/people/sebubeck/book.html

[2] ***"Understanding Machine Learning: From Theory to Algorithms"***
by Shai Ben-David and Shai Shalev-Shwartz
Source: http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/index.html

[3] ***"Optimization Methods for Large-Scale Machine Learning"***
by Léon Bottou, Frank E. Curtis, and Jorge Nocedal
Source: https://arxiv.org/pdf/1606.04838v1.pdf

Note: All presented material, slides, and scribes, will be posted online, on the course's webpage.

Prerequisites

ECE/CS 761: Advanced Machine Learning

This course is ideal for advanced graduate students, who are interested in applying novel research concepts to their own research. Students are expected to be familiar with basic concepts in optimization and machine learning, and a solid background in linear algebra and probability.