EECE 7268 / CS 7268 - Verifiable Machine Learning
Fall 2025

Instructor	Prof. Michael Everett
HCAs	TBD
Lectures	MW 2:50pm-4:30pm, Location: TBD
Contact	Please post all course-related questions to Piazza to enable efficient sharing of knowledge between classmates and course staff.
Gradescope	TBD

Course Description

This class provides an introduction to the emerging field of verifiable machine learning with the goal of preparing students with (i) an understanding of key limitations of current learning-based methods with respect to safety and robustness properties, and (ii) a theoretical and practical understanding of ongoing research that aims to address these issues. Topics covered will include: adversarial robustness, uncertainty quantification, safety verification, reachability analysis, safe/robust reinforcement learning, and neural network relaxations, with applications in robotics/control, image classification, medical diagnosis, and language modeling.

Prerequisites

There are no prerequisites for this course.

Helpful Background Knowledge

Students will benefit from having some prior experience with machine learning and/or reinforcement learning. Familiarity with a deep learning framework (e.g., PyTorch, TensorFlow, JAX) would also be helpful.

Intended learning outcomes

Upon completion of this course, you will be able to:

1. Attack learned models to cause them to behave undesirably
2. Identify various forms of uncertainty and develop models that quantify confidence in their decisions
3. Understand the theory and practical tradeoffs of various neural network relaxations
4. Implement formal safety/robustness verification on a learned model
5. Deploy these techniques in safety-critical applications that interest you

Course materials

Textbooks: There are no required textbooks for this course.

Papers: We will also refer to selected papers from the literature throughout the course, which will all be available online.

Computing resources: We will distribute a Jupyter notebook in Colab for each coding assignment. You are welcome to do the assignments in Colab, which would only require a computer with internet access (Colab will use free cloud resources to run your code). If you want to develop your code locally, which many students prefer over Colab, it will be up to you to configure your environment. This will probably just involve installing the right python packages.

Course format

The course will consist of two meetings per week. In general, we will alternate between an introductory lecture on Friday, and a class-wide discussion of state-of-the-art research papers the following Wednesday (students have the weekend + Monday + Tuesday to read the papers). The lectures will aim to cover a framework for thinking about the topic and an introduction to relevant mathematical tools. The paper discussion meeting will be split into 2 halves: in the first half, each student will present for 3-5min about a recent paper of their choosing on the topic; in the second half of the meeting, we will have an instructor-led discussion of one paper that everyone has read.

To provide hands-on experience with state-of-the-art research in verifiable ML, there will be 3-4 coding assignments. There will be a final project, in which students will either (i) apply or extend ideas from the course into a problem related to their research, or (ii) re-implement a state-of-the-art research paper "from scratch."

Course assessment

The course will be assessed on the basis of the homework assignments (25%), class participation (25%) and final project (50%); there will be no exams.

This course will employ a *mastery-based* grading scheme; in particular, ***there will be no curving***. Final grades will be assigned according to the final scale:

Total %	Final grade
$\geq 90\%$	A
$\geq 80\%$	B
$\geq 70\%$	C
$\geq 60\%$	D
$< 60\%$	F

Late Policy

Each student can miss 2 of the group discussion/presentation days without penalty. After that, each missed discussion/presentation day will be marked as a 0 for that day's participation.

For coding assignments, late solutions will not be accepted.

We also understand that things come up during the semester. For circumstances that allow planning ahead (e.g., conference travel), please contact Prof. Everett ahead of time to come up with a reasonable plan. For less predictable circumstances (e.g., medical or family issue), we can give more flexibility on an individual basis.

Expectations and policies

Collaboration and academic integrity: I encourage you to discuss the course with your colleagues! This is one of the best ways to both sharpen and expand one's own thinking on a subject. However, if you collaborate on any part of the homework, **you must declare** who you worked with and what was discussed. Most importantly, **any work that you submit must be written independently by you, and must reflect your own understanding**. More generally, please review Northeastern's Academic Integrity Policy.

Violations of the academic integrity policies will result in penalties ranging from a zero on an assignment to an F in the course to a referral to the dean, at the instructor's discretion.

Modifications to the course: the policies and course outline in this syllabus are subject to change, as needed, as the course proceeds.

Feedback & general problem-solving: My goal is for this course to be both enjoyable and informative. To that end, I welcome and encourage feedback (whether positive or negative) on any aspect of the course at any time. In particular, if some feature of it (or some extraneous circumstance) is making it difficult for you to learn, **please say something** – the sooner the better!

Example Schedule / Topics

EECE 7398: Verifiable ML Schedule (Fall 2024)

Date	Topic	Notes
W 9/4	L01: Course Intro & Recap of Neural Networks	HW 1 out
F 9/6	L02: Adversarial Attacks	
W 9/11	Student Paper Presentations: Adversarial Attacks Group Paper Discussion: Paper	
F 9/13	L03: Formal Verification (Part 1)	HW 1 due HW 2 out
W 9/18	Student Paper Presentations: Formal Verification Finish L03	
F 9/20	**No Class**	
W 9/25	L03.5: CROWN & Duality	
F 9/27	Student Paper Presentations: Formal Verification Paper Discussion: Convex Relaxation Barrier	
W 10/2	Student Paper Presentations: Formal Verification L04: Formal Verification (Part 2)	
F 10/4	Paper Discussion: Beta-CROWN L05: Robust Training	HW2 due HW3 out
W 10/9	Student Paper Presentations: Verifying Neural Feedback Loops L06: Verifying Neural Feedback Loops	
F 10/11	Paper Discussion: OVERT	
W 10/16	**No Class**	
F 10/18	L07: Backward Reachability	
W 10/23	Student Paper Presentations: Verifying Neural Feedback Loops / Backward Reachability Paper Discussion: Backward Reachability	
F 10/25	L08: Safe / Robust RL	HW3 due
W 10/30	**No Class**	
F 11/1	L09: CLFs / CBFs	

W 11/6	Paper Discussion: CLFs/CBFs Student Paper Presentations	
F 11/8	L10: Uncertainty Quantification	
W 11/13	Finish L10 Student Paper Presentations	
F 11/15	Guest Lecture: Sushant Veer (Nvidia) Student Paper Presentations	
W 11/20	Guest Lecture: Nick Rober (MIT) Student Paper Presentations	
F 11/22	Guest Lecture: Alex Robey (CMU) ***** Class in EXP 610 *****	
W 11/27	Fall Break / Thanksgiving	No class
F 11/29	Fall Break / Thanksgiving	No class
W 12/4	Final project presentations	
F 12/6	Final project presentations	

Paper Discussion: Safe / Robust RL
Paper Discussion: Uncertainty Quantification

Recording of Classes

Depending on the recording capabilities available in the lecture room, we aim to record the Friday lectures to enable all students to review material at their own pace. Please contact Prof. Everett if you have any concerns.

Student Accommodations

Northeastern University and the Disability Resource Center (DRC) are committed to providing disability services that enable students who qualify under Section 504 of the REHABILITATION ACT and THE AMERICANS WITH DISABILITIES ACT AMENDMENTS ACT (ADAAA) to participate fully in the activities of the university. To receive accommodations through the DRC, students must provide the Disability Resource Center (DRC) with appropriate documentation that demonstrates a current substantially limiting disability.

For more information, visit http://www.northeastern.edu/drc/getting-started-with-the-drc/.

Diversity and Inclusion

Northeastern University is committed to equal opportunity, affirmative action, diversity and social justice while building a climate of inclusion on and beyond campus. In the classroom, members of the University community work to cultivate an inclusive environment that denounces discrimination through innovation, collaboration and an awareness of global perspectives on social justice. It is my intention that students from all backgrounds and perspectives will be well served by this course, and that the diversity that students bring to this class will be viewed as an asset. I welcome individuals of all ages, backgrounds, beliefs, ethnicities, genders, gender identities, gender expressions, national origins, religious affiliations, sexual orientations, socioeconomic background, family education level, ability – and other visible and nonvisible differences. All members of this class are expected to contribute to a respectful, welcoming and inclusive environment for every other member of the class. Your suggestions are encouraged and appreciated.

Please visit http://www.northeastern.edu/oidi/ for complete information on Diversity and Inclusion

TITLE IX

Title IX of the Education Amendments of 1972 protects individuals from sex or gender-based discrimination, including discrimination based on gender-identity, in educational programs and activities that receive federal financial assistance. Northeastern's Title IX Policy prohibits Prohibited Offenses, which are defined as sexual harassment, sexual assault, relationship or domestic violence, and stalking. The Title IX Policy applies to the entire community, including male, female, transgender students, faculty and staff. In case of an emergency, please call 911.

Please visit www.northeastern.edu/titleix for a complete list of reporting options and resources both on- and off-campus.