Indian Institute of Technology Kanpur

Department of Computer Science and Engineering

Proposal for a New Course

Course Title: Theory of Multi-Armed Bandits

Course No.:

Credits: 3-0-0-0 [9]

Duration of Course: Full Semester

Prerequisites: Knowledge of undergraduate algorithms and discrete probability

Proposing Department: CSE

Other Departments/IDPs which may be interested in the proposed course: Mathematics, Electrical Engineering

Proposing Instructor: Gunjan Kumar, Department of CSE

Course Description:

In many real-world scenarios, decision-makers face environments where the best action is unknown, and they must make choices under uncertainty. The challenge lies in balancing exploration—learning which actions yield the highest rewards—and exploitation—leveraging known actions to maximize reward. This tension, known as the exploration-exploitation trade-off, is central to the theory of multi-armed bandits, a framework that has been studied for nearly a century. These algorithms are widely applicable in areas such as online advertising, recommendation systems, auctions, network routing, e-commerce, and any scenario where information is collected progressively over time.

In this course, we will study a variety of bandit algorithms designed for different random environments to maximize rewards over the duration of play. The course will cover both foundational concepts and the latest research developments. It will serve as a complement to courses on online convex optimization and reinforcement learning.

Course Contents: 27 lectures/40 hours

S.No	Broad Title	Topics	No. of Lectures
1	Overview and motivation of the course, recap of probability	Concentration inequalities	2
2	Bandits with IID Rewards	Explore-then-commit (ETC) algorithm, Successive Elimination algorithm, UCB1 algorithm	2
3	Lower Bounds on Regret	Review of Information theory, $\Omega(\sqrt{KT})$ lower bound, $\Omega(\log T)$ instance-dependent lower bound	3
4	Bayesian Bandits	Thompson Sampling algorithm	1
5	The Best-Expert Problem (Full Feedback & Adversarial Rewards)	Majority vote algorithm, Weighted majority algorithm	2

6	Adversarial Bandits	Exp3 and Exp4 algorithm	2
7	Lipschitz Bandits	Continuum-armed bandits, Lipschitz MAB, Zooming algorithm	2
8	Contextual and Linear Bandits	Regret analysis of adversarial and stochastic contextual bandits, stochastic linear bandits	3
9	Adversarial Linear Bandits	Foundations of convex analysis, Exp3 for adversarial linear bandits, Follow-the-Regularized-Leader and Mirror Descent, The Relation between Adversarial and Stochastic Linear Bandits	4
10	Pure Exploration	Uniform exploration, Best arm identification: fixed budget and fixed confidence settings, lower bounds	3
11	Miscellaneous Topics	Based on availability of time and interest of students some topics from the following list will be covered: 1. Combinatorial Bandits 2. Markov Decision Processes 3. Bandits and Games 4. Bandits with Knapsacks	3

Reference books:

1. Lattimore, T., & Szepesvári, C. (2020). *Bandit algorithms*. Cambridge University Press.
2. Slivkins, Aleksandrs. "Introduction to multi-armed bandits." *Foundations and Trends® in Machine Learning* 12.1-2 (2019): 1-286.

Dated 10-09-2024 Proposer: Gunjan Kumar

Dated _____ DPGC Convener:

The course is approved/not-approved

SPGC/SUGC Chairperson

Dated: _____