

Special Colloquium

Modern Classification with Big Data

Boxiang Wang (University of Minnesota)

Abstract: Rapid advances in information technologies have ushered in the era of "big data" and revolutionized the scientific research. Big data creates golden opportunities but has also arisen unprecedented challenges due to the massive size and complex structure of the data. Among many tasks in statistics and machine learning, classification has diverse applications, ranging from improving daily life to reaching the new frontiers of science and engineering. This talk will discuss the envisions of broader approaches to modern classification methodologies, as well as computational considerations to cope with the big data challenges. I will present a modern classification method named data-driven generalized distance-weighted discrimination. A fast algorithm with an emphasis on computational efficiency for big data will be introduced. Our method is formulated in a reproducing kernel Hilbert space, and learning theory of the Bayes risk consistency will be developed. In addition, I will use extensive benchmark data applications to demonstrate that the prediction accuracy of our method is highly competitive with state-of-the-art classification methods including support vector machine, random forest, gradient boosting, and deep neural network.

Wednesday, January 24 at 3:00 PM in SEO 636