|  |
| --- |
| **ML for cybersecurity program schedule -- *Outline*** |
| Date Created: August 24, 2020 |
| Scheduled date for completion: September 21 (TBC) |
| **Prerequisites:**  To be best prepared to succeed in this program, students should have basic familiarity with:   * Programming: Proficiency with one or more programming languages such as Python/C/C++/MATLAB/Java/JavaScript * Basic Probability and Statistics: You should know the basics of probabilities, gaussian distributions, mean, and standard deviation * Linear Algebra: You should be comfortable with matrix/vector notation and operations * Computer Security: Basic knowledge of cybersecurity or applied computer security |
| **Target Audience:**  Information security managers, engineers, and professionals whose role includes working in applied computer security or cybersecurity.  Prior knowledge of machine learning is not required. |
| **Course Registration Survey:**  **Title/organization/ industry/years of experience/ highest degree obtained/major/ other coursework**  Rate your level of experience with the listed skills. On our scale, 1 = no knowledge about this area, and 5 = expert knowledge in this area.   * Web (HTML/Javascript) * Java * Python * C/C++ * Databases * Data Analytics * Jupyter Notebooks * Linear Algebra * Machine Learning * Computer Security   Please describe why you are interested in this program and what you hope to achieve by participating in it.  What topics or concepts are you most interested in discussing in this course? |

**Timeline:**

1. **September 21st: Program Outcomes (2-3), Module Outline & Learning Objectives (2-3) Finalized**
2. **October 5th: Slides Finished**
3. **Oct. 12th Video Recording at Home Begins (Suggestion to begin recording Module 1 Video the week of October 5 to build in more feedback and editing time)**
4. **Course materials available online (~ November 3)**
5. **Tuesday/Thursday 7 - 9 pm (November 10, 12, 17, 19)**
6. **Total Instruction Time per Session - 2 hours** 
   1. **Syn/Live Session Time:** ~ **2 Hour**
   2. **Asynchronous Content:** ~ **1 Hour**

**Course Delivery:**

Remote instruction, a mix of asynchronous and synchronous learning activities**:**

* Recorded videos, corresponding to the module content (about 20-45 minutes of video content per lecture slot).
* Online Zoom Q&A/discussion and case study sections to further discuss the content of each lecture in more depth/as a class (~ 90 mins)
* Additional opportunities for peer networking and 1:1 time with faculty (~ 30 mins)
* Final project: Case study development

**Lecture Format:**

* UCPE template in ppt
* Intro slide includes CDAC logo
* Include handwritten notes (via Goodnotes?) recorded during lecture (as needed)

Recording at Home - UCPE to send details on equipment and recording applications

|  |  |
| --- | --- |
| **Module 1** | November 10, 2020, 7-9pm CT |
| **Topic:** | Foundations of Machine Learning and Data Science for Security |
| **Description:** | A module focused on machine learning fundamentals, with applications to security. This module will offer an introduction to the data science pipeline, and teach fundamental building blocks, from data ingestion and feature engineering to machine learning model selection. |
| **Faculty Leads:** | Yuxin Chen (Lead), Nick Feamster |
| **Asynchronous Content:** | * Pre-course survey to understand student background, familiarity with concepts, what problems/topics they are interested learning about. * Pre-recorded video lectures(~ 20- 45 mins -- to be watched ahead of sync) * Videos are broken up into shorter 5-10 min videos focusing on:   + Industry use cases (motivation for concept/topic)   + Fundamental concepts   + Applications   + Failure cases (real-life examples where the application of ML to security has failed and the dangers of such failures) * Post-video survey/quiz to check for understanding/provide students with opportunity to ask questions that can be answered synchronously in class |
| **Synchronous Content:** | * Group discussion of core concepts and how they relate to students experiences in industry/work * Case studies/group work using jupyter notebooks/dummy data to provide hands on experiences with concepts * Networking -- potentially pair students based on skills/background |

**Module 1 Learning Objectives & Outline:**

**Module 1: Synchronous Schedule (via Zoom):**

|  |  |  |
| --- | --- | --- |
| **Topic** | **Time** | **Notes:** |
| Introduction/  Core Concepts Recap | 15 min | * Introduction * Recap of basic concepts in videos * Group discussion about experience with module concepts in industry/work experience |
| Introduction to case study/group work | 5 min |  |
| Simulation/  Case study/Group Work | 40 min | * TBD hands-on lab leveraging virtual case study (Jupyter Notebooks) * Divide students into Zoom breakout rooms * Pair students based on skill levels |
| Break | 10 min |  |
| Discussion/  Wrap up | 20 min |  |
| *Networking Opportunity* | *30 min* | * *TBD Virtual Happy Hour /speed meet-a-thon* * *Potentially led by class facilitator?* |

|  |  |
| --- | --- |
| **Module 2** | November 12, 2020, 7-9pm CT |
| **Topic:** | Data-Driven Network and Computer Security |
| **Description:** | A system-oriented security course, with grounding in fundamentals that we expect our audience will be familiar with, pivoting towards more data- driven and oriented concepts. For example, we aim to bridge the divide between familiar concepts such as signature-oriented anomaly detection to statistical anomaly detection. The course will also provide training in the underlying mechanics of machine learning as applied to practical problems in cybersecurity. |
| **Faculty Leads:** | Nick Feamster (Lead), Blase Ur |
| **Asynchronous Content:** | * TBD Pre-recorded video lectures(~ 20-45 mins) |
| **Synchronous Content:** | * TBD Case/Lab/Group Work * TBD Networking |

|  |  |
| --- | --- |
| **Module 3** | November 17, 2020, 7-9pm CT |
| **Topic:** | Machine Learning in the presence of adversaries |
| **Description:** | A course involving evasion and adversarial behavior in machine learning for security, including considerations of adversarial attempts to evade detection, pollution attacks on classifiers, backdoors in neural networks, and related topics in adversarial ML. |
| **Faculty leads:** | Yuxin Chen (Co-Lead), Nick Feamster (Lead), Blase Ur |
| **Asynchronous Content:** | * TBD Pre-recorded video lectures(~ 20-45 mins) |
| **Synchronous Content:** | * TBD Case/Lab/Group Work * TBD Networking |

|  |  |
| --- | --- |
| **Module 4** | November 19, 2020, 7-9pm CT |
| **Topic:** | Ethics, Fairness, Responsibility, and Transparency in Data-Driven Cybersecurity |
| **Description:** | The course will engage students in a series of programming assignments and case studies to expose them to ethical considerations associated with automated decision-making in the context of security. |
| **Faculty Lead:** | Blase Ur (Lead) |
| **Asynchronous Content:** | * TBD Pre-recorded video lectures(~ 20- 45 mins) |
| **Synchronous Content:** | * TBD Final Case Presentations * TBD Networking |