NYU TANDON SCHOOL OF ENGINEERING

New York University Tandon School of Engineering
Computer Science and Engineering

# CS-GY 9223 L
# Language Models, Reinforcement Learning, Reasoning, Fall 2026

Professor Pavel Izmailov
Thursday 2:00-4:30pm, Jacobs Hall, 6 Metrotech, Room 204 (Brooklyn Campus)
September 2, 2026 – December 14, 2026

## PROFESSOR CONTACT

**Email:** pi390@nyu.edu.

**Office Hours:** See course webpage.

## COURSE ASSISTANTS

**TBD** : Course assistants will be announced before the first lecture.

**Office Hours:** See course webpage.

## ESSENTIAL LINKS

**Course Web page:** TBD.

**Brightspace:** TBD.

**Ed Discussion:** TBD.

## COURSE DESCRIPTION

This is a graduate seminar-style course on modern large language models (LLMs), reinforcement learning (RL), and reasoning. The goal is to bring students to the frontier of the field: we will read and discuss contemporary research papers and technical reports, and work through the engineering and scientific decisions behind today's state-of-the-art systems.

Topics include:

- **Modern LLM architectures.** We will take a recent frontier-scale model and go through it end to end: tokenization, attention variants, mixture-of-experts, long-context mechanisms, normalization and initialization choices, optimizers, parallelism, and the many implementation tricks that make these systems work.
- **Reinforcement learning foundations.** Markov decision processes, value-based methods, policy gradients, actor-critic methods, DQN, and the AlphaGo / AlphaZero line of work.
- **RL for language models.** RLHF and preference optimization, PPO, GRPO and its many variants, reinforcement learning with verifiable rewards, reward modeling, and reward hacking.
- **Reasoning and inference-time scaling.** Chain-of-thought, self-consistency, search and verification at inference time, reasoning models, and agentic / tool-use training.
- **Scaling laws.** Pre-training scaling laws, compute-optimal training, data-constrained scaling, and emerging results on scaling reinforcement learning.
- **Video generation and world models.** Diffusion and autoregressive video models, world models, and JEPA-style predictive architectures.

Evaluation is primarily project-based. There are no written exams.

## COURSE PREREQUISITES

This is an advanced course. Students are expected to have taken a graduate-level machine learning or deep learning course (e.g., CS-GY 6923 or equivalent) and to be comfortable with the following:

1. **Probability and statistics**: random variables, expectation and variance, conditional and joint distributions, Monte Carlo estimation, maximum likelihood.
2. **Linear algebra**: matrix and vector operations, norms, eigen- and singular value decompositions.
3. **Calculus and optimization**: gradients, chain rule, stochastic gradient descent and its variants, backpropagation.
4. **Machine learning and deep learning**: supervised learning, generalization, neural network architectures, training and regularization; familiarity with transformers is strongly recommended.
5. **Programming**: fluency in Python and experience with a deep learning framework (PyTorch or JAX). Projects will involve non-trivial implementation work.

If you are not familiar with these topics, you will need to study them on your own; they will not be re-taught in class.

## COURSE STRUCTURE AND GRADING

There is one 2.5-hour class meeting per week. Meetings combine instructor lectures with student-led paper presentations and discussion. Most of your learning will happen through reading papers, participating in discussion, and working on your course project.

**Class participation and reading responses (20% of grade).** Students are expected to come to class having read the assigned paper(s) and to contribute to discussion. Short written reading responses will be submitted before class for a subset of the lectures.

**Paper presentation (40% of grade).** Each student (or small group, depending on enrollment) will present one research paper or technical report to the class, including background, method, results, and a critical assessment.

**Course project (40% of grade).** Students will complete a substantial research or implementation project, individually or in small groups. The project grade is broken down as: project proposal (5%), midterm milestone report (5%), final presentation (10%), and final report (20%). Projects may be empirical (e.g., training or fine-tuning models, reproducing a result, running a careful ablation), analytical, or a thorough survey with original synthesis.

**The grading weights above are preliminary.** Final weights will depend on enrollment and course format, and will be announced at the start of the semester. There are no midterm or final exams in this course.

## COURSE POLICIES

**Assignments and submissions:** Reading responses, project reports, and presentation materials are turned in via Gradescope / Brightspace. Reports should be prepared in LaTeX using a standard machine learning conference template (e.g., NeurIPS or ICML).

**Late work:** Reading responses are not accepted late, since they support in-class discussion. Extensions on project deliverables are only possible under extenuating circumstances and with prior permission from the instructor.

**Collaboration policy:**

- Project work is expected to be done within your project group. Discussion of ideas across groups is encouraged, but each group must write its own code and its own report.
- Reading responses must be written individually.

- We have a zero tolerance policy for copied text, code, or results: any submissions with duplicate or very similar material will receive a zero, for both the student who copied and the student they copied from.
- If your project builds on existing code, models, or datasets, you must clearly state what was pre-existing and what you contributed.

**ChatGPT, Claude Code, and other AI tools:** You may freely use AI tools for your project work and for reading assistance — these tools are part of modern research practice, and this is a course about them. However, (i) you are fully responsible for the correctness of everything you submit, (ii) any substantive use of AI tools in your project should be described in your report, and (iii) make sure the tools do not impede your own learning. Reading responses and in-class discussion should reflect your own understanding.

**Grade changes:** Fair grading is very important to me. If you feel that you were incorrectly docked points, please contact the course assistants, who will escalate to me as necessary. Do not wait until the end of the semester.

**Questions about performance:** If you are struggling in the class, contact me as soon as possible so that we can discuss what's not working. It is difficult to address questions about performance at the end of the semester, or after final grades are submitted: by Tandon policy no extra-credit or makeup work can be used to improve a student's grade once the semester closes. So if you are worried about your grade, seek help from me and the course assistants *early*.

## READINGS

*There is no textbook to purchase.* The primary course material consists of research papers, technical reports, and lecture slides, all posted on the course webpage. Useful background references, all freely available online, include:

- *Reinforcement Learning: An Introduction* (2nd edition), by Richard Sutton and Andrew Barto — the standard reference for RL foundations. Available here.
- *Deep Learning*, by Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Available online.
- *Probabilistic Machine Learning: Advanced Topics*, by Kevin Murphy. Available here.

## COURSE SCHEDULE

Because this is a fast-moving field, the lecture schedule and reading list will be finalized shortly before the semester begins and adjusted as the semester progresses. An up-to-date schedule will be maintained on the course webpage.

## COURSE OBJECTIVES

1. Students will develop a working understanding of how modern large language models are built, from architecture and pre-training through post-training and deployment.
2. Students will learn the foundations of reinforcement learning and how those foundations are adapted — and modified — when applied to language models.
3. Students will be able to read, critically evaluate, and present contemporary research papers and technical reports in this area, and to identify their strengths, weaknesses, and open questions.
4. Through the course project, students will gain hands-on experience designing and running experiments with language models and RL methods, and communicating the results in the form of a research report.
5. A major goal is to prepare students to contribute to research or advanced engineering work on frontier AI systems, and to keep up with a field that changes on a timescale of months.

# ADDITIONAL INFORMATION

## MOSES CENTER STATEMENT OF DISABILITY.

If you are student with a disability who is requesting accommodations, please contact New York University's Moses Center for Students with Disabilities (CSD) at 212-998-4980 or mosescsd@nyu.edu. You must be registered with CSD to receive accommodations. Information about the Moses Center can be found here.

## NYU SCHOOL OF ENGINEERING POLICIES AND PROCEDURES ON ACADEMIC MISCONDUCT.

As an NYU community member, it is your responsibility to know your rights and responsibilities around academic misconduct. Please click here to review the NYU Tandon Student Code of Conduct. Any questions about how this policy relates to this class, please ask your professor.

A. Introduction: The School of Engineering encourages academic excellence in an environment that promotes honesty, integrity, and fairness, and students at the School of Engineering are expected to exhibit those qualities in their academic work. It is through the process of submitting their own work and receiving honest feedback on that work that students may progress academically. Any act of academic dishonesty is seen as an attack upon the School and will not be tolerated. Furthermore, those who breach the School's rules on academic integrity will be sanctioned under this Policy. Students are responsible for familiarizing themselves with the School's Policy on Academic Misconduct.

B. Definition: Academic dishonesty may include misrepresentation, deception, dishonesty, or any act of falsification committed by a student to influence a grade or other academic evaluation. Academic dishonesty also includes intentionally damaging the academic work of others or assisting other students in acts of dishonesty. Common examples of academically dishonest behavior include, but are not limited to, the following:

   1. Cheating: intentionally using or attempting to use unauthorized notes, books, electronic media, or electronic communications in an exam; talking with fellow students or looking at another person's work during an exam; submitting work prepared in advance for an in-class examination; having someone take an exam for you or taking an exam for someone else; violating other rules governing the administration of examinations.
   2. Fabrication: including but not limited to, falsifying experimental data and/or citations.
   3. Plagiarism: intentionally or knowingly representing the words or ideas of another as one's own in any academic exercise; failure to attribute direct quotations, paraphrases, or borrowed facts or information.
   4. Unauthorized collaboration: working together on work meant to be done individually.
   5. Duplicating work: presenting for grading the same work for more than one project or in more than one class, unless express and prior permission has been received from the course instructor(s) or research adviser involved.
   6. Forgery: altering any academic document, including, but not limited to, academic records, admissions materials, or medical excuses.

## NYU SCHOOL OF ENGINEERING POLICIES AND PROCEDURES ON EXCUSED ABSENCES.

The complete policy can be found here.

A. Introduction: An absence can be excused if you have missed no more than 10 days of school. If an illness or special circumstance has caused you to miss more than two weeks of school, please refer to the section labeled Medical Leave of Absence.

B. Students may request special accommodations for an absence to be excused in the following cases:

   1. Medical reasons
   2. Death in immediate family
   3. Personal qualified emergencies (documentation must be provided)
   4. Religious Expression or Practice

Deanna Rayment, advocacy.tandonstudentlife@nyu.edu, is the Coordinator of Student Advocacy, Compliance and Student Affairs and handles excused absences. The Office of Student Advocacy is located in 5 MTC, LC240C and they can assist you should it become necessary or if you have any questions.

## NYU SCHOOL OF ENGINEERING ACADEMIC CALENDAR

The full calendar can be found here. Please pay attention to notable dates such as Add/Drop, Withdrawal, etc. For confirmation of dates or further information, please contact Susana Garcia: sgarcia@nyu.edu.