ML for Security in Cyber-Physical Systems
EE/CSC 7700

Instructor: Xugui Zhou
Assistant Professor, ECE/CSE
xuguizhou@lsu.edu

Teaching Assistants: N/A

Class Time: Monday, Wednesday, and Friday; 10:30-11:20 am
Location: Patrick F. Taylor Hall 1218

Office Hours: Friday 3-4 pm at ERAD Building 333
TA Office Hours: N/A

Course Description

This course provides a general exploration of the intersection between machine learning (ML) techniques and security principles within the context of cyber-physical systems (CPS). Cyber-physical systems integrate computational and physical components to monitor and control various processes, including industrial systems, autonomous vehicles, and medical devices. As these systems become increasingly interconnected and autonomous, ensuring their security and safety against cyber threats becomes paramount.

The course begins with an overview of fundamental concepts in machine learning and cybersecurity, laying the groundwork for understanding their application in CPS environments. Students will then delve into advanced ML algorithms and techniques tailored to address security challenges in CPS, such as anomaly detection, threat prediction, and hazard mitigation.

Throughout the course, students will engage in hands-on exercises, simulations, and projects to reinforce theoretical concepts and develop practical skills in ML-based security for cyber-physical systems. By the end of the course, students will be equipped with the knowledge and tools necessary to design, implement, and evaluate ML-driven security solutions tailored to the unique challenges of CPS environments.

Topics covered include (tentative):

- Introduction to cyber-physical systems and their security challenges
- Fundamentals of machine learning and its applications in CPS security
- Supervised, unsupervised, and reinforcement learning techniques
- Adversarial machine learning and its implications for CPS security
- Anomaly detection and outlier detection in CPS data streams
- Threat modeling and risk assessment in cyber-physical environments
- Runtime hazard prediction and mitigation for CPS control systems
- Secure machine learning techniques for protecting ML-enabled CPS

• Case studies and real-world applications of ML-based security solutions in CPS domains (e.g., autonomous driving)

Course Objectives

- Understand the fundamental concepts of cyber-physical systems and their security challenges.
- Learn and apply various machine learning techniques to identify and mitigate security threats in CPS.
- Develop skills to design and implement secure CPS using ML algorithms.
- Analyze real-world case studies to understand the application of ML in securing CPS.

Assessment and Evaluation

Attendance: 5%
Class Activities: 10%
Assignment: 20% (2 assignments)
Paper Presentation: 30%

- Slides, presentation, and discussion: 20%
- Blog post about the presentation and discussion: 10%

Final Exam/Project: 35%

- Initial plan: 5%
- Progress report: 5%
- Presentation: 10%
- Report: 15%

During the week of the paper presentation, one group of students (2 or 3 members) will be responsible for preparing the topic and leading the discussion during the class meeting, and another group of students will be tasked with writing a blog post about the class. Both groups should collaborate closely on the posted write-up.

Prerequisites:
A basic knowledge of computer science and programming (e.g., c/c++/Python). A basic understanding of fundamental ML concepts.

Course Schedule (tentative)**:**
xugui-zhou.github.io/teaching/EECS7700/schedule.html

Course Policy

Communication: Email is the primary mode of communication for course-related queries. Students should use their university email accounts and include the course code in the subject line. The instructor's office hours are available for additional help and discussion. Appointments can be scheduled outside of these hours if necessary.

Attendance: Students are required to attend each class in person. If a student can't attend a class for a valid reason (e.g., illness), he/she should inform the course instructor before the class. Any absence of a class without a valid reason will get a **10**% penalty per time.

Late Submission: Students are allowed to submit their assignments up to 5 days after the required deadline. However, a **10%** penalty will be applied to late submissions per day. For example, if an assignment is submitted 2 days late, 20% of the total marks will be deducted. Assignments will not be accepted if they are more than **5 days** late. After 5 days, a grade of **zero** will be assigned.

Academic Integrity: By enrolling in this course, you have agreed to abide by and uphold the LSU Honor Code (https://www.lsu.edu/saa/students/academicintegrity/index.php) of Student Conduct (https://www.lsu.edu/saa/students/codeofconduct.php). All students are expected to fully comply with all the provisions of the University's Honor Code. Any attempt to take credit for work done by another person is considered plagiarism and a violation of the honor code. You are encouraged to study in groups and work together for final projects or group activities, but you are expected to finish the assignments independently. Acts of academic misconduct, including cheating, plagiarism, fabrication, and unauthorized collaboration, are strictly prohibited and will be forwarded to the Honor Committee and result in severe penalties such as failing grades.

Discrimination and Power-based Violence: The university is dedicated to providing a safe and equitable learning environment for all students.
1. Power-based personal violence will not be tolerated.
2. Everyone has a responsibility to do their part to maintain a safe community on Grounds.

Week 1: Introduction

- L1: Introduction to CPS
- L2: Security Threats in CPS
- L3: Existing security solutions in CPS

HW1 release, due in one week.

Week 2: Machine Learning for Security

- L4: Basic Concepts of Machine Learning
- L5: Supervised learning
- L6: Unsupervised learning

Week 3: Machine Learning Application

- L7: Tools and platforms
- L8: Data collection, labeling, training, testing
- L9: Evaluation metrics and performance considerations

HW2 release, due in two weeks.

Week 4: Safety Validation in CPS

- L10: Basic concepts of safety validation
- L11: Fault injection techniques in CPS
- L12: Safety validation applications

Week 5: Safety Monitoring and Mitigation in CPS

- L13: Intrusion and Anomaly Detection
- L14: Safety Monitoring
- L15: Hazard Mitigation

HW3 release, due in two weeks.

Week 6: Adversarial Machine Learning

- L16: Introduction to adversarial machine learning
- L17: Adversarial attack techniques
- L18: Adversarial defense techniques

Week 7: Reinforcement Learning

- L19: Basic concepts
- L20: Dynamic Programming
- L21: RL platforms

HW4 release, due in two weeks.

Week 8: Reinforcement Learning Practice

- L22: Temporal Difference Learning
- L23: Q-learning
- L24: RL application in CPS security

Lxx: Context-Aware Safety Assurance

Week 9-13: Paper Presentations
Week 14-15: Final Project and Presentation

Homework 1:

1. Describe your understanding of CPS. (20 pts)
2. Explain how a CPS works using a specific example application (e.g., autonomous driving systems) (20 pts)
3. List a couple of safety/security threats to CPS. (20 pts)
4. What are the challenges in advancing CPS security/safety? (20 pts)
5. What do you expect to learn from this course? (20 pts)

Homework 2:
Train a model for image classification.
Add noise and train a defense model

Homework3:
Train an RL agent