博 士 後 演 講

講 題：MPCGA: A Tree-Based Chebychev's Greedy Algorithm

演講人：Mr. Yan-Shuo Pan（潘彥碩 博士候選人）
(國立清華大學統計研究所)

時 間：2023年5月31日(星期三)，14:00-15:00

地 點：中央研究院統計科學研究所 B1F 演講廳

Abstract

Prediction and feature selection are essential topics for statistical and machine learning (ML) methods when dealing with high-dimensional data. However, they come with limitations: statistical methods may exhibit lower predictive ability than ML methods, while ML methods are often criticized as black boxes. In this paper, we introduce a tree-based algorithm, the Multipath Chebyshev Greedy Algorithm (MPCGA), which enhances the predictive performance of statistical methods and feature selection capabilities under model misspecification. This algorithm extends the Chebychev's Greedy Algorithm (CGA) and High Dimensional Information Criterion (HDIC) into a tree-expanded structure, allowing for the simultaneous consideration of multiple models. MPCGA outperforms traditional statistical methods when models are misspecified, while maintaining high feature selection precision. Furthermore, we propose accelerated algorithms to boost the computational speed of MPCGA handling indicator features in binary outcome cases. The paper includes a case study on a lung cancer dataset, demonstrating that utilizing ML methods with the feature set selected by MPCGA leads to suitable results.

※ 茶 會：15：00開始   ※ 中文演講，實體與線上視訊同步進行。