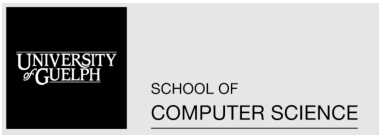
CIS*6670 Special Topics In Cybersecurity, Summer 2024



Course Information

Instructor:	Dr. Wenjing Zhang
Email:	wzhang25@uoguelph.ca
Lecture Time:	TBD
Lecture Type:	In-Person
Lecture Location:	University of Guelph Research Park, 150 Research Lane
Office Hours:	Online, Wednesdays 10:00 AM - 11:00 AM

Course Pages

1. CourseLink – www.courselink.uoguelph.ca (Primary).
2. Slido – In Class Quizzes and Attendance.

Course Description

CIS*6670 Summer 2024 is a specialized 12-week graduate-level course focused on the intersection of Artificial Intelligence/Machine Learning (AI/ML) and cybersecurity. Through three distinct parts, students will explore AI/ML applications in enhancing cybersecurity measures, delve into the security and privacy risks associated with the AI/ML lifecycle, and examine the impact of Generative AI and Large Language Models (LLMs) on the cybersecurity landscape.

Designed to equip students with a thorough understanding of AI/ML's benefits and challenges in cyber defense, the course emphasizes critical evaluation and application of AI/ML technologies within robust cybersecurity frameworks. It includes three practical assignments and an extensive project, enabling students to apply and demonstrate their knowledge through research or project work focused on AI/ML and LLMs within the cybersecurity context. Upon completion, students will be skilled in integrating AI/ML into cybersecurity strategies, ready to address emerging cyber threats, and capable of contributing to the advancement of sophisticated security solutions using AI/ML. More critically, they will demonstrate profound insights into research challenges and exhibit their capability for independent research in the interdisciplinary realms of Cybersecurity, AI, and Machine Learning.

Prerequisites (By Topics)

Fundamentals of Cybersecurity, Statistical Machine Learning, Deep Learning, Optimization, Probability Theory, Algorithms, Linear Algebra, and Programming.

Textbooks

TBD.

Course Learning Outcomes

Upon the completion of this course, students should have achieved the following objectives:

Part 1: AI/ML for Security and Privacy (Weeks 1-4)

- To understand the role and capabilities of AI/ML in cybersecurity.
- To analyze AI/ML approaches for intrusion detection, malware analysis, and network security.
- To evaluate the effectiveness of AI/ML applications in real-world cyber defense scenarios.
- To understand and apply various machine learning techniques to assess, enhance, and maintain privacy within data-driven systems.

Part 2: Security and Privacy Risks in AI/ML Lifecycle (Weeks 5-8)

- To identify and mitigate security risks during data collection, ensuring data integrity.
- To explore vulnerabilities in model training and deployment, developing strategies for secure operations.
- To analyze privacy breaches, apply privacy-preserving techniques such as Federated Learning and Differential Privacy.
- To understand adversarial attacks on machine learning models, their impacts, and how to defend against them using techniques such as adversarial training and robust optimization.

Part 3: Generative AI and LLMs in Cybersecurity (Weeks 9-11)

- To grasp the capabilities and functions of generative AI and Large Language Models (LLMs) in detecting and responding to cyber threats.
- To scrutinize security vulnerabilities specific to generative AI and LLMs, preparing defensive strategies.
- To investigate the application of generative AI and LLMs for enhancing security operations and incident response.

The course culminates in Week 12 with student presentations, where participants will demonstrate their understanding and application of the concepts learned through research or project work related to AI/ML and LLMs in cybersecurity. Upon completion, students will be adept at integrating AI/ML technologies into cybersecurity practices, prepared to tackle emerging cyber threats, and able to contribute to the development of advanced security solutions using AI/ML technologies. More importantly, students will be able to demonstrate a deep understanding of the research problem and showcase the their ability to conduct independent research in the interdisciplinary field of Cybersecurity, AI, and Machine Learning.

Course Schedule (12-Week Format, Tentative)

Part 1: AI/ML for Security and Privacy (Weeks 1-4)

Week 1: Introduction to AI/ML in Cybersecurity

- Course Introduction
- Overview of AI/ML
- Cybersecurity applications of AI/ML

Week 2: AI/ML for Security - IDS, Malware and Anomaly Detection

– AI/ML for Intrusion Detection Systems (IDS)
– AI/ML for malware analysis and generation
– AI/ML for anomaly and behavior pattern detection

Week 3: AI/ML for Security - Web and Network Security

– Phishing detection and spam filtering using AI/ML
– AI/ML techniques in password security and cracking
– Firewall rule testing with ML algorithms
– AI/ML role in incident response and content filter testing

Week 4: AI/ML for Privacy

– Estimation of privacy leakage in data-driven IoT systems using machine learning
– Design of machine learning-based privacy protection mechanisms
– Generation of privacy-aware synthetic data using machine learning techniques
– Data privacy-utility trade-offs for research and development purposes

Part 2: Security and Privacy Risks in AI/ML Lifecycle (Weeks 5-8)

Week 5: Security Risks in Model Training

– Common attacks including model poisoning, adversarial examples, backdoor attacks, model inversion attacks, and membership inference attacks.
– Mitigation strategies, such as data sanitization, adversarial training, Differential Privacy, and Secure Multi-party Computation.

Week 6: Security Risks in Data Collection and Model Deployment

– Common attacks including data poisoning, data leakage, data injection, data integrity attacks.
– Mitigation strategies, such as data validation and sanitization, anomaly detection: robust authentication and encryption: regular audits and monitoring: redundancy and backup.
– Ensuring security during the deployment phase
– Protecting deployed models from evasion, extraction, and adversarial exploitation, e.g., model stealing attacks

Week 7: Privacy-Preserving Machine Learning

– Exploration of real-world instances where privacy was breached in machine learning systems.
– Exploring techniques to ensure privacy in ML, e.g., Federated Learning, Differential Privacy.
– Examination of challenges and emerging trends in the field of privacy-preserving ML.

Week 8: Adversarial Machine Learning

– Definition, motivations, and examples of adversarial attacks
– Types of attacks, techniques used, and potential impact on machine learning models
– Techniques for defending against adversarial attacks, including adversarial training and robust optimization

Part 3: Generative AI and LLMs in Cybersecurity (Weeks 9-11)

Week 9: Generative Adversarial Networks (GANs)

– Definition, underlying principles, and primary applications.
– Security and privacy risks in GANs: exploration of how GANs can be used in creating deepfakes, potential data leakage, misuse of synthetic data, and defense mechanisms.

Week 10: Introduction to Generative AI and Large Language Models (LLMs)

– Capabilities and functions of generative AI and LLMs in cybersecurity
– The role of LLMs in security analytics and threat detection

Week 11: Security Risks and Defensive Strategies for Generative AI and LLMs

– Analyzing security vulnerabilities specific to Generative AI and LLMs
– Developing defenses against the misuse of generative AI and LLMs

Week 12: Student Presentations

– Final group project presentations
– Course wrap-up, discussion, and feedback session

Grading Policy

Grades will be assigned based on the performance *projects*, *practical assignments*, and *class participations*. The weights assigned to each component are listed as below.

Criteria	**Percent of Grade**
Project	67%
(Initial Proposal, Due Week 4, via Dropbox on Courselink)	(for approval only)
(Status Report, Due Week 8, via Dropbox on Courselink)	(20%)
(Final Project Presentation, Due Week 11, In-Class)	(15%)
(Final Report, Due Week 12, via Dropbox on Courselink)	(22%)
(Code Demonstration Video, Due Week 12, via Dropbox on Courselink)	(10%)
Assignments	30%
(Assignment 1, Open Week 4, Due Week 6, via Dropbox on Courselink)	(10%)
(Assignment 2, Open Week 6, Due Week 8, via Dropbox on Courselink)	(10%)
(Assignment 3, Open Week 8, Due Week 10, via Dropbox on Courselink)	(10%)
Class Participation (Attendance will be taken randomly during three lectures)	3%

Table 1. Grading Criteria

Assessment Details

1. Project (67%)
   Projects will be conducted in groups of 2-3 students. Unless in exceptional circumstances, students in the same group will receive the same grade.
   (1) Initial Proposal (not for grading, for approval only): due Week 4, via Dropbox in Courselink. Students will propose a research problem/topic related to the interdisciplinary field of Cybersecurity, AI, and Machine Learning, along with appropriate datasets. They will then employ various AI/ML approaches to solve the problem. By Week 4, students are required to submit a 1-page proposal, which will be reviewed and approved by the instructor within one week.

(2) Status Report (20%): due Week 8, via Dropbox in Courselink.
For the status report, students are expected to provide an update on the progress of their AI and machine learning research project. The report should include:
- Overview: Briefly summarize the research problem and the chosen datasets.
- Progress Update: Describe the progress made so far in addressing the research problem. Include details about any experiments conducted, algorithms implemented, and any preliminary results obtained.
- Challenges Encountered: Discuss any challenges or obstacles faced during the research process. This could include issues with data collection, algorithm implementation, or unexpected results.
- Plan for Remaining Weeks: Outline the plan for the remaining weeks of the project. Specify the tasks that need to be completed, any additional experiments to be conducted, and the timeline for completion.
- Next Steps: Provide a brief overview of the next steps in the research process, including any adjustments to the research plan based on the progress made so far.

The status report should be submitted via Dropbox in Courselink by Week 8. It should be well-organized, clear, and concise, with a focus on providing a comprehensive update on the research project's status and future direction.

(3) Project Presentation (15%): due Week 11, in class project presentation.
Each group will present their results and assess other groups presentations.

(4) Final Report (22%): due Week 12, via Dropbox in Courselink.
The final report serves as a comprehensive documentation of your research project. It should include:
- Introduction: Provide an overview of the research problem, objectives, the significance of the project, novelty (if any), and your contributions.
- Literature Review: Summarize relevant literature and previous work related to the research problem.
- Methodology: Describe the methodology followed in addressing the research problem, including details about data preprocessing, feature engineering, model selection, and evaluation metrics.
- Experimental Results: Present the results obtained from the experiments conducted during the research process. Include tables, graphs, or figures to illustrate the findings.
- Discussion: Analyze the results and discuss their implications. Compare the performance of different AI/ML approaches and interpret the findings in the context of the research problem.
- Conclusion: Summarize the key findings of the research and provide recommendations for future work.
- References: List all the references (at least 12) cited in the report using a standard citation format.

The final report should be well-organized, clearly written, and properly formatted. It should demonstrate a deep understanding of the research problem and showcase the student's ability to conduct independent research in the interdisciplinary field of Cybersecurity, AI, and machine learning.

(5) 10-minute Pre-recorded Code Demonstration Video (10%): due Week 12, via Dropbox in Courselink.
In addition to the final report, students are required to submit their code in a ZIP file and a pre-recorded demonstration video showcasing their projects. The video should include:

- Code Demonstration:
  - Code Correctness: Demonstrate how the code works without encountering errors, including data preprocessing, model training, and evaluation.
  - Code Structure: Describe the organization of your codebase, including the main modules or packages.
  - Key Functions: Dive into key functions or classes, explaining the logic and its importance to the project.
- Results Showcase: Showcase the results obtained from running the code, including any visualizations or performance metrics.
- Discussion: Discuss any insights gained from the results and how they relate to the research problem. Discuss any significant challenges you encountered during development and how you addressed them.

The video should be well-produced, clear, and engaging. It should effectively communicate the student's understanding of the project and demonstrate their ability to implement machine learning algorithms to address real-world problems.

The final deliverable will be a report in the format of the International Conference on Machine Learning — please use LATEX style file to generate a PDF document for your report. A template will be provided on CourseLink. The report is due by Week 12 and must be submitted via the CourseLink Dropbox tool.

2. Assignments (30%)

(1) Assignment 1 (10%): Released Week 4, due Week 6, via Dropbox on Courselink.
(2) Assignment 2 (10%): Released Week 6, due Week 8, via Dropbox on Courselink.
(3) Assignment 3 (10%): Released Week 8, due Week 10, via Dropbox on Courselink.

The instructor expects the students to attempt solving each assignment individually. However, if a student is stuck on a problem, collaboration is encouraged under the following conditions:

- Students may discuss problems with any classmate, work on brainstorming solutions together, and go through potential solutions, but should avoid sharing complete solutions.
- After solving a problem, each student must independently write up their solution without referring to others' solutions or sharing their own.
- The write-up for each problem discussed must include the names of all students involved in the discussions. This will not impact their grade.

Detailed instructions for each assignment will be distributed during the term.

3. Class participation (3%)
Attendance will be taken randomly during three lectures, with each contributing 1% to the overall grade.

Course Policies

Presentations that are made in relation to course work - including lectures - cannot be recorded or copied without the permission of the presenter, whether the instructor, a student, or guest lecturer. Material recorded with permission is restricted to use for that course unless further permission is granted. Please note that the distribution of course materials is prohibited without the explicit permission of the instructor.

Submission Policies

Please be aware of the importance of adhering to deadlines. Assignments and Projects must be submitted by their respective due dates. Any late submissions will incur a penalty of 10% of the total grade for each day they are delayed. Please note that submissions will not be accepted if they are more than 2 days late, except in cases where prior arrangements have been made with the instructor. Additionally, once the solutions are posted, no further assignments will be accepted under any circumstances.

If you have any concerns about your grades, please discuss them with the instructor during office hours or reach out via email. Note that you should do this within one week of the grade being posted. Also, remember that a single point usually constitutes a small portion of your total grade. Therefore, I kindly ask that you consider the significance of your query to avoid overburdening the instructor with minor concerns.

Accessibility

The University of Guelph prioritizes establishing an environment free from barriers. Supporting students is a collective responsibility shared by students, faculty, and administrators. This cooperative effort is grounded in mutual respect for individual rights, the dignity of every individual, and a commitment from the University community to foster an open and supportive learning environment. Students in need of services or accommodations, whether due to a long-term identified disability or a short-term disability, are encouraged to reach out to Student Accessibility Services (SAS) at the earliest opportunity. For additional information, please contact Student Accessibility Services (SAS) at 1.519.824.4120 ext 56208, email accessibility@uoguelph.ca, or visit Wellness.uoguelph.ca/accessibility.

Academic Integrity

The University of Guelph is committed to the highest standards of academic integrity and honesty. Students are expected to be familiar with these standards, and must abide by the applicable policies (see The Academic Misconduct Policy in the Undergraduate Calendar).

The Academic Misconduct Policy is detailed in the Undergraduate Calendar:
http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-amisconduct.shtml

Health & Wellness

Should you encounter any personal challenges, please feel free to reach out to the instructor. Remember, the University of Guelph offers various resources to support you.

For medical concerns, contact Student Health Services at 1.519.824.4120 ext 52131.

For threats of violence or personal safety issues, contact Campus Police at 1.519.824.4120 ext 2000.

For psychological or emotional concerns, get in touch with Counselling Services at 1.519.824.4120 ext 53244.

For accessibility concerns, reach out to SAS at 1.519.824.4120 ext 56208.

In case of sexual assault, contact Campus Police at ext 2000, or Counselling Services at 1.519.824.4120 ext 53244.

For mental health concerns, visit https://wellness.uoguelph.ca/mental-health-support-services.