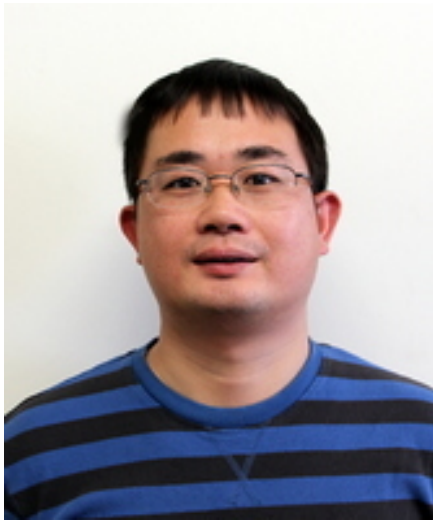# Longhai Li's Short CV

Longhai Li is a professor at the University of Saskatchewan, located in Saskatoon, SK, Canada. He received his Ph.D. degree in statistics from the University of Toronto, where his supervisor was Prof. Radford Neal. Before that, he received B.Sc honours in statistics from the University of Science and Technology of China. He served on the NSERC IDG EG 1508 committee (Mathematics and Statistics) during 2022-2025. His research activities focus on developing and applying statistical machine-learning methods for high-throughput data and complex-structured data. His research has been funded by NSERC, CFI, CFREF, and MITACS. His research papers have appeared in high-impact journals, such as Journal of American Statistical Association, Bayesian Analysis, Statistics in Medicine, Statistics and Computing, Scientific Reports, and BMC Bioinformatics. A predictive model comparison method called integrated importance sampling introduced by him was included in a textbook by Peter Congdon: Bayesian Hierarchical Models: With Applications Using R.

Keywords summarizing his expertise

Statistical machine learning, survival analysis, model checking, residual diagnostics, model comparison, cross-validation, information criterion, zero-inflated models, high-throughput data, and mixed-effects models.

A brief description of my current research activities

My current research activities can be described with two themes: (1) statistical machine learning for bioinformatics data and (2) model evaluation for public health data. My research on theme 1 aims to develop new tools for honestly measuring the predictivity (such as error rate, AUC) of selected features, and new tools for identifying truly predictive features and for building sharper predictive models for phenotypes. I am particularly interested in uncovering the molecular mechanisms behind Alzheimer's and Parkinson's diseases. My research on theme 2 aims to develop new tools for evaluating Bayesian/non-Bayesian models with complex structures. Today, increasingly complicated models are being proposed for a variety of correlated data such as temporal, spatial, and repeated measurement data. There is a gap between developing new modelling methods and model validation methods. I am working on developing new residual diagnostic methods for checking the adequacy of statistical models.