INSTITUTE OF MATHEMATICAL SCIENCES
UNIVERSITI MALAYA

SIRI SEMINAR KUMPULAN PENYELIDIKAN

Title: Machine Learning and Statistical Strategies in High-dimensional Predictive Modelling.
Speaker: Professor Dr. Syed Ejaz Ahmed.
Date: 12 August 2024.
Time: 11.00 a.m. to 12.00 p.m.
Venue: MM3, Level 2, Institute of Mathematical Sciences, Faculty of Science, Universiti Malaya.

ABSTRACT

There are a host of buzzwords in today's data-centric world. We encounter data in all walks of life, and for analytically and objectively minded people, data is crucial to their goals. However, making sense of the data and extracting meaningful information from it may not be an easy task. The rapid growth in the size and scope of data sets in a host of disciplines has created a need for innovative statistical strategies to understand such data. A variety of statistical and machine learning methods would be useful in revealing the hidden data story. Complex big data analysis is a very challenging but rewarding research area as data sets include a larger number of features, data contamination, unstructured patterns, and so on. A host of models are now data-driven with a large number of predictors, namely high-dimensional data (HDD).

For HDD analysis, many penalized methods were introduced for simultaneous variable selection and parameter estimation when the model is sparse. However, a model may have sparse signals as well as with number of predictors with weak signals. In this scenario variable selection methods may not distinguish predictors with weak signals and sparse signals. For this reason, we propose a high-dimensional shrinkage strategy to improve the prediction performance of a submodel. We demonstrate that the proposed high-dimensional shrinkage strategy performs better than the penalized and machine-learning methods in many cases. The relative performance of the proposed strategy is appraised by both simulation studies and real data analysis. I will also discuss some open research problems and possible solutions.

Reference:

S. Ejaz Ahmed, Feryaal Ahmed and B. Yuzbasi (2023). Post-Shrinkage Strategies in Statistical and Machine Learning for High Dimensional Data. CRC Press, USA.

All are Welcome