14D006

Stochastic Models and Optimization

Winter Term - 3 ECTS
Elective Course
Prof. Hrvoje Stojic and Prof. Gergely Neu

Prerequisites to Enroll

None

Overview and Objectives

Reinforcement learning (RL) is a model-based theory of sequential decision-making under uncertainty. It is currently a dominant theoretical framework for understanding and building autonomous agents that can learn and act in the environment on their own. The objective of this course is to introduce students to the main challenges and techniques of modern RL, particularly focusing on computational aspects of dealing with the dynamic nature of the RL problem, and on the statistical challenges posed by the uncertainty of the environment. On both fronts, the goal is to provide a strong understanding of the most common methods and provide a basic algorithmic toolbox for building RL systems. The course puts a strong emphasis on crucial challenges that set RL problems apart from other machine learning problems. Students taking the course are expected to gain the capability to identify and tackle such challenges in various application domains. Various applications relevant to Data Science will be highlighted throughout the course.

Course Outline

1. Reinforcement learning preliminaries
   a. The characteristics of reinforcement learning
   b. Showcase applications and current frontier
   c. Agent-Environment, Rewards
   d. Markov decision processes
   e. Challenges: optimization and statistics
2. Multi-armed bandits
   a. Definition, exploration-exploitation trade off in details
   b. Simple algorithms - eGreedy, Softmax
   c. More sophisticated - UCB, Thompson sampling
   d. Adversarial bandits
   e. Combinatorial bandits
3. Contextual multi-armed bandits
   a. Definition, Bayesian optimization
   b. Linear models and UCB
   c. Intro to Gaussian processes
   d. GP-UCB and other variants
4. Bandit applications
   a. A/B testing, clinical trials
   b. Optimizing hyperparameters
   c. News recommendation
5. Dynamic programming basics
   a. The Markov property and its implications
   b. Policies, value functions, optimality
   c. The Bellman equations

   d. Solving MDPs: value iteration, and policy iteration
   e. Two fundamental operations: policy evaluation and improvement
6. Temporal difference methods
   a. Monte Carlo and TD(0) for policy evaluation
   b. Eligibility traces and TD(lambda)
   c. Policy evaluation and improvement: Q-learning and SARSA
   d. TD learning with parametrized value functions
7. Value-based and policy-based methods
   a. Least-squares TD for policy evaluation and improvement
   b. Fitted value iteration and deep Q networks
   c. The policy gradient theorem
   d. Actor-critic methods
8. Robust reinforcement learning
   a. Conservative policy iteration, Trust-region policy optimization
   b. Relative Entropy Policy Search
   c. Regularized Markov decision processes
9. Exploration in Markov decision processes
   a. PAC reinforcement learning
   b. Regret minimization in MDPs: optimistic MDPs and value functions
   c. Thompson sampling and randomized exploration in MDPs

Required Activities

Attendance to classes, completing individual problem sets and a group-based final project.

Evaluation criteria

Problem sets account for 50% of the final grade, while the remaining 50% comes from a final project. The problem sets are carried out by each student individually, whereas the final project in teams of up to three students.

Competences

- Construct a global vision of the situation of the problem based on knowledge of the synergies between advanced statistical methods, computing and business analysis to generate added value.
- Modeling and predicting high-dimensional data with advanced statistical methods in the field of data science in order to improve strategic decision making.
- Apply the knowledge of programming languages, computer programs and advanced services in the Cloud to solve the problems that are presented to the data scientist.
- Solve the real problems that arise in the fields of study through the accurate analysis of the data.
- Communicate with conviction in English the results and implications of the required analytical study using a language related to the receiver.
- Work in a heterogeneous team of researchers in the field of the economic analyst using specific group techniques.
- Own and understand knowledge that provides a basis or opportunity to be original in the development and / or application of ideas, often in a research context.
- That students know how to apply the acquired knowledge and their ability to solve problems in new or unfamiliar environments within broader (or multidisciplinary) contexts related to their area of study.

- That the students be able to integrate knowledge and face the complexity of making judgments based on information that, being incomplete or limited, include reflections on the social and ethical responsibilities linked to the application of their knowledge and judgments.
- That the students know to communicate their conclusions and the knowledge and last reasons that sustain them to specialized and non-specialized publics in a clear and unambiguous way.
- That students have the learning skills that allow them to continue studying in a way that will be largely self-directed or autonomous.

Learning Outcomes

- Apply mathematical and computational analysis of social, business and economic networks knowing the theory and optimization algorithms.
- Apply search algorithms and estimation methodologies in networks through observation of data.
- Apply optimization algorithms in business and marketing problems.
- Apply learning and optimization methods in marketing environments.
- Apply mathematical and statistical analysis using economic theory in complex problems with high-dimensional data.
- Apply mathematical theory and statistics on data sets from disparate disciplines.

Materials

Sutton, R.S., & Barto, A.G. (2018). Reinforcement learning: An introduction. MIT Press.

Lattimore, T., & Csaba, S. (2018). Bandit algorithms.

Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian processes for machine learning. MIT Press.