CAE IN CYBERSECURITY COMMUNITY

Machine learning applications in cybersecurity: From development to deployment

Dr. Holly Yuan

University of Wisconsin-Stout

Menomonie, WI

yuanh@uwstout.edu

Introduction

- Program Director, Computer Networking & Cybersecurity
- Professor
- Cybersecurity Research Center Director
- Certifications: CCIE Enterprise, CCNP, CISM, AWS, VMware, CEH, Linux
- Teaching Focus: Cybersecurity, Networking, Capstone courses
- Research Interests: Cloud Security, Zero Trust, Workforce Development in Cybersecurity

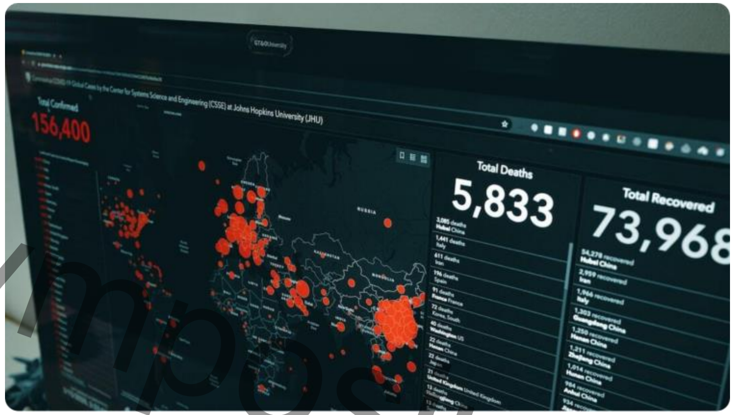
How familiar are you with AI Applications in Cybersecurity?

Please share your response! ☺

A. Very Familiar

B. Somewhat Familiar

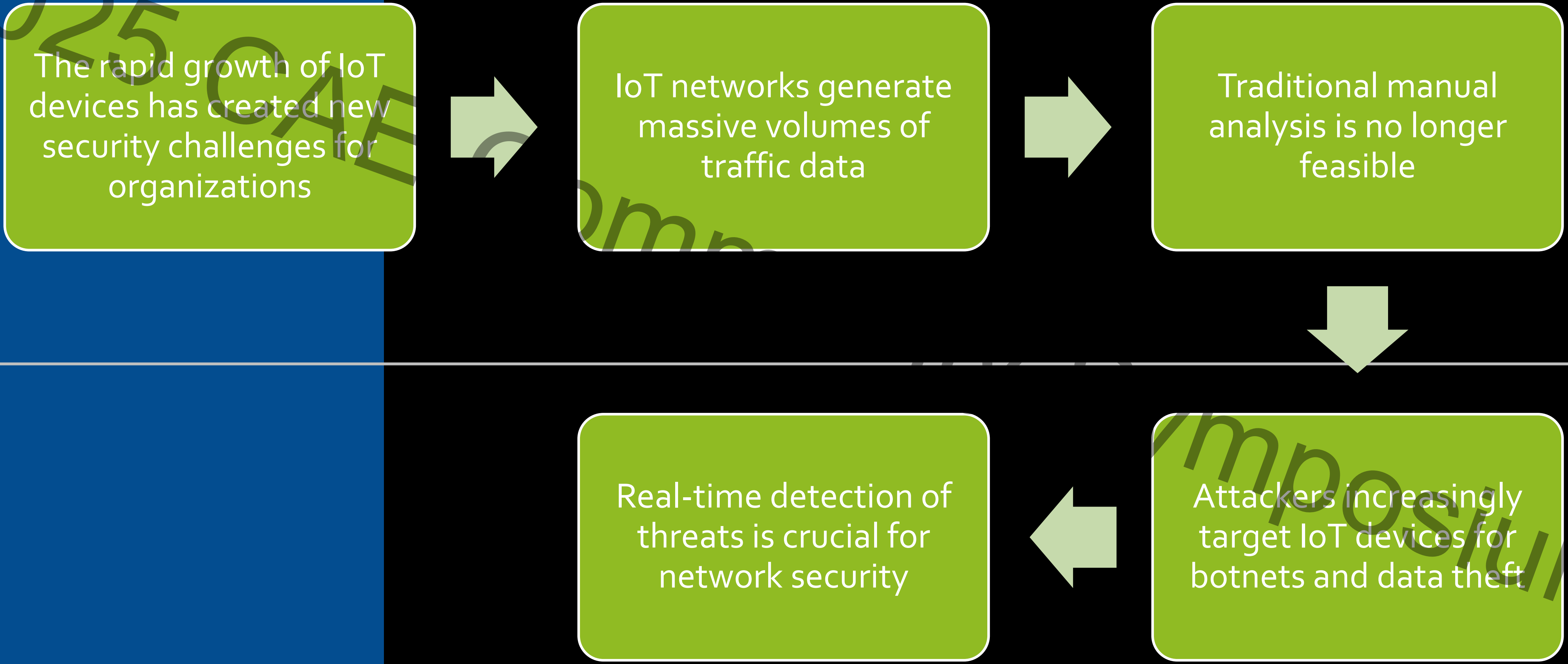
C. Not Familiar At All



Complex Threat Landscape

- **Data Volume:** Massive data sets challenge real-time analysis.
- **Attack Speed:** Rapid execution leaves minimal reaction time.
- **Complex Patterns:** Evolving threats bypass traditional detection methods.

AI in Action: Business Context



The rapid growth of IoT devices has created new security challenges for organizations

IoT networks generate massive volumes of traffic data

Traditional manual analysis is no longer feasible

Attackers increasingly target IoT devices for botnets and data theft

Real-time detection of threats is crucial for network security

Our Goal: Use a dataset (RT-IoT 2022) to train a machine learning model to automatically detect malicious network traffic in real-time IoT environments.

Dataset Description

- **The RT-IoT2022 Dataset is a collection of real-world network traffic data from IoT environments.:**
- Size: 123,117 network traffic samples
- Features: 85 network traffic characteristics
- Types: Mix of normal and attack traffic patterns
- Source: Real-world IoT infrastructure data
- Attack types: DDoS, ARP poisoning, and malware
- **Key Features Include:**
- Protocol types
- Service types
- Flow duration
- Packet statistics
- Network behavior patterns

Prepare Data for Training

separate features (X) from Target (y), splits data: 80% for training, 20% for testing

Train the Model

Implemented supervised ML.

Evaluate the Model

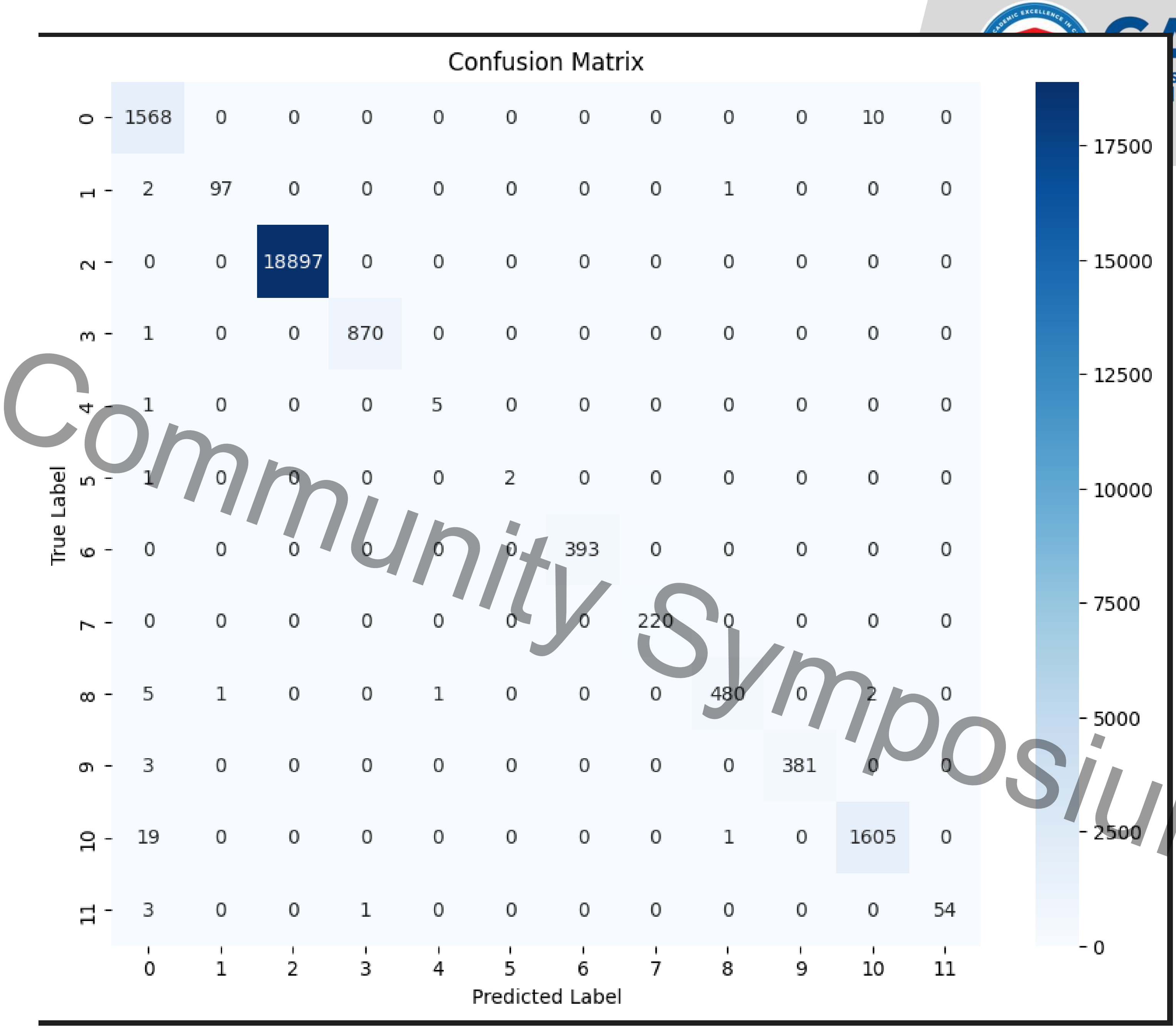
Makes predictions on test data, show accuracy of overall correct predictions

Random Forest Classifier

- Works by creating multiple decision trees (forest).
- Each tree makes a prediction (votes)
- The final prediction is the majority vote from all tree

Evaluating Effectiveness

99.7% Accuracy
in identifying anomalies.



Confusion Matrix

True Label \ Predicted Label	0	1	2	3	4	5	6	7	8	9	10	11
0	1568	0	0	0	0	0	0	0	0	0	10	0
1	2	97	0	0	0	0	0	0	1	0	0	0
2	0	0	18897	0	0	0	0	0	0	0	0	0
3	1	0	0	870	0	0	0	0	0	0	0	0
4	1	0	0	0	5	0	0	0	0	0	0	0
5	1	0	0	0	0	2	0	0	0	0	0	0
6	0	0	0	0	0	0	393	0	0	0	0	0
7	0	0	0	0	0	0	0	220	0	0	0	0
8	5	1	0	0	1	0	0	0	480	0	2	0
9	3	0	0	0	0	0	0	0	0	381	0	0
10	19	0	0	0	0	0	0	0	1	0	1605	0
11	3	0	0	1	0	0	0	0	0	0	0	54