Advanced Artificial Intelligence and Machine Learning: Reinforcement Learning

LMH Summer Programmes are provided by Lady Margaret Hall, a college in the University of Oxford

Course:	Advanced Artificial Intelligence and Machine Learning: Reinforcement Learning
Available:	Programme Session 1: 24th June 2024 to 12th July 2024
Lectures:	18 Hours
Seminars:	12 Hours
Tutorials:	3 Hours
Independent Study:	Approximately 120 Hours
Recommended Credit:	15 CATS / 7.5 ECTS / 4 US Credits

About this Course:	Getting things wrong is part of what makes us human, and our natural intelligence helps us learn from our mistakes. Reinforcement learning is an area of machine learning which enables artificial intelligence to learn from its mistakes as well, for example allowing a robot to use trial-and-error to interact with a new environment and achieve an objective. This advanced course examines the fundamentals of reinforcement learning and explores the varied applications of dynamic programming methods. The course will begin with a thorough grounding in the key theoretical concepts of reinforcement learning, familiarising you with agents, environments, and rewards, before introducing Markov decision processes, dynamic programming, and Monte Carlo methods. As the course progresses you will explore a wide range of reinforcement learning methods and techniques, including policy gradient methods and how they optimise policies, policy search methods such as evolutionary strategies and hill-climbing, and the cross-entropy method for policy optimisation. The final part of the course will introduce even more advanced topics, including multi-agent reinforcement learning. This intensive course offers students theoretical understanding and practical experience in a range of reinforcement learning concepts and techniques, offering career skills as well as excellent foundations for future research.
Course Overview:	Week 1 • Introduction to Reinforcement Learning ○ Fundamentals of Reinforcement Learning, including agents, environments, and rewards. • Markov Decision Processes

Lady Margaret Hall, University of Oxford
Norham Gardens, Oxford, OX2 6QA, United Kingdom
vacation.programmes@lmh.ox.ac.uk
+44 (0)1865 274295

  - The formal framework of MDPs and how they model sequential decision-making.
- Dynamic Programming, Policy Iteration, and Value Iteration
  - Policy evaluation, policy improvement, value iteration, and iterative processes for optimizing policy and value functions in MDPs.
- Monte Carlo Methods and Temporal Difference Learning
  - Using Monte Carlo methods for estimating value functions and optimizing policies.
  - Using TD methods such as SARSA and Q-Learning for model-free reinforcement learning.
- Function Approximation (Neural Networks, DNNs) and Deep Q-Networks
  - Use of function approximation techniques, such as linear approximation and neural networks.
  - DQN and their role in approximating Q-values in high-dimensional state spaces.

Week 2

- Policy Gradient Methods, REINFORCE, and Actor-Critic Architectures
  - How policy gradient methods optimize policies directly.
  - Using REINFORCE algorithm and actor-critic architectures for policy optimization.
- Exploration and Exploitation
  - The exploration-exploitation trade-off, epsilon-greedy, and UCB.
- Multi-Armed Bandit problems and their Relevance to Reinforcement Learning
- Policy Search Methods and the Cross-Entropy Method
  - Policy-search methods including evolutionary strategies and hill-climbing.
  - The cross-entropy method for policy optimization and its variations.
- Model-Based Reinforcement Learning and Model Predictive Control (MPC)
  - Approaches that use models of the environment for decision-making and planning.
  - How MPC utilizes learned models to make sequential decisions.

Week 3

- Exploration Strategies and the Upper Confidence Bound
  - Thompson sampling and information gain.
  - The UCB algorithm and its applications in bandit problems and reinforcement learning.
- Off-Policy Learning and Deep Deterministic Policy Gradients (DDPG)
  - Off-policy methods including importance sampling and the use of experience replay.
  - How DDPG combines DQN-techniques with policy gradients for continuous actions.
- Multi-Agent Reinforcement Learning, and Decentralised and Centralized Approaches
  - Exploring scenarios where multiple agents interact and learn, including cooperative and competitive settings.
  - Different strategies for training multiple agents, both decentralized and centralized.
- Advanced Topics in Reinforcement Learning

	- Meta Reinforcement Learning, Transfer Learning, and Lifelong Learning. - Ethical and Societal Implications - Ethical considerations and real-world implications of reinforcement learning algorithms.
Key Texts:	Sutton, R.S., and Barto, A.G., *Reinforcement Learning: An Introduction*, Cambridge MA, 2020. Wiering, M., and Van Otterlo, M., *Reinforcement Learning: State of the Art*, Berlin, 2012.
Learning Outcomes:	By the end of this course, you will: - Understand the fundamentals of reinforcement learning, including agents, environments, and rewards. - Be able to assess and utilise a range of reinforcement learning approaches. - Be able to evaluate the efficacy of a range of reinforcement learning methods. - Understand different strategies for training multiple agents, both decentralised and centralised. - Demonstrate familiarity with current research.
Admissions Requirements:	LMH Summer Programmes are designed for students who want to gain and develop knowledge in their chosen subject area. LMH Summer Programmes are intensive courses of study aimed at undergraduates who have completed one, two, or three years of their degree, or entry level postgraduate students. We will consider each applicant's academic ability and expect successful applicants to have a minimum grade point average equivalent to 2:1 level on the British grading scale. For example, this would mean at least a 3.2 GPA on the 4.0 grading scale in the United States, and 80% in China. **This course would suit STEM students with intermediate level experience in artificial intelligence and machine learning concepts and techniques, including those undertaking, or looking ahead to, graduate level study or research.** **Specifically, students on this course must have experience of the following topics:** **Knowledge of the deep learning libraries.** **Understanding of deep learning, neural networks and basic dynamic programming.** **Strong background in optimization and probability.** **Familiarity with the Python programming language.** To participate fully in the programme all students will need to have proficiency in English. English language requirements for students who are not native English speakers: - Overall TOEFL score of 85; - or IELTS score of 6.5 (no less than 6.0 in each component); - or CET-4 at 550 or CET-6 at 520. If the language of instruction in your home institution is English you do not need to provide evidence of your English proficiency.
Teaching Methods:	Core syllabus material will be covered in lectures. Students attend four lectures each week and each lecture lasts 90 minutes. Seminars in smaller groups offer students space to discuss and debate, to dig deeper into difficult concepts, and to explore their own ideas. Student contribution to seminars is vital, and tutors will ensure

Lady Margaret Hall, University of Oxford
Norham Gardens, Oxford, OX2 6QA, United Kingdom
vacation.programmes@lmh.ox.ac.uk
+44 (0)1865 274295

	everyone takes part in discussions. Seminars last 1 hour and students will take part in four seminars each week. Independent study is a crucial part of an LMH Summer Programme and of the Oxford teaching model. Tutors will recommend important reading to do between lectures and seminars that will enable students to come to class equipped to understand the information presented and prepared to take part in discussion and debate. Each week students will have an assignment of independent work to complete and submit in advance of the tutorial. There is an appropriate amount of space in the timetable to complete the necessary reading, preparation, and assignments. Students should expect to do around 40 hours of independent study each week. The final class each week is a tutorial, a very small class typically including only 2-4 students and central to the teaching methods used by the University of Oxford and on LMH Summer Programmes. Guided by their tutor, students will receive feedback on their assignments and be challenged to defend, justify, or even rethink their work and ideas. These rigorous academic discussions help develop and facilitate learning in a way that cannot be done with lectures and seminars alone.
Assessment:	On a three-week LMH Summer Programme students produce one piece of assessed work every week, which is submitted to the tutor and then discussed in a tutorial. At the end of each week students will receive a percentage grade for their submitted work. Each week's work counts for a third of the final percentage grade, so the final grade is an average of the mark received for each piece of work. Students who stay for six or nine weeks will receive a separate grade for each 3-week course.
Academic Credit:	Lady Margaret Hall will provide a transcript of students' assessed work, and can send this directly to your home institution if required. LMH Summer Programmes are designed to be eligible for academic credit, and we will communicate with home institution to facilitate this as needed. As a guide, we recommend the award of 15 CATS / 7.5 ECTS / 4 US Credits for each 3-week course.

Lady Margaret Hall, University of Oxford
Norham Gardens, Oxford, OX2 6QA, United Kingdom
vacation.programmes@lmh.ox.ac.uk
+44 (0)1865 274295