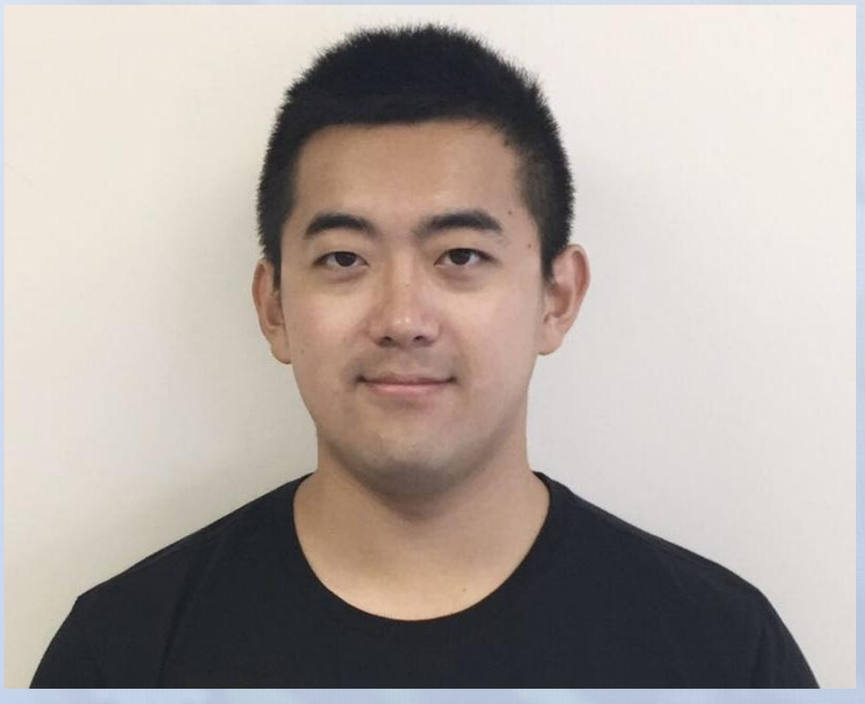
数据科学与深度学习
青年学者及博士生论坛

Why Does Differential Privacy Noise Have Little Impact on Fine-Tuning? A Representation Learning Perspective



王晨笛

（厦门大学）

时间：2025年5月12日 14：00-15：00

地点：光华楼东主楼 2001

Abstract:

Pre-training on public data significantly boosts differentially private (DP) learning in downstream tasks. We analyze this via representation learning, showing that strong pre-trained models like Vision Transformers yield better last-layer features. Despite this, DP fine-tuning is less robust than non-private training. We propose techniques—feature normalization and PCA—that improve DP accuracy. For intermediate layers, we assess how DP noise impacts feature separability, finding that careful hyperparameter tuning can maintain high accuracy even under strong privacy. Our results show that pre-training and targeted finetuning can effectively reduce the privacy-utility trade-off in DP deep learning.

个人简介：

Dr. Chendi Wang is an Assistant Professor at the Paula and Gregory Chow Institute, WISE, and the School of Economics at Xiamen University. He received his Ph.D. from The Hong Kong Polytechnic University and conducted postdoctoral research at the Wharton School, University of Pennsylvania, and the Shenzhen Research Institute of Big Data. His work on data privacy and machine learning has appeared in leading conferences such as ICML, ICLR, and NeurIPS, with an oral presentation at ICML 2024 and media coverage by New Scientist.