CSI5340 / ELG5214 – Winter 2026

Deep Learning & Deep Reinforcement Learning

Instructor: Colin Bellinger
Email: colin.bellinger@uottawa.ca
Office: SITE 5037
Office Hours: Prof. Bellinger - Fridays, 3:00–4:00 PM (SITE 5340)
Office Hours: TA. Alex De Furia - By request <adefu020@uottawa.ca>
Lecture: Thursdays, 2:30–5:20 PM, 200 Lees A130
Course dates: January 12 to April 15

Teaching and course evaluation period: March 30 to April 12

Reading week: February 15 to 21

Course Description

This course will focus on deep reinforcement learning (DRL), situating it within modern deep learning. You will develop theoretical foundations, master key frameworks (primarily PyTorch), and gain hands-on experience developing research questions, designing, training, and evaluating DRL agents. The course includes paper presentations, critiques and a reproducible research project.

Prerequisites

- Undergraduate-level courses in **linear algebra, calculus, probability & statistics.**
- Strong **Python** programming skills and fluency in **PyTorch**.
- **Prior exposure to basic DL/RL** concepts is expected.
- **JAX** familiarity is not required but beneficial for assignments/projects.

Students are expected to independently address any gaps in foundational skills.

Communication:

- Brightspace announcements are the primary channel.
- The teaching assistant (TA) is available to provide support on homework, presentations, and the course project. They will address any inquiries related to these areas. The default is to pose questions on the course discussion boards.
- The TA and Professor **will not responded** to emails **on the weekend**.

Change of Status

- Last day to withdraw and receive a financial credit: January 30, 2026
- Last day to withdraw from a course or an activity (no financial credit): March 27, 2026

Learning Outcomes

By the end of the course, students will be able to:

1. **Learn fundamentals** including optimization, universal approximation theorem, back-propagation
2. **Understand and train deep learning models** (MLPs, CNNs, RNNs/Transformers) in PyTorch.
3. **Formulate RL problems using MDPs** and apply value-based and policy gradient methods.
4. **Understand and train** DRL algorithms (DQN variants, PPO, SAC, TD3, DDPG) and **analyze their behaviour**.
5. **Evaluate DRL algorithms** including on-policy vs. off-policy methods, model-based RL, and offline RL.
6. **Critically review and present** current DRL literature and communicate findings clearly.
7. Develop curiosity and **provide constructive feedback** to peers.
8. **Design, implement a reproducible a well-scoped DRL research** project with clear baselines, ablations, and documented code.

Texts & References (all available online / via uOttawa library)

Goodfellow, Bengio, Courville – *Deep Learning* (MIT Press, 2016). http://www.deeplearningbook.org

Bishop, Christopher M. – *Deep Learning: Foundations and Concepts* (Springer, 2024). Library link: https://ocul-uo.primo.exlibrisgroup.com/permalink/01OCUL_UO/17kkpb9/cdi_askewsholts_vlebooks_9783031454684

Sutton & Barto – *Reinforcement Learning: An Introduction* (2nd ed. draft). http://incompleteideas.net/book/RLbook2020.pdf

François-Lavet et al. – *An Introduction to Deep Reinforcement Learning* (FnT ML, 2018). https://arxiv.org/pdf/1811.12560

Additional tools & docs (recommended)

- Basics of reinforcement learning (RL)
  - University of Alberta's RL specialization on Coursera
  - David Silver's RL course on YouTube
- PyTorch,
- Stable-Baselines3: https://stable-baselines3.readthedocs.io/en/master/
- CleanRL: https://docs.cleanrl.dev/
- Gymnasium: https://gymnasium.farama.org
- Mujoco
  - Mujoco tutorial
  - Creating custom environments in Mujoco
- JAX: https://docs.jax.dev/en/latest/jax-101.html

Course topics

The topics below are subject to change based on the pacing and interest of the class.

DL Foundations

- Optimization (SGD variants, momentum, Adam), loss landscapes.
- Regularization (weight decay, dropout, early stopping).

DL Architectures

- MLP, CNNs; RNNs, Transformers, Diffusion.

RL Foundations

- MDPs, Bellman equations, dynamic programming, policy evaluation/ improvement, Monte Carlo & TD learning.

Value-Based Deep RL

- Q-learning; DQN and variants (Double, Dueling, Prioritized Replay).

Policy Gradient & Actor–Critic

- REINFORCE; on-policy actor–critic; A2C/A3C; entropy regularization.

Continuous Control & Exploration

- DDPG, TD3, SAC; exploration strategies & noise processes.

Model-Based RL

- Planning, MPC, Dyna-style methods; world models (e.g., Dreamer-style).

Advanced topics — Offline RL & Safety

- Offline RL, Safe RL Active Measure RL, hierarchical RL, etc.

Assignments

- Implementation and analysis focused
  - Implement core algorithms, run baselines, and produce clear analysis.
  - Submit code, a short report (PDF), logs/plots, and a **reproducibility checklist** (environment file, seed control, instructions).

Presentations

- **Student-led presentation of course content.**
  - Groups of 2
  - Topic / paper will be provided
  - Must explain the topic to the class, link it to the bigger deep RL picture, contextualize contributions, assess assumptions/limitations, and compare with related work.
- **Poster presentation**
  - Groups of 2 (sub group from project group)
  - DL/RL paper selected by the group from top conference or journal (e.g. NeurIPS/ ICML/ICLR/AAAI/TPAMI/JMLR/ICRA/CORL)
  - Recommendation: select one related to your project

Research Project

- Groups of **3–4** (randomly assigned).
- Topics: novel algorithmic idea, systematic benchmarking, application (robotics, games, operations, RLHF), replication study with extensions. Other project ideas are welcome but should be discussed with the TA in advance of the proposal submission.

- **Deliverables:** Proposal, Final Report & Artifact (Week 12), Presentation.
- **Reproducibility Requirements:**
  - Public **Git repository** (private until grading) with README, environment specification, seed control.
  - Clear **experiment scripts** and **config files**; ablations and baselines.
  - **AI Usage Log** (if tools used).
- **Peer Evaluation:** Individual contribution self/peer assessments may adjust individual project grades.

Grading

Participation – 10%: Engagement in in-class and online discussions, paper critiques, in-class activities.

Assignments – 20%: 3 programming-heavy assignments focused on DL/DRL implementation, analysis, and reproducibility.

Late Policy: Late assignments will not be graded (see *Extensions & accommodations* below).

Course topic presentation – 15%: Group (2) in-class presentation, online peer review, in-class question and answers. No AI use permitted.

Research paper poster – 15%: Group (2) in-class poster presentation, online peer review, in-class question and answers. No AI use permitted.

Project – 40%: Group (4) DRL research project with proposal, milestone, final report, and presentation. **AI use permitted** for idea generation/debugging/programming support with **full disclosure** (see *AI Policy*).

Course Policies

Late Work:

- Assignments submitted after the deadline will **not be graded**.
- For documented accommodations (medical, religious observance, etc.), contact the instructor **before** the deadline when possible.

Collaboration:

- Discuss concepts freely; **code and written work must be your own** unless explicitly group-based.
- Cite all external resources (libraries, blogs, papers, AI). Do not copy solutions.

Academic Integrity:

- Plagiarism and academic fraud—including uncredited use of **Internet resources or AI tools**—are serious offences and will be handled per **University of Ottawa** regulations.
- Providing or receiving unauthorized assistance is equally subject to sanctions.

Accessibility & Accommodations:

- Students requiring accommodations should contact **Student Accessibility Services** and inform the instructor early in the term.

AI Policy

You are here to learn. AI such as ChatGPT, Gemini, and Copilot can be a great tool to support this. However, it can also be a hindrance, preventing you from developing your deep understanding of the concepts that will support you to succeed in your graduate endeavours and future employment. Based on this, the following rules are defined to support you in developing your knowledge base and gaining experience using AI tools in research.

Best Practices
- Use AI as a supplement, not a substitute.
- Validate outputs against official documentation and research papers.

Assignments & Presentations: Use of AI tools (e.g., ChatGPT, Gemini, Copilot) are not permitted.

Project: AI tools are permitted for ideation, debugging, and coding assistance with explicit disclosure:

- Include an AI Usage Log (tool, how used).
- Cite AI-assisted sections in code comments and the report.
- You must understand and be able to explain all code/content generated.
- You are responsible for verifying correctness and avoiding plagiarism.

Software & Compute

- **Python 3.10+**, **PyTorch 2.x**, **Gymnasium/OpenAI Gym**, **Stable-Baselines3** (optional), **CleanRL** (optional).
- **Morningstar:** Docs provided by a previous TA: https://morningstar.uottawa.ca/help/index.html
- **JAX/Flax**.
- GPU strongly recommended (local or cloud). If needed, use **Google Colab** or **Morningstar**. It is important to be aware that Morningstar may be overloaded close to assignment and project deadlines. It is your responsibility to plan ahead if you are relying on Morningstar or other external resources.

Expected Workload

- ~9–12 hours/week outside of class for readings, assignments, and project work.
- Readings posted weekly; students are responsible for staying current.

Tentative Deliverable Timeline

- Course Topic Presentation (Group): Present a lecture related topic - January 27 - March 26, 2026
- Homework 0 (individual): Setup & Reproducibility - January 23, 2026
- Homework 1 (individual): JAX Versus Pytorch MLP - February 3, 2026
- Project Proposal (Group): 1 page document - February 24, 2026
- Poster Session (Group): Present a paper related to group project - February 26, 2026
  - **Note**: you are welcome to print your poster on multiple standard sheets of paper or a basic poster paper. The cost of printing black and white on standard paper at uOttawa is 9 cents per page and 25 cents for colour printing. See: uOttawa Printing
- Homework 2 (individual): DQN Versus REINFORCE - March 17, 2026
- Homework 3 (individual): Paper critic - March 31, 2026
- Project Presentation (Group): Presentation - April 2, 2026
- Project Presentation (Group): Presentation - April 9, 2026
- Project Paper and Github Code (Group): Submit final project documents - April 24, 2026

Attendance to classes

Class attendance is mandatory. As per academic regulations:

"*to be admitted to the final examination in a subject, a student must attend a minimum of 80% of classes and must not have more than five unauthorized or unjustified absences in that subject.*"

Plagiarism

Academic fraud is an act by a student that may result in a false evaluation (including papers, tests, examinations, etc.). It is not tolerated by the University. Any person found guilty of academic fraud will be subject to severe sanctions. Here are some examples of academic fraud:

- Plagiarism or cheating of any kind;
- Present research data that has been falsified;
- Submit a work for which you are not the author, in whole or part;
- Submit the same piece of work for more than one course without the written consent of the professors concerned.

Please consult this webpage: it contains regulations and tools to help you avoid plagiarism. An individual who commits or attempts to commit academic fraud, or who is an accomplice, will be penalized. Here are some examples of possible sanctions:

- Receive an “F” for the work or in the course in question;
- Imposition of additional requirements (from 3 to 30 credits) to the program of study;
- Suspension or expulsion from the Faculty.

You can refer to the regulations on this webpage.

Information sharing and copyright

All documents prepared by the course instructor, including assignments, course notes, and exams, are protected by copyright. Copying, digitizing, or publishing on a web site is therefore a violation of copyright and is illegal.

Health and Counselling service

Please take care of your physical and mental health during the term. Physical activity and getting out in nature (even during an Ottawa winter) may relieve stress, promote happiness and support an successful academic term.

The University offers many opportunities for health, fitness and sport. Explore options here.

There are many reasons to take advantage of the Counselling Service. They offer:

- Personal counselling
- Career counselling
- Study skills counselling

Further information is available here.

Access service

The Access Service acts as an intermediary between students, their faculty and other University offices to ensure that the special needs of these students are addressed and that the best possible learning conditions are being offered.

Note that the University of Ottawa is affiliated with AERO and ACE services for the adaptation of accessible academic materials for students with perceptual disabilities. If you have any questions, please contact the Accessibility Librarian or check the Access services for textbooks.

Policy – Prevention of sexual violence

The University of Ottawa will not tolerate any act of sexual violence. This includes acts such as rape and sexual harassment, as well as misconduct that take place without consent, which includes cyberbullying. The University, as well as various employees and student groups, offers a variety of services and resources to ensure that all uOttawa community members have access to confidential support and information, and to procedures for reporting an incident or filing a complaint. For more information, please visit https://www.uottawa.ca/about-us/respect/sexual-violence-support-and-prevention

Academic regulation A-1 on bilingualism

Except in programs and courses for which language is a requirement, all students have the right to produce their written work and to answer examination questions in the official language of their choice, regardless of the course's language of instruction.