CS 59200 — Distributed Optimization for Machine Learning

Instructor: Brian Bullins
Email: TBD

Course Description

As the size of the networks and datasets for modern machine learning models have experienced tremendous growth, so too has there been an increasing need for finding better ways to handle the enormous problems in this setting. Distributed optimization presents one such opportunity, whereby these large datasets/models are split across many machines so that much of the computation occurs in parallel. In this course, we will explore the many recent advances in developing algorithms, many of which are based on stochastic gradient methods, for this important area of research.

We will approach the problem from both theoretical and practical angles. Throughout the course, we will aim to formally describe and analyze various distributed optimization settings in terms of e.g. the number of machines we have, how many rounds of communication are conducted, what oracle models are accessed, etc. At the same time, we will look into the practical implications of these methods' choices for efficiently solving many relevant problems.

Course Format

The course will be mostly based on reading recent research papers in the distributed optimization literature, as published at conferences such as ICML, NeurIPS, COLT, etc. Students will be expected to present the contributions and results of papers, and will further lead a discussion on the works, with the expectation that the other students will be able to contribute to the discussion. At the beginning of the course, I will present a few introductory lectures on some of the basics of distributed optimization theory.

Learning Outcomes

By the end of this course, students should be able to:

- Understand the importance and relevance of distributed optimization for modern machine learning, along with their current limitations
- Have a well-informed overview of state-of-the-art results in distributed optimization research
- Critically analyze and interpret the contributions of relevant papers in the research area

- Determine novel directions of research and think effectively about techniques to make progress for such open problems

Course Materials

While there is no required textbook, the following resources are recommended:

Convex Optimization, by Boyd and Vandenberghe
Convex Optimization, Algorithms and Complexity, by Sébastien Bubeck
Numerical Optimization, by Nocedal and Wright

Grading Overview

Reading / Paper Discussions: 20%
Paper Presentations: 30%
Final Project: 50%