

Tame the Beast: Practical Theories for Real-World Responsible AI Deployment

Zhun Deng, PhD

Postdoctoral researcher at Columbia University



Modern digital systems powered by machine learning have permeated various aspects of society, playing an instrumental role in many high-stakes areas such as disease diagnosis and finance. Therefore, it is crucial to ensure that machine learning algorithms are deployed in a “responsible” way so that digital systems are more reliable, transparent, and aligned with societal values. In this talk, Dr. Deng will provide an overview of my research on building practical theories to guide real-world responsible deployment of machine learning. First, he will introduce our recent work on distribution-free uncertainty quantification for a rich class of statistical functionals of quantile functions to avoid catastrophic outcomes and unfair discrimination in the deployment of black-box models. The power of our framework is shown by end-to-end applications on tumor detection and medical summaries with large language models. Second, he will describe an extension to the previous framework on the group-based fairness notions so as to protect every group that can be meaningfully identified from data. Finally, he will further demonstrate how to ensure fairness generalization for complex white-box models, such as neural networks, under data imbalance along with experimental results on image recognition.

Tuesday, January 23, 2024, 3:30-4:30PM Eastern

2308 McGavran-Greenberg Hall

Zoom Link:

<https://unc.zoom.us/j/94304183888?pwd=SW5kQmJ1bkVVOFN5U2ZiQzR0MThnQT09>

Meeting ID: 943 0418 3888

Passcode: 050276