Introduction

Course objectives

1. Read, present, and participate in discussing academic work that uses machine learning in security & privacy research
2. Engage in research through two assignments and a group project

Course objectives

You should take this course if...

- You want to learn to do research
- You are interested in the subjects discussed in this course: machine learning, security & privacy
- You want to challenge yourself; you are not here for an easy grade

This course is designed as a graduate-level course cross-listed for undergraduate students.

Class

Phase 1: 4 weeks of lectures.

- Machine Learning
- Security
- Research

Heavily condensed; ML and Security are full courses...

- You won't be able to master everything in class; you need to spend time offline on these topics.

Two long breaks in each lecture.

Class

Phase 2: 8 weeks of reading academic papers.

- For each paper, 20 minutes of presentation followed by 20 minutes of discussions
- One student will present each paper
- Papers and signup link are on CourSys website
- 3 or 4 papers per week
  - If there are 3 papers, I will end with a review of the papers and the field

Grading

Detailed rubrics are on CourSys.

- Participation: 20%
- Assignments: 20%
- Paper Presentation: 10%
- Project: 30%
- Oral Exam: 20%

Participation

- In Phase 2, you are expected to attend each class for discussion
- Take notes when reading each paper so you can have something to ask
- Answering questions and adding to answers is also encouraged
- Please don't be afraid to argue
- Don't be afraid to veer off-track; you are encouraged to talk about other papers and non-academic topics

Assignments

Assignment 1: Critically evaluate an academic paper that uses machine learning in security & privacy.

- Choose a paper that isn't in Phase 2.
- Identify weaknesses in threat model, experiments, data, methodology, presentation, and replicability.
- Provide evidence for your claims.
  - If the data is malformed: fix it, and show how it changes the results.
  - If a table is poorly presented: Show how you would present it.
  - If the evaluation is lacking: Present your own evaluation with their data.

Assignments

Assignment 2: Re-evaluate a security & privacy problem using new ML classifiers, methods, and tools

- Example: Can transformer architectures help in detecting DDoS?
- Example: Do CNNs succeed in smudge attacks?
- Example: Can we attack malware detection tools with adversarial perturbations?
- Start by finding published data, or creating it

Further details will be posted soon.

Paper presentation

Present a paper listed on the course website.

- You are encouraged to make your own diagrams/figures
- Make your presentation understandable
- Demonstrate/develop your speaking and presenting skills

Project

Up to 3 students per group. Conduct original research and present the results.

- You can extend from Assignment 1/2.
- There will be a presentation in the last week; submit your report after.
- Start group work and resolve issues early!
- I am considering that you have limited data collection capacity and computational power; do your best.

You can submit a project proposal and ask me for feedback. This is optional.

Oral Exam

In the exam period, you will have an oral exam; you will talk to me about your project and the course material.

- Time: 20 minutes
- You will be asked specific details about your project. The intent is to verify your understanding of your own work.
- You will be asked to solve problems. These problems may be related to your project, a paper we read, or the lectures.
- Possible grades are fail (0%), pass (12%), satisfactory (20%), and excellent (25%), with minor deviations from those grades if necessary.
- Grading details and sample questions are on the course website.

AI Usage Policy

- You may use AI to assist you in writing.
  - However, you must fully disclose how you are using AI. Include your prompts and explain your process.
  - Anything in your work that could not have resulted from your disclosed AI usage process will be investigated as academic dishonesty.
  - Please note that AI-generated text often contains hallucinations, irrelevant information dumping, disorganized structure, inability to maintain focus on one concept, mysterious terminology, lack of rigor, and excessive self-praise. These issues will result in poor grades with or without the use of AI.
- You may use AI to assist with programming/engineering tasks. You do not need to explain your process.