Abstract for 6th Chinese Computational and Cognitive Neuroscience Conference (CCCN 2024) (July 3-5, 2024)

Neural Network Training: Theory Framework and Challenges

Ruoyu Sun

School of Data Science, Chinese University of Hong Kong, Shenzhen, P. R. China

Email: sunruoyu@cuhk.edu.cn

Artificial neural networks have achieved remarkable progress in the past decades. The achievements are impossible without the ability to train large neural networks, and many people thought successful training of neural nets is mostly engineering. In this talk, we will demonstrate a theoretical framework for the successful training of neural networks. This framework decouples the training error into representation error and optimization. We present a few theories that link the widely used training tricks such as initialization methods and Adam to various kinds of errors. We will also briefly discuss a few challenges in training large foundation models.