Stat364: Scalable Statistical Inference for Big Data with Applications
Time: 10:30am-11:45am, Monday, Spring 2022
Lead Instructor: Xihong Lin

Description: Unprecedented advances in digital technology have produced a revolution that is transforming science and society. Big data have been rapidly generated in many disciplines, such as genomics, health, physical and social sciences. The value of big data lies in effective analysis using statistical inference and machine learning methods that are computationally scalable and efficient. This seminar course is offered in response to these needs. Discussions will be co-led by several faculty members in the Departments of Statistics and Biostatistics at Harvard with active participation by students, postdoctoral fellows and faculty, with the goal of stimulating and discussing several cutting-edge ongoing research areas on scalable statistical inference for big data.

The discussions include current research activities, challenges and open problems. Examples include dense and sparse signal detection, feature selection, distributed learning, boosting, high-dimensional prediction and testing, sparse and low-rank matrix estimation, tensor methods, fast PCA, post PC inference and eigenvector theory, efficient concave and non-concave penalized likelihood maximization and inference, nonparametric function estimation, network analysis, convolutional neural networks, random projection, sketching, cloud computing, variational Bayesian methods, and estimating (possibly) low dimensional functionals in high dimensional models. Applications of these methods in various areas will be discussed.

Class schedule and reading list: http://bit.ly/stats364_presentation

Responsibilities:

- Active class participation
- Serve as a discussant in a group and present group discussions

Class slack account:

stat364.slack.com

Signup Sheet

http://bit.ly/stat364_signup

Zoom link

https://harvard.zoom.us/j/99894185420?pwd=SkxoaWdCZVI0WXZnYjhaNDk2RjFTQT09