Statistical machine learning for Financial Prediction and Inference

Jianqing Fan

Department of Operations Research and Financial Engineering

Princeton University

ABSTRACT

This talk first gives an overview on the genesis of machine learning and AI and how statistical and computational methods have evolved with growing dimensionality and sample sizes and become the foundation of modern machine learning and AI.

We will introduce recent developments of statistical machine learning methods for analysis of Big Data in Finance. Motivated by stylized features such as heavy-tails and cross-sectional dependence, we introduce a simple method for dependence adjustment and robustfication principles for dealing with heavy tail gig-data issues in finance. We then apply factor models to extract latent factors for prediction, Factor-Adjusted Robust Multiple test (FARM-test) and model selection (FARM-select). We will highlight three applications: robust forecast of bond risk premia via augmented factor models, AI investing via novel estimation of Sharpe ratio, and predictability of the momentum and duration in high-frequency finance.