Statistics and Actuarial Science Seminar

Title: Non-stationary Gaussian process discriminant analysis with variable selection for high-dimensional functional data

Speaker: Sara Wade (University of Edinburgh)

Date: Mon 21st March 2022 at 12:00PM

Location: Online

Abstract: High-dimensional classification and feature selection tasks are ubiquitous with the recent advancement in data acquisition technology. In several application areas such as biology, genomics and proteomics, the data are often functional in their nature and exhibit a degree of roughness and non-stationarity. These structures pose additional challenges to commonly used methods that rely mainly on a two-stage approach performing variable selection and classification separately. We propose in this work a novel Gaussian process discriminant analysis (GPDA) that combines these steps in a unified framework. Our model is a two-layer non-stationary Gaussian process coupled with an Ising prior to identify differentially-distributed locations. Scalable inference is achieved via developing a variational scheme that exploits advances in the use of sparse inverse covariance matrices. We demonstrate the performance of our methodology on simulated datasets and two proteomics datasets: breast cancer and SARS-CoV-2. Our approach distinguishes itself by offering explainability as well as uncertainty quantification in addition to low computational cost, which are crucial to increase trust and social acceptance of data-driven tools.

Join the seminar: https://ucd-ie.zoom.us/j/68316324831