Mathematics, Statistics, and Computer Science **@ UIC**

Statistics and Data Science Seminar

Model-free variable selection via learning gradients

Lei Yang (UIC)

Abstract: Variable selection is popular in high-dimensional data analysis to identify the truly informative variables. Many variable selection methods have been developed under various model assumptions, such as linear model and additive model. However, their success largely rely on validity of the assumed models. In this talk, I will introduce a model-free variable selection method based on gradient learning. The key idea is that if a variable is informative is equivalent to if its corresponding gradient function is substantially non-zero. The proposed method is formulated in a framework of learning gradients equipped with a flexible reproducing kernel Hilbert space. Computationally, a blockwise majorization decent (BMD) algorithm is introduced for efficient computation. Theoretically, without assuming explicit models, the estimation and variable selection consistencies are established. A variety of simulated examples and real-life examples are provided to evaluate the performance.

Wednesday, March 4 at 4:00 PM in SEO 636