Abstract:
Nowadays, data is exploding at a faster rate than computer architectures can handle. For that reason, mathematical techniques to analyze large-scale data need be developed. Stochastic iterative algorithms have gained interest due to their low memory footprint and adaptability for large-scale data. In this talk, we will present the Randomized Kaczmarz algorithm for solving extremely large linear systems of the form $Ax = y$. In the spirit of large-scale data, this talk will act under the assumption that the entire data matrix $A$ cannot be loaded into memory in a single instance. We consider different settings including when a only factorization of $A$ is available, when $A$ is missing information, and a time-varying model. We will also present applications of these Kaczmarz variants to problems in data science.