Syllabus

Course Number: CSCE 489/704

Course Title: Data Analytics for Cybersecurity

Instructor: Prof. Martin “Doc” Carlisle

Course Description

This course is an introduction to the theoretical foundations, algorithms, and methods of data analytics for cybersecurity. Cybersecurity is a societally critical topic with impacts across computing systems and networks, social and web-based communities, industrial control systems, and personal devices (among many others). Through the study and application of data analytics – including cluster analysis, supervised machine learning, anomaly detection, and visualization – this course will address a suite of cybersecurity topics including cyber-attacks, anomaly detection, vulnerability analysis, strategic manipulation and propaganda.

Course Prerequisites

Students should have taken a statistics course and a data structures course.

Course Learning Outcomes

At the end of this course the students should be able to:

1. Define and explain the key concepts and models relevant to data analytics for cybersecurity, including anomaly detection, data-driven vulnerability analysis, and supervised machine learning.
2. Design, implement, and evaluate the core algorithms underlying an end-to-end data analytics for cybersecurity workflow, including the experimental design, data collection, mining, analysis, and presentation of information derived from large datasets.
3. Apply "best practices" in data analytics for cybersecurity, including facility with modern tools.

Textbook and/or Resource Materials

There is no required textbook. Course readings will be available on Canvas and drawn from a variety of scholarly papers and other resources.

Assessments

Homework: There will be 4 individual programming-based homework assignments, each tackling a problem in data analytics for cybersecurity (e.g., anomaly detection in a large dataset of credit card transactions). Homework assignments will be in Python, using Jupyter notebooks.

Final Project: For the project, students will work in teams of two to three on a cybersecurity problem that demonstrates the students' facility with data analytics. At the end of the semester, teams will deliver a Jupyter notebook, paper, and two-minute video summarizing the project work.

Exams: Both exams are closed book. For all exams, students may bring one standard 8.5" by 11" piece of paper with any notes deemed appropriate or significant (front and back).

Graded Class Participation: Attendance in class and participation in the discussion are both important to success in the course. Students are expected to come to class, to ask questions and engage with the material, and to be an active participant on the course discussion group.

Graded for Stacked Course (UG/GR):

CSCE 704: your research goal should be advancing the body of knowledge by implementing data analysis—something that might be suitable for a conference publication.

CSCE 489: your research goal could be advancing the body of knowledge or attempting to reproduce results of others. If the latter, it is not sufficient to simply run someone else's code, you need to show that you have implemented a data analysis yourself.

The final project will be graded with these goals in mind.

Course Schedule

Modules and Lessons	Readings	Assignments
1,2: Introduction, Basics of Python and Jupyter Notebooks	https://codingthesmartway.com/getting-started-with-jupyter-notebook-for-python/ (Links to an external site.) Week1.ipynb	
3: Basics of Anomaly Detection	"An Intrusion Detection Model" Denning "The Base-Rate Fallacy…" Axelsson "Anomaly detection: A survey" Chandola et al.	
4: Basics of Machine Learning	Python Data Science Handbook, Chapter 5 (Links to an external site.) "Fast, Lean…Password Guessability…" by Melicher et al. "Modeling Password Guessing with Neural Networks" by de Castro et al.	Homework #1
5: Anomaly Detection in Practice	"CARDWATCH: A Neural Network…" by Aleskerov et. Al "Credit Card Fraud.. A review" by Abdou and Pointon A Behavior-cluster Based … Method for Credit Card Fraud Detection by Li and Xie "Detecting organized eCommerce fraud" Marchal and Szyller "A Survey of Data Mining… for Cyber Security Intrusion" by Buczak and Guven "A neural network approach for misuse and anomaly"… by Yu et al. "A training-resistant anomaly detection system" by Muller et al.	Homework #2 Homework #3
Midterm Exam		Exam #1
6: ML in practice: Malware detection	"JStap: … malicious JavaScript…" by Fass et al. "Accurate Malware Detection" by Copty et al.	
7: ML in practice: Phishing detection	Textual and Visual Content-Based Anti-Phishing… by Zhang et al. Diverse Datasets and A Customizable Benchmarking… by Zeng et al. Learning to Detect Phishing Emails by Fette et al	

8: Social Network Security	Analyzing Spammer's Social Networks Yang et al. Robust De-anonymization of Large Sparse Datasets by Narayanan and Shmatikov SocialSpamGuard by Jin et al. "Automatically Detecting Bystanders in Photos to Reduce Privacy Risks" by Hasan et al. "Browsing Unicity: On the Limits of Anonymizing Web Tracking Data" by Deuber et al.	Homework #4
9: Dark Web: Crawling, analysis, uncovering networks	Cybercrime in the Deep Web- Black Hat EU 2015 BlackWidow: Monitoring the Dark Web by Schafer et al. Identification of Illegal Forum Activities Inside the Dark Net by Alnabulsi and Islam IEDs in the Dark Web by Chen Darknet and Deepnet Mining for … Cybersecurity Threat Intelligence by Nunes et al.	Project Proposal
10: Strategic Manipulation, Propaganda, and Fake News	A Survey of Fake News by Zhou and Zafarani Extremist Propaganda Tweet Classification… by Nizzoli et al. Mixed-code text analysis for the detection of online hidden propaganda by Tundis et al. "FakeDetector: Effective Fake News Detection…" by Zhang et al. "A Sensitive Sylistic Approach to Identify Fake News…" by de Oliveira et al. "Unsupervised Fake News Detection" by Gangireddy et al.	Project Checkpoint

11: Insider Threats	A Survey of Insider Attack Detection Research by Ben Slaem et al "A New Take on Detecting Insider Threats…" by Rashid et al.	Project Checkpoint
Midterm Exam		Exam #2
12: Project Presentations		
Final Exam Period		Final presentation and report due during final exam