Department of Data Science
香港城市大學
City University of Hong Kong

DS SEMINAR

Theoretical and Algorithmic Perspectives on Decision Making: Efficiency and Robustness

Date: 26 February 2025 (Wednesday)

Time: 9:30am - 10:30am

Seminar Link: https://cityu.zoom.us/j/84568279465

Mr. Jian QIAN

ABSTRACT

AI-powered decision-making systems have revolutionized many fields by transforming complex data into actionable insights. However, challenges remain in optimizing their training efficiency and ensuring robustness against adversarial attacks. To address these challenges effectively, it is essential to develop solutions grounded in rigorous theoretical principles. In this talk, I will address two critical aspects of AI design that directly respond to these challenges: (i) Efficiency: I will introduce a general framework for interactive decision-making, encompassing multi-armed bandits and reinforcement learning, which are core models behind systems like AlphaGo and large language models (LLMs). I will present the learning limits and a general algorithm design principle under this framework, which enables the development of new algorithms with provable near-optimal guarantees. (ii) Robustness: I will examine clean-label data poisoning, where seemingly correct but strategically manipulated data misleads the learning process. My work exposes vulnerabilities in max-margin classifiers and presents a theoretically optimal defense strategy.

By offering these theoretical insights, this talk deepens our understanding of AI's fundamental limits and guides the development of systems that are both efficient and robust.

GUEST SPEAKER'S PROFILE

Jian Qian is a Ph.D. student in the Department of Electrical Engineering and Computer Science at MIT, advised by Prof. Alexander Rakhlin. His research focuses on understanding the fundamental aspects of machine learning and decision making. He is particularly interested in topics including efficiency and robustness. His work has been published in top-tier ML conferences. He is also enthusiastic about exploring new research directions that bridges the theory and practice of machine learning.

Enquiries: ds.go@cityu.edu.hk

All are welcome