

Name: Qing Lu

Email: lucienq@ufl.edu

## **Kernel Neural Networks for Analyzing Large-scale Substance Use Disorders Genetic Data**

Xiaoxi Shen; Xiaoran Tong; Qing Lu

Texas State University; University of Kentucky; University of Florida

Artificial intelligence (AI) is a thriving research field with many successful applications in areas such as computer vision and protein structure prediction. Neural-network-based methods, such as deep learning, play a central role in modern AI technology. While neural-network-based methods also hold great promise for substance use disorders (SUD) genetic research, the high dimensionality of genetic data, the massive amounts of study samples, and the complex relationships between genetic variants and SUD bring tremendous challenges to use these methods in SUD genetic research. To address these challenges, we propose a new AI method based on kernel neural networks (KNN) for analyzing large-scale genetic datasets involving millions of variants and tens of thousands of samples. Unlike the classic neural networks, KNN models the overall genetic effects as random effects, which could substantially alleviate the overfitting issue and reduce the complexity of network structures. Kernel-based neural networks and Minimum norm quadratic unbiased estimation are implemented in KNN to model complex relationships (e.g., interactions) and accelerate the computation. Through simulations, we demonstrate the advantages of KNN over existing methods in terms of prediction accuracy and computational efficiency. We also apply KNN to the large-scale UK Biobank dataset, investigating high-dimensional risk prediction models for SUD.