Machine Learning for Information Assurance in SCADA Systems

Steve H. Kim, Peter Bayiokos, Constanza Cabrera-Mendoza, Sabrin Kaur Guron, Wildenslo Osias, Charles C. Tappert and Avery Leider

Seidenberg School of Computer Science and Information Systems
Pace University

What is a SCADA System?

- ➔ Supervisory Control and Data Acquisition Systems
- ➔ Field: Industry Setting
  - ◆ i.e. power plants, manufacturing and assembly lines, chemical plants, water supply networks
- ➔ Use: Remotely control industrial machines

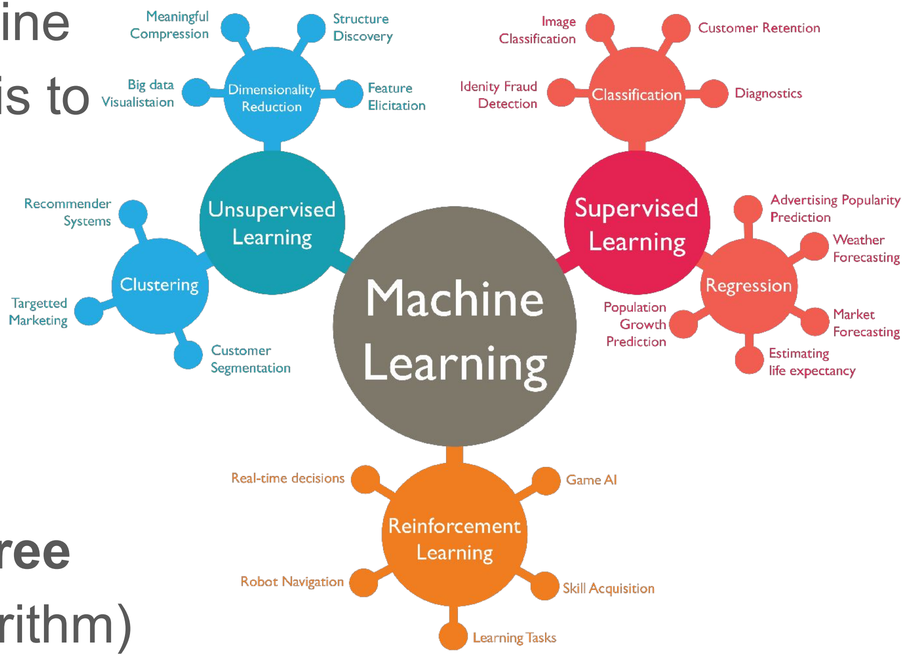
What is the DANGER?: Technologically dependent systems have high levels of risk if a threat finds a vulnerability. This has a tremendous impact on industries infrastructure and clients.

Machine Learning: A Quick Review

ML a variety of methods to teach a machine to recognize specific patterns and use this to identify and classify data

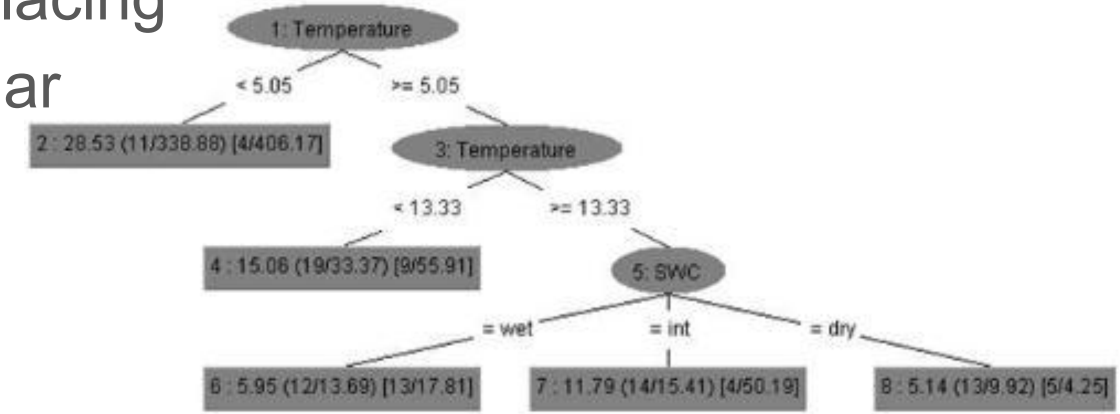
➔ Unsupervised Learning
➔ Supervised Learning
➔ Reinforcement Learning

We selected: **Reduced Error Pruning Tree** (Classified as a supervised learning algorithm)



Reduced Error Pruning Tree

- Decision making tree algorithm
- Based off of C4.5, data mining algorithm for data classifying.
- Reduces errors by replacing nodes with most popular classes



MITRE ATT&CK for Industrial Control Systems Framework

- Knowledge base to determine actions taken by attackers
- Describes tactics, techniques, software, and groups
- Use framework to gain clarity on attack type, severity level, and remediation steps.

Literature Review Performed

- ➔ Security Issues in SCADA Networks, (2006)
- ➔ Sustainable Security for Infrastructure SCADA, (2003)
- ➔ Guide to Industrial Control Systems (ICS) Security, (2011)
- ➔ A Survey of Approaches Combining Safety and Security for Industrial Control Systems, (2015)

Our Research

- ➔ We will be using a data processing tool, Weka, to classify our data set.
- ➔ The machine learning algorithm being used to classify will be the REPTree to better make decisions of the type of attack.
- ➔ Once the ML algorithm determines the attack based on our data points, we move into the framework.
- ➔ The ATT&CK for ICS framework will assist in determining the type of attack tactic as well as the severity level.
- ➔ We will combine all of this to create an automated alerting system for SCADA and Industrial Control systems.