ENSURE RESEARCH GROUP

Energy-Aware, Secure and Reliable Computing

≡ MENU

Search … SEARCH

EL-GY-9163: MACHINE LEARNING FOR CYBER-SECURITY

Class Meeting Times: Thu 5 PM – 7.30 PM (ONLINE OR IN DIBNER LC 102)

Artificial intelligence (AI) and machine learning (ML) techniques are being increasingly deployed in cyber-security settings. Examples of critical applications include network anomaly detection, biometric authentication, spam detection, and data analytics based financial fraud detection. At the same time, advanced ML algorithms also give attacker's an advantage, setting up a complex interplay between attackers and defenders.

At the same time, ML systems are susceptible to new attacks. These include "adversarial input perturbations" which have to been shown to be particularly pernicious for deep neural networks and ML "backdooring" attacks that compromise the training process. For these reasons, there is growing interest in techniques to develop and deploy verifiably safe and secure ML systems, adopting and adapting techniques from the software security domain. A final vulnerability involves the fact that modern ML systems and especially deep learning systems are trained and executed in the cloud, raising concerns about the privacy of the user's data. New solutions are being developed to address these privacy concerns.

LECTURES

#	Topics	Link
1	*Foundations:* Introduction and Basics I: Point estimation, MLE, linear regression, bias-variance trade-offs	Slides
2	*Foundations:* Introduction and Basics II: Linear classification, clustering, feature selection	Slides

3	*Application*: Spam Filtering	Slides
4	*Security Vulnerability:* Adversarial Attacks on spam filters	Slides
5	*Foundations:* Deep learning basics	Slides
6	*Foundations, Application and Ethics:* Deep learning basics contd., Face recognition, Ethical concerns.	Slides
7	*Security Vulnerability:* Training data-poisoning attacks on deep learning	
8	*Security Vulnerability:* Adversarial input attacks on Deep Learning	
9	*Privacy:* Training data and model reconstruction attacks; differential privacy	
10	*Application*: Deep fakes and fake news detection.	
11	*Societal Implications:* Model accountability and interpretability	
12	*Societal Implications:* Investigating bias and fairness concerns	
13		
14		

PRINCIPAL INVESTIGATOR