

Simple Bias In Deep Learning

Date: 7 December 2023 (Thursday)

Time: 10:00am - 11:00am

Venue: Room 6-211, Lau Ming Wai Academic Building, City University of Hong Kong

ABSTRACT

Why do neural network models that look so complex usually generalize well? To understand this problem, we find two simple biases in deep learning training. The first is the frequency principle that neural networks learn from low frequency to high frequency. To mitigate the difficulty of learning high-frequency information, we developed Multi-scale DNN. The second is the parameter condensation, which makes the number of effective neurons in large networks far less than the actual number of neurons. We utilize the condensation effect to reduce network size, and use an example of solving high-dimensional ODEs to demonstrate the reduction method. These implicit biases all reflect the characteristic that neural networks tend to use simple functions to fit data in the training process, to achieve better generalization. In the last part, I will show some AI-based algorithms for solving high dimensional ODEs in combustion.



Dr Zhi-Qin Xu

GUEST SPEAKER'S PROFILE

Dr Zhi-Qin John Xu is an Associate Professor at Shanghai Jiao Tong University (SJTU). Dr Zhi-Qin obtained B.S. in Physics (2012) and a Ph.D. degree in Mathematics (2016) from SJTU. Before joining SJTU, Zhi-Qin worked as a postdoc at NYUAD and Courant Institute from 2016 to 2019. He is interested in studying deep learning theory and AI for Science. In deep learning theory, he discovered frequency principle, parameter condensation and embedding principles in deep learning, and developed multi-scale neural networks. In AI for Science, he developed AI-based algorithms for solving high dimensional ODEs. He is one of the managing Editors of Journal of Machine Learning.