NXP and Microsoft Demonstrates Edge-to-Cloud Machine Learning Solution for Predictive Maintenance

NXP and Microsoft Bring Microsoft Azure Sphere Security to the Intelligent Edge with a New Energy-Efficient Processor

Microsoft announces latest Windows IoT innovations for intelligent edge devices at Embedded World

By David Lemson / Director of Program Management, Windows IoT and Networking
4-step plan to make smart devices

1. Gather data
2. Label data
3. Compute ML model
4. Deploy ML model
Where Machine Learning, Security & Privacy Intersect

Machine Learning can **contribute** to IoT Security – but Machine Learning itself must be **secured**.

**Main topic**

- Confidentiality
- Adversarial examples
- Integrity & Authenticity
- Privacy

Improve safety and security of ML Systems

**Security of ML**

- Intrusion detection
- Fraud detection
- Control Flow protection
- SCA
- API/protocol

Apply ML in products to help defeat security attacks

Defend against attacks enabled by ML
01. Interpolation vs. Extrapolation
• Biggio, Corona, Maiorca, Nelson, Srdic, Laskov, Giacinto, Roli: Evasion attacks against machine learning at test time. In Machine Learning and Knowledge Discovery in Databases, 2013.
• Szegedy, Vanhoucke, Ioffe, Shlens, Wojna: Rethinking the inception architecture for computer vision. In IEEE conference on computer vision and pattern recognition, 2016.
Model Cloning

Image source: Matrix Revolutions movie poster
Example: Microsoft Azure Emotion Recognition

MLaaS: Machine Learning as a Service

https://azure.microsoft.com/en-us/services/cognitive-services/emotion
Example: Microsoft Azure Emotion Recognition

MLaaS

Clone made for < $350 with 98.6% accuracy of original

Model cloning

ML system

Training data

ML model

Input

Answer

Data collection + labelling takes time & effort

Non-problem domain data is cheap

Clone model

Labeling through queries

Very high similarity in terms of accuracy

Data collection + labelling takes time & effort
02. Boundary placement
Adversarial Examples

“Optical Illusions” for Machines

Image by artist Joseph Jastrow, published in 1899 in Popular Science Monthly
Adversarial examples

ML model

Global noise

x100

feature y

Dog

Cat

feature x

Cats

ML model
Sharif, Bhagavatula, Bauer, Reiter: Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition. In ACM SIGSAC 2016

03.
Changing boundaries
Incremental learning | Anomaly detection in practice

Training data

ML Model

Re-training/incremental learning
Data poisoning

Hmmm, more data → better model
Data poisoning in anomaly detection

Training data

Incremental learning/re-training

ML Model
04. Feature selection
What does an ML Model learn?

- size
- # teeth
- # tennis balls

ML Model

- feature y
- Dogs
- Cats

- feature x
- tail length
- pupils
- amount of yarn
Interpretability $\rightarrow$ Explainability

ML training algorithm learns features automatically *without* knowing what they represent

**The Good**

Montavon, Lapuschkin, Binder, Samek, Müller. "Explaining nonlinear classification decisions with deep taylor decomposition." *Pattern Recognition* 2017


**The Bad**

Detecting and Removing Bias


05. Probability maximization
ML Model

Dogs

Cats

Dogs

Cats
Fredrikson, Jha, Ristenpart. "Model inversion attacks that exploit confidence information and basic countermeasures." In CCS 2015.

Figure 1: An image recovered using a new model inversion attack (left) and a training set image of the victim (right). The attacker is given only the person's name and access to a facial recognition system that returns a class confidence score.
Privacy - Use case: Smart Grid
Forecasting power consumption

- **Suppliers** need forecast to buy energy generation contracts that cover their clients
- **Distribution supply operators** require longer term forecasts to ensure the necessary network capacity is available
- Forecasting could allow for dynamic price determination

Privacy Concerns in the Smart-Grid

Energy consumption reveals

Patterns
- Another microwave meal?

Invalid usage
- Insurance or warranty

Real-time information
- Number of people in a household
- Are you on holidays?

Computing on Encrypted Data

1. Enter medical data
2. Encrypt values with personal password
3. Send encrypted values to Azure
4. Cloud runs prediction algorithm on encrypted data
5. Cloud returns encrypted prediction
6. Decrypt prediction with personal password

**Machine Learning using Encrypted Data**

- Forecast power consumption for next half hour in $\approx 2.5$ seconds to evaluate
- Simple Neural network
  - Inputs: 51
  - Hidden layers: $3 \ (8 \rightarrow 4 \rightarrow 2)$
  - Output: 1

**Funded by the Horizon 2020 Framework Programme of the European Union**

- Bos, Castryck, Iliashenko, Vercauteren: *Privacy-friendly Forecasting for the Smart Grid using Homomorphic Encryption and the Group Method of Data Handling*. AFRICACRYPT 2017
Summary

Model Cloning
• How to protect IP sensitive trained model from extraction / cloning?

Adversarial Examples
• Safety & Security impact

Data Poisoning
• Incremental learning is often essential for deployment → How to detect, prevent or harden?

- Large-scale deployment + acceptance by the masses needs explainability → detect and prevent bias
- How to enable privacy-enhancing technologies?
  ✓ Crypto to the rescue: FHE, MPC, …
Conclusions

• Machine learning can improve quality of life due to availability of huge amounts of data
• Security is one of the biggest challenges in large scale deployment of machine learning
• A lot of open security, trust & privacy challenges
• On top of these all ‘classical’ attacks remain
  - Platform security is non-trivial
• Expect zero-day attacks against interesting valuable machine learning models
• Very active field → cat and mouse game
SECURE CONNECTIONS FOR A SMARTER WORLD

&MACHINE LEARNING