CS-456

Deep reinforcement learning

Gulcehre Caglar

Cursus	Sem.	Type
Computational biology minor	E	Opt.
Computational science and Engineering	MA2, MA4	Opt.
Computational science and engineering minor	E	Opt.
Computer science	MA2, MA4	Opt.
Cybersecurity	MA2, MA4	Opt.
Data Science	MA2, MA4	Opt.
Data science minor	E	Opt.
Digital Humanities	MA2, MA4	Opt.
Electrical Engineering		Obl.
Financial engineering	MA2, MA4	Opt.
Life Sciences Engineering	MA2, MA4	Opt.
Neuro-X minor	E	Opt.
Neuro-X	MA2, MA4	Opt.
Quantum Science and Engineering	MA2, MA4	Opt.
SC master EPFL	MA2, MA4	Opt.
Statistics	MA2, MA4	Opt.

Language of teaching	English
Credits	6
Session	Summer
Semester	Spring
Exam	Written
Workload	180h
Weeks	14
Hours	**4 weekly**
Lecture	2 weekly
Exercises	1 weekly
Labs	1 weekly
Number of positions	

Summary

This course provides an overview and introduces modern methods for reinforcement learning (RL.) The course starts with the fundamentals of RL, such as Q-learning, and delves into commonly used approaches, like PPO and DQN. The course will introduce students to practical applications of RL.

Content

- *Introduction and Overview (What is RL?)*
- *An overview of neural networks and deep learning approaches*
- *Deep learning frameworks*
- *Supervised learning of behaviors (behavior cloning)*
- *Value function methods and related theory*
- *Policy gradient methods and related theory*
- *Actor-Critic Algorithms (A2C, A3C)*
- *Deep RL with Q functions (DQN, R2D2)*
- *Deep Policy Gradient and Optimization methods (PPO, TRPO, Impala, MPO)*
- *Model-based RL and Planning (Alphago, Alphazero, Dreamer)*
- *Exploration and credit assignment in Deep RL*
- *Offline RL (BVE, CQL, CRR, ...)*
- *Deep Imitation learning and Learning from demonstrations (DAGGER, DQFD, R2D3, Learning from play, Third*

person imitation)

- *RL from human feedback and alignment (InstructGPT, DPO, ReST, etc.)*
- *Advanced continuous control approaches (DDPG, D4PG, SAC)*
- *A selection of extra topics from:*

- MPO, IMPALA
- Distributional RL
- Multi-agent RL (Centralized Training, Decentralized Execution)

Keywords

Deep learning, reinforcement learning, TD learning, SARSA, Actor-Critic Networks, policy gradients, alphago, alphastar, planning, alignment, RLHF, PPO

Learning Prerequisites

Required courses

- Analysis I, II
- Linear Algebra
- Probability and statistics (MATH-232)
- Algorithms I (CS-250)

Recommended courses

- Introduction to machine learning (CS-233)
- Machine learning (CS-433)

Important concepts to start the course

- *Regularization in machine learning,*
- *Gradient descent. Stochastic gradient descent.*
- *Expectation, statistics*
- *Linear algebra and probabilities*
- *programming*

Learning Outcomes

By the end of the course, the student must be able to:

- Apply Understand and define basic problems and tasks in reinforcement learning (like Markov decision process, model-based and model-free RL, on-policy vs off-policy RL)
- Assess / Evaluate Formulate a real-world problem as an RL setting to apply the approaches taught in the class.
- Elaborate Implement standard deep RL algorithms.
- Judge Understand the failure modes of these models and learning algorithms.
- Propose Read and review academic papers to understand their contributions and learn how to evaluate them critically.
- Apply Students gain the skills and knowledge necessary to tackle complex problems in autonomous robotics, game-playing, and other domains through lectures, hands-on coding exercises, practical applications, and course

projects.

Transversal skills

- Continue to work through difficulties or initial failure to find optimal solutions.
- Access and evaluate appropriate sources of information.
- Write a scientific or technical report.
- Manage priorities.

Teaching methods

- Lectures
- Lab sessions
- Individual course projects
- Paper reading
- Group projects

Expected student activities

- Work on miniproject
- Solve all exercises
- Attend all lectures and take notes during lecture, participate in quizzes.
- If you cannot attend a lecture, then you must read the recommended book chapters
- Work on a project

Assessment methods

- Written final exam (25%)
- Assignments (25%)
- Course project (50%)

Supervision

Office hours	No
Assistants	Yes
Forum	Yes
Others	TAs are available during exercise sessions. Every week one of the exercises is run as 'integrated exercise' during the lecture.

Resources

Bibliography

• Textbook: Reinforcement Learning by Sutton and Barto (MIT Press). Pdfs of the preprint version of the book are availble online

Ressources en bibliothèque

• Reinforcement Learning / Sutton

Moodle Link

• https://go.epfl.ch/CS-456