國立中央大學數學系

專題演講

主 講 人：吳沛遠副教授 (臺灣大學電機工程學系)

演講題目：From Stochastic Bandits to Kernel Function Approximation in Offline Reinforcement Learning

演講茶會：2026年12月03日(星期四) 15:30~16:00

茶會地點：中央大學鴻經館 M306

演講時間：2026年12月03日(星期四) 16:00~17:00

演講地點：中央大學鴻經館 M107

Abstract：

In this tutorial, we begin with stochastic bandits, where a learner chooses an action each round, observes only that action's noisy reward, and aims to minimize cumulative regret relative to always choosing the action with the highest expected reward. We use this setting to introduce *optimism in the face of uncertainty* and review prior work on Linear UCB and its regret guarantees, Kernel UCB for nonlinear rewards, and LSVI and LSVI-UCB for value estimation and exploration in reinforcement learning. We conclude by connecting these ideas to our lab's research on *kernel Function Approximation in Offline Reinforcement Learning*. Standard offline RL learns from a fixed dataset of transitions with recorded rewards. Our work studies how kernel methods can use both labeled and unlabeled transitions to learn a policy when reward labels are available for only part of the dataset.

敬 請 公 告 ， 歡 迎 參 加