Machine Learning Security, Privacy & Fairness Group

1. Introduction

Hello there, this is a weekly reading group on ML security, privacy & fairness. Generally, we will talk about these issues in ML systems and models, with related topics such as adversarial robustness of ML systems, privacy preserving models in different learning settings (e.g. federated learning, byzantine ML) and encrypted ML, etc.

Beyond this learning scopes, it is also our aim to encourage collaborations in this field through the reading group. Currently this reading group starts on Wednesday 4:30-5:30 pm, if you are interested in joining us (or to give a presentation), email Tao (tyu@cs.cornell.edu) to be added into the email list.

2. Schedule

1. Semi-supervised Knowledge Transfer for Deep Learning from Private Training Data. (ICLR 2017, Oral) Presenter: Shengyuan Hu, 9.18.2019
2. Differential Privacy Has Disparate Impact on Model Accuracy. (NeurIPS 2019); Deep Learning with Differential Privacy. (CCS 2016). Presenter: Tao Yu, 9.25.2019
3. CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy (ICML 2016); Low Latency Privacy Preserving Inference (ICML 2019). Presenter: Shengyuan Hu, 10.02.2019
4. Membership inference attacks against machine learning models. (2017 IEEE Symposium on Security and Privacy (SP)). Presenter: Tao Yu, 10.16.2019
5. NATTACK: Learning the Distributions of Adversarial Examples for an Improved Black-Box Attack on Deep Neural Networks. (ICML 2019); Black-box Adversarial Attacks with Limited Queries and Information. (ICML 2019). Presenter: Shengyuan Hu, 10.23.2019
6. Distillation as a Defense to Adversarial Perturbations against Deep Neural Networks. (2016 IEEE Symposium on Security and Privacy (SP)). Presenter: Junxiong Wang, 10.30.2019
7. Defense-GAN: Protecting Classifiers Against Adversarial Attacks Using Generative Models. (ICLR 2018); The Robust Manifold Defense: Adversarial Training using Generative Models. Presenter: Yiwei Bai, 11.06.2019
8. A tutorial on Machine Learning fairness. Papers included: On fairness and calibration. (NIPS 2017); How Do Fairness Definitions Fare?: Examining Public Attitudes Towards Algorithmic Definitions of Fairness. (AIES 2019) Presenter: Nathan Yan, 11.13.2019
9. A introduction on secure multi-party computation. Papers included: Communication-Efficient Unconditional MPC with Guaranteed Output Delivery. (IACR-CRYPTO-

2019); Lecture Notes: https://www.cs.umd.edu/~jkatz/gradcrypto2/f13/lecture13.pdf. Presenter: Ke Wu, 11.20.2019