PURDUE UNIVERSITY | Elmore Family School of Electrical and Computer Engineering

Faculty Candidate Seminar -- Purdue Computes: AI/ML

Chaoyue Liu

Postdoc, Halicioglu Data Science Institute
UC San Diego

Tuesday, March 19, 2024
10:30 A.M. – 11:30 A.M.
MSEE 112

The Power of Opening the Black Box of Deep Learning

Abstract

We are now observing an ongoing "AI spring" powered by the emergence and successful implementation of deep neural network models. However, neural networks are often considered as black box models lacking a theoretical and fundamental understanding, leaving deep learning mostly driven by heuristics, not interpretable and not fully reliable. My research focuses on (1) opening the black boxes by discovering fundamental properties of neural networks and deep learning algorithms, and (2) applying these findings to practice and/or theory to push forward the boundary of deep learning.

In this talk, I will illustrate that neural networks and deep learning algorithms can be theoretically understood, to a greater extent than you expected, and show how powerful these findings can be when applied. Specifically, in the first part, I will introduce a new and interesting property of neural networks – transition to linearity, and then apply it to optimization theory to obtain fast convergence guarantee of (stochastic) gradient descent on the non-convex loss function of neural networks. In the second part, I will show that the popularly used Nesterov momentum algorithm does not provide an expected acceleration when implemented onto SGD, even for linear regression problems. Instead, our new algorithm, MaSS, is provable to provide an acceleration in the stochastic setting, and experimentally outperforms the commonly used algorithms in deep learning. Lastly, I will showcase a new finding both theoretically and experimentally that the ReLU activation function has an effect of decreasing the condition number and hence helps in accelerating neural network training.

Bio

Chaoyue Liu is a postdoc at Halıcıoğlu Data Science Institute (HDSI), UC San Diego, working with Prof. Misha Belkin. His research focuses on the foundation of deep learning and its applications. He obtained his Ph.D. degree in Computer Science from The Ohio State University in 2021, where he was advised by Prof. Misha Belkin. After that, He spent one year at Meta Platforms Inc., as a research scientist. He also holds B.S. and M.S. degrees in physics from Tsinghua University.