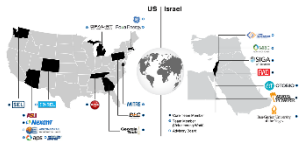


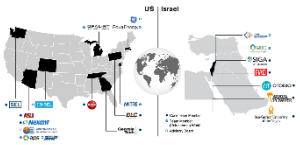
Enhancing Cybersecurity of Grid Operations

Lalitha Sankar
Associate Professor
Arizona State University

28 September 2022

ARIZONA STATE UNIVERSITY

Enhancing Cybersecurity of Grid Operations

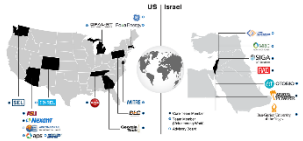


Task 5: Generate event-mimicking attacks ✓

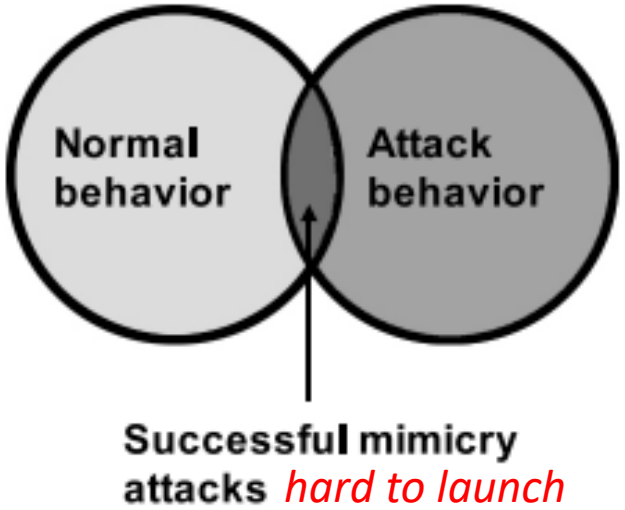
Task 8: Detect event-mimicking attacks

Commercialization: Evaluate attacks on Nexant's (Resource Innovations) EMS Platform ✓

Task 5: Event-mimicking Attacks and Countermeasures



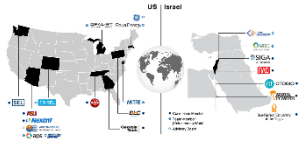
- Modern grid with renewables is more stochastic in operations and requires real-time monitoring to detect/identify real events (oscillations/outages) and attacks.
- ML-based detectors can be easily evaded by attacks that mimic events, ultimately, causing significant damage on grid operations.



mimicry attack: a careful cyberattack on data that throws off ML detector

Source: https://towardsdatascience.com/evasion-attacks-on-machine-learning-or-adversarial-examples-12f2283e06a1

Task 5: Mimicking Attacks in IT Systems



A practical mimicry attack against powerful system-call monitors

Authors: Chetan Parampalli, R. Sekar, Rob Johnson Authors Info & Claims

ASIACCS '08: Proceedings of the 2008 ACM symposium on Information, computer and communications security • March

Mimicry attacks on host-based intrusion detection systems

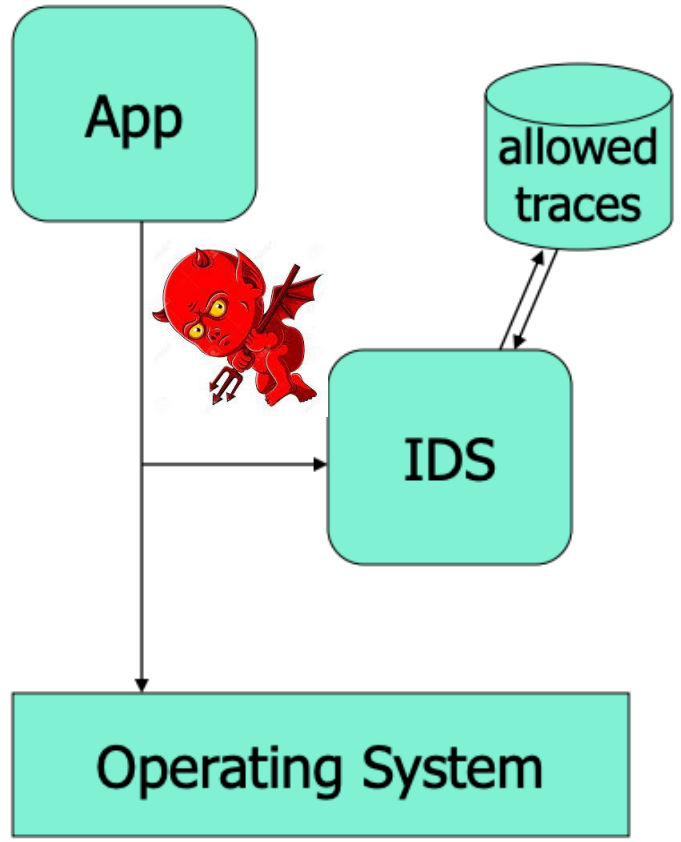
Authors: David Wagner, Paolo Soto Authors Info & Claims

CCS '02: Proceedings of the 9th ACM conference on Computer and communications security • November 2002 • Pages 255–264 • https://doi-org.ezproxy1.lib.asu.edu/10.1145/586110.586145

IEEE TRANSACTIONS ON COMPUTERS, VOL. 64, NO. 1, JANUARY 2015 139

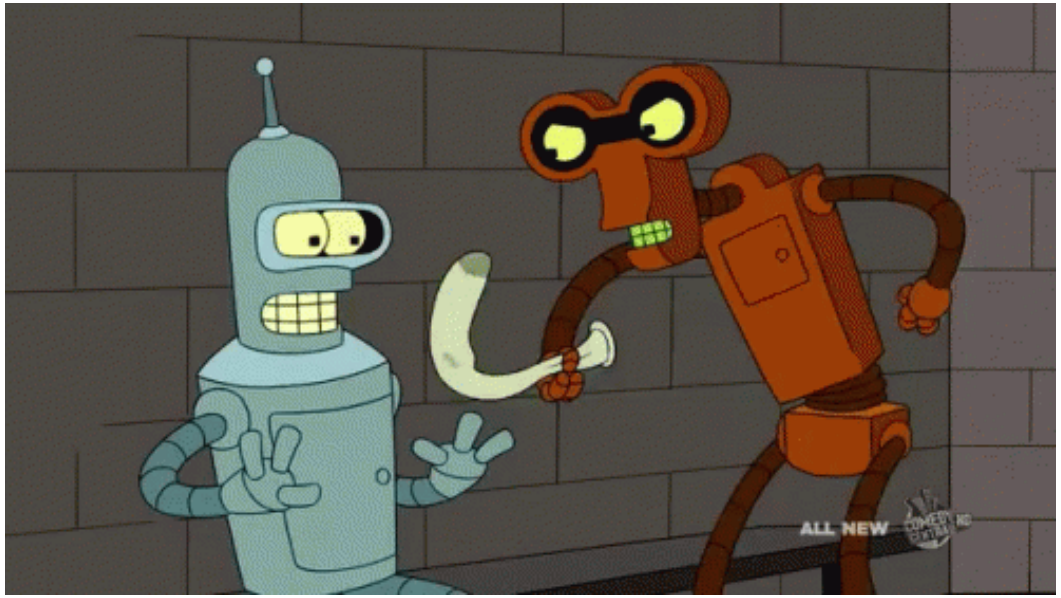
Fool Me If You Can: Mimicking Attacks and Anti-Attacks in Cyberspace

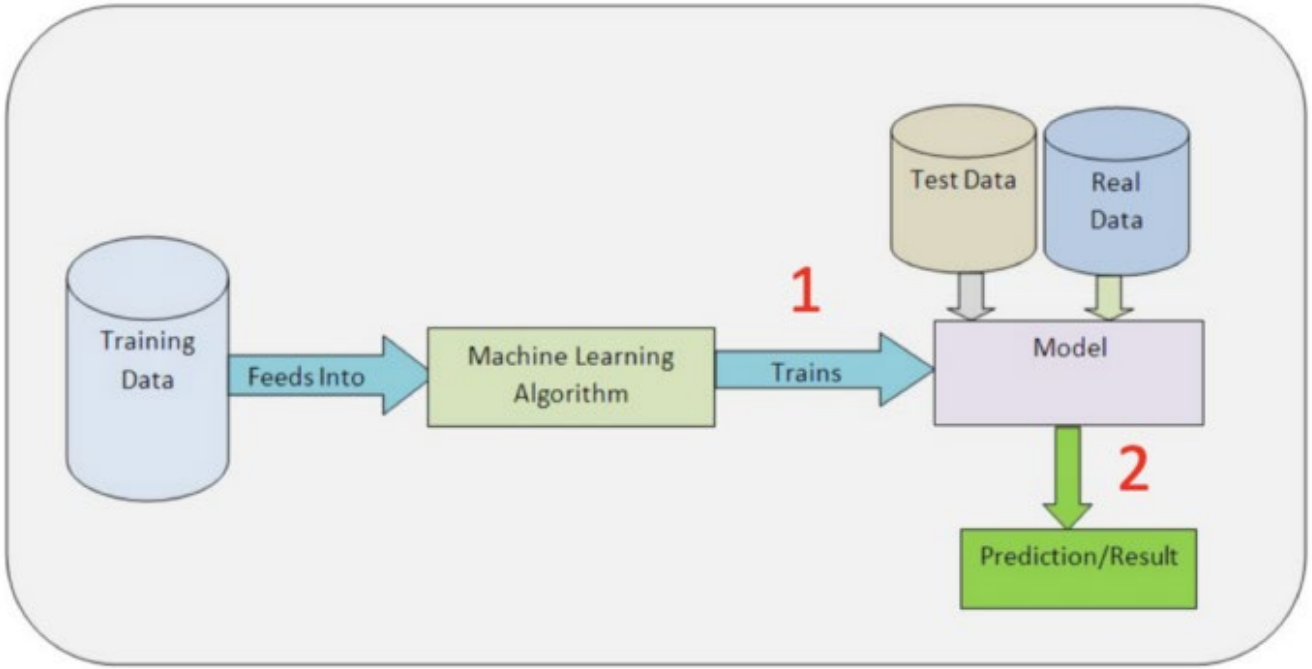
Shui Yu, *Senior Member, IEEE*, Song Guo, *Senior Member, IEEE*, and Ivan Stojmenovic, *Fellow, IEEE*



attacks target software internal to a computer

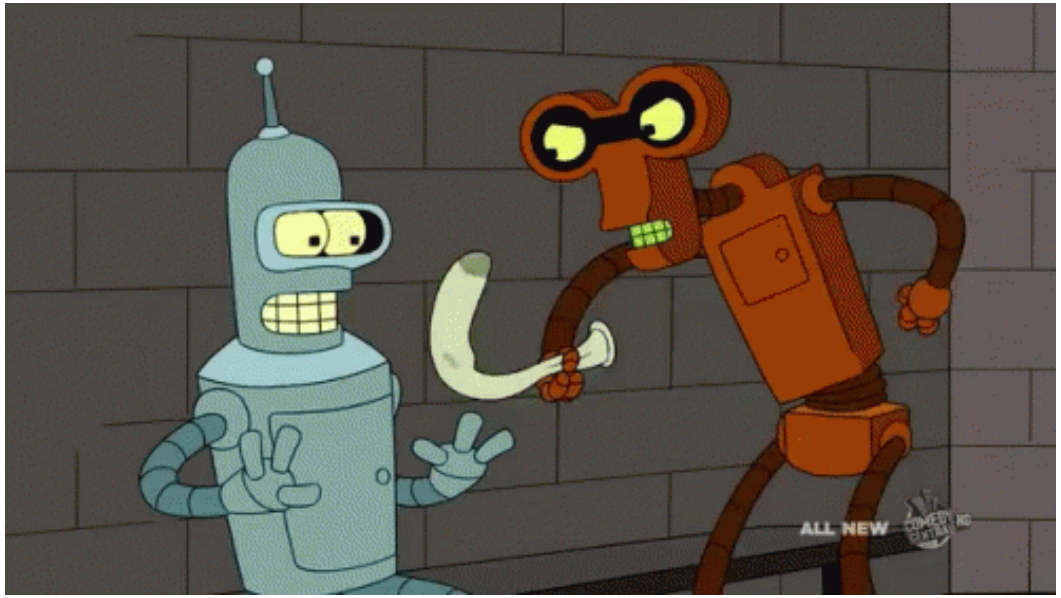
Where can Attackers target in OT Systems?

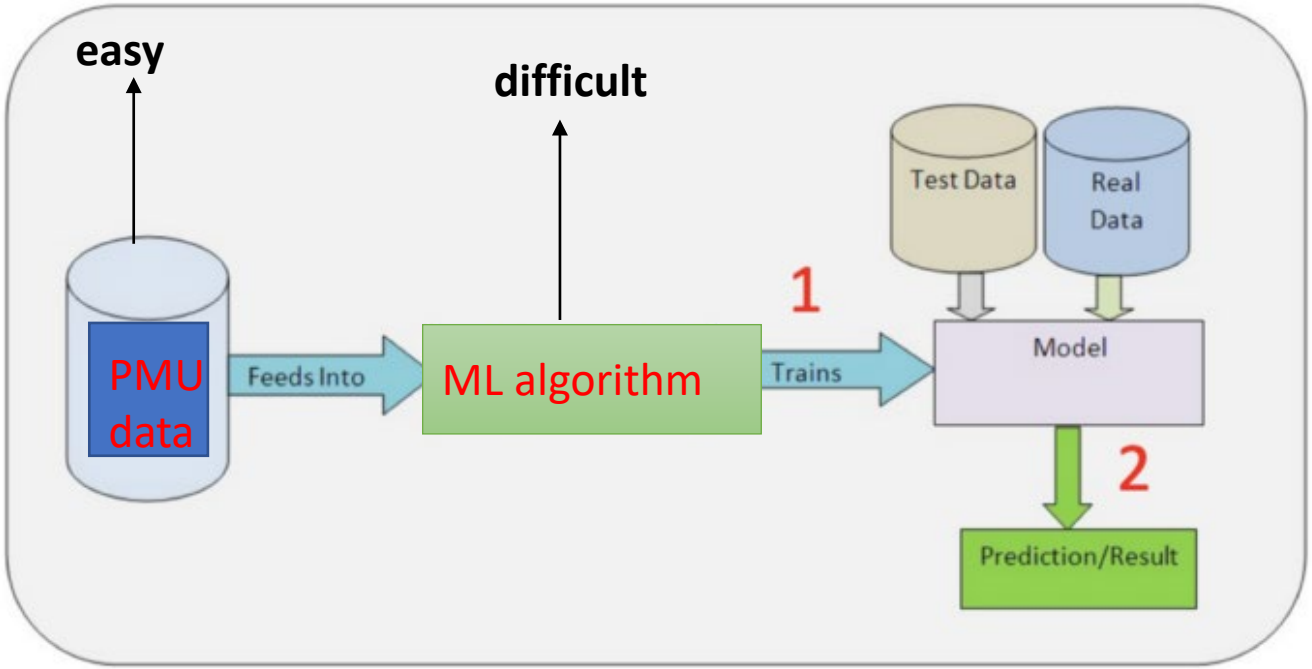




Source: https://towardsdatascience.com/will-my-machine-learning-be-attacked-6295707625d8

Where can Attackers target in OT Systems?





….attackers are like electricity: they chose the path of least resistance…..

Data is a potentially feasible pathway for attacks

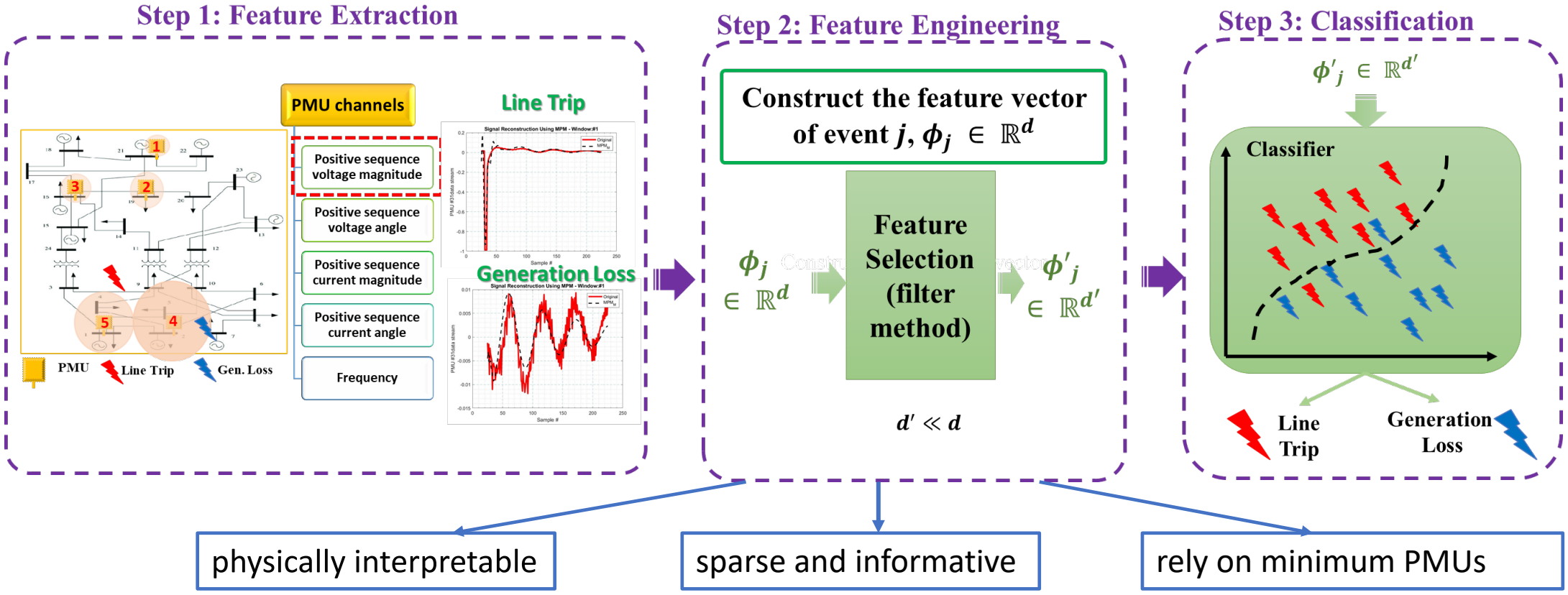
But for mimicking event attacks, need to explore:

- *how to tamper data?*
- *how many PMUs to tamper?*
- *how long to tamper?*

Source: https://towardsdatascience.com/will-my-machine-learning-be-attacked-6295707625d8

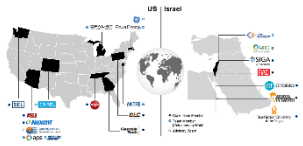
Task 5 (a): Learn Event Signatures from Measurements

- A typical ML-based attack detector maps "event signatures" into "feature space"
- Features are later used to classify events (e.g., line trip or generation loss)

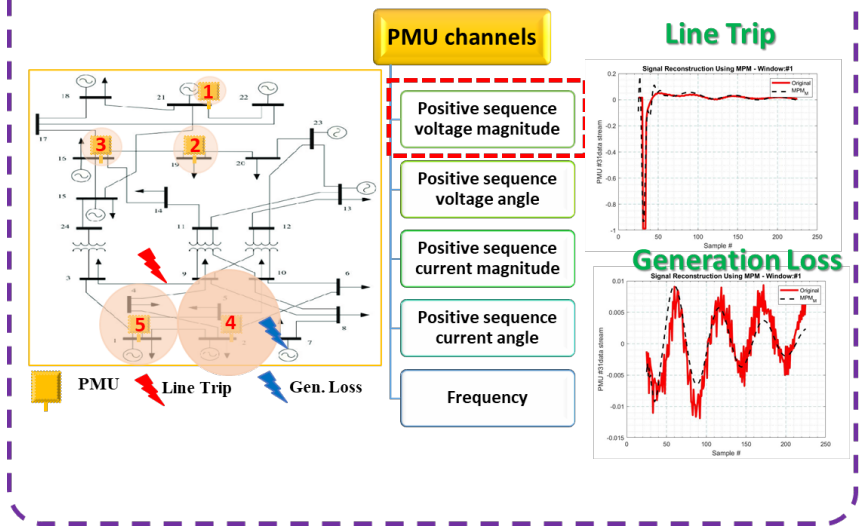


[3] N. Tahipourbazargani et.al (2022) A Machine learning framework for event identification via modal analysis of PMU data, under review, IEEE PES.

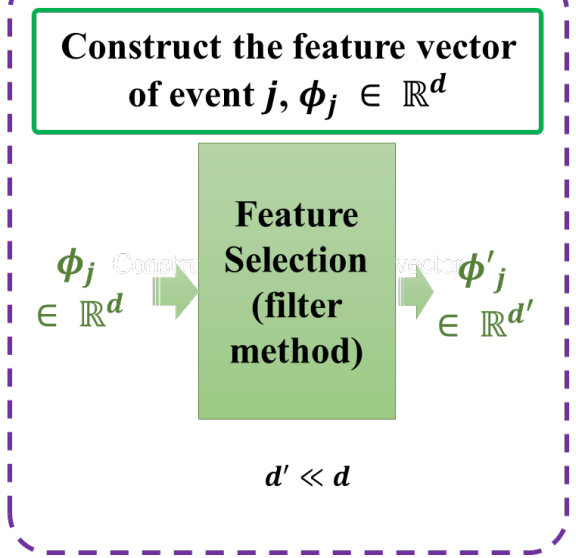
Task 5 (a): Learn Event Signatures from Measurements



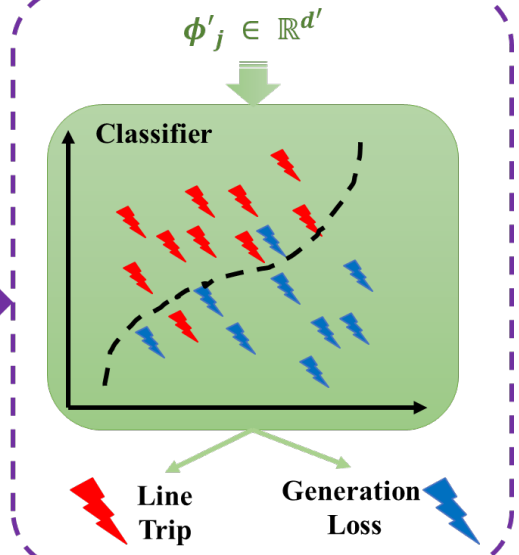
Step 1: Feature Extraction

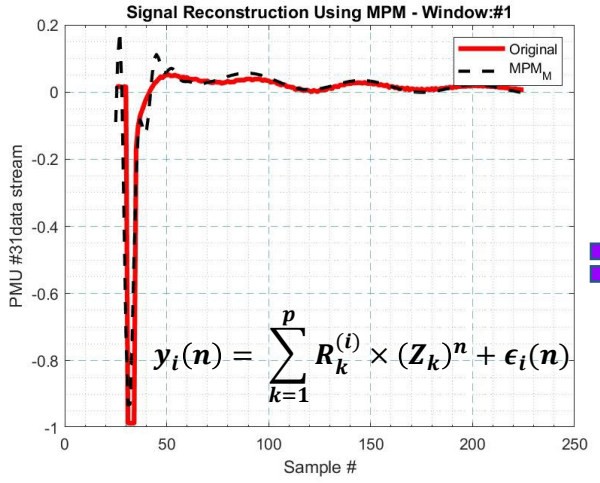


Step 2: Feature Engineering

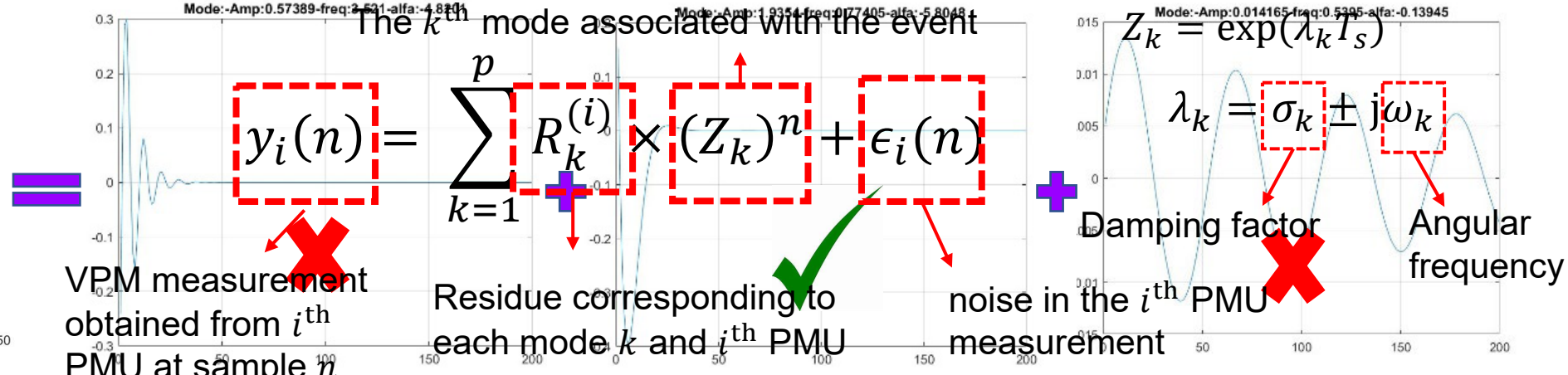


Step 3: Classification

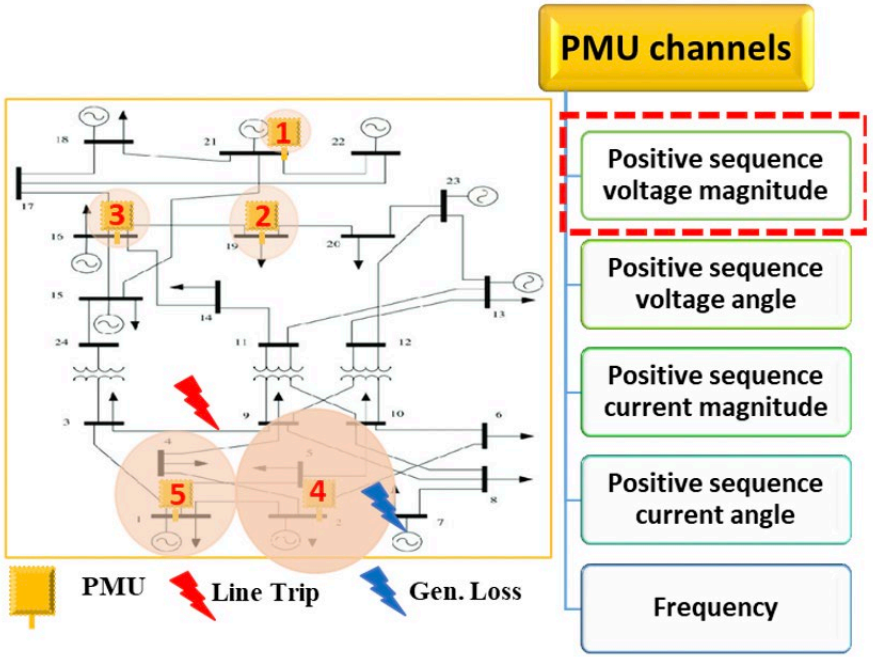


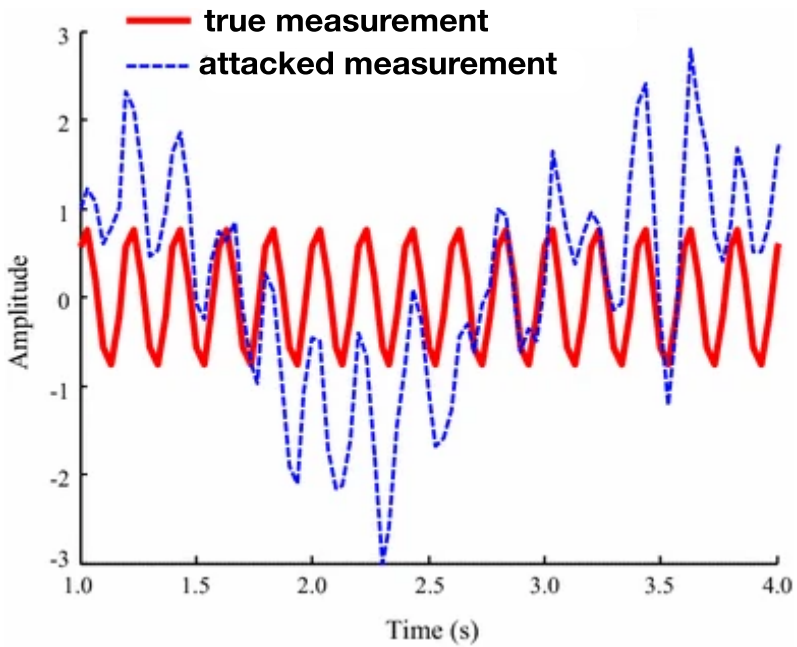


Not all features are meaningful or useful – identify best features that capture event



Task 5 (a): Learn Event Signatures from Measurements



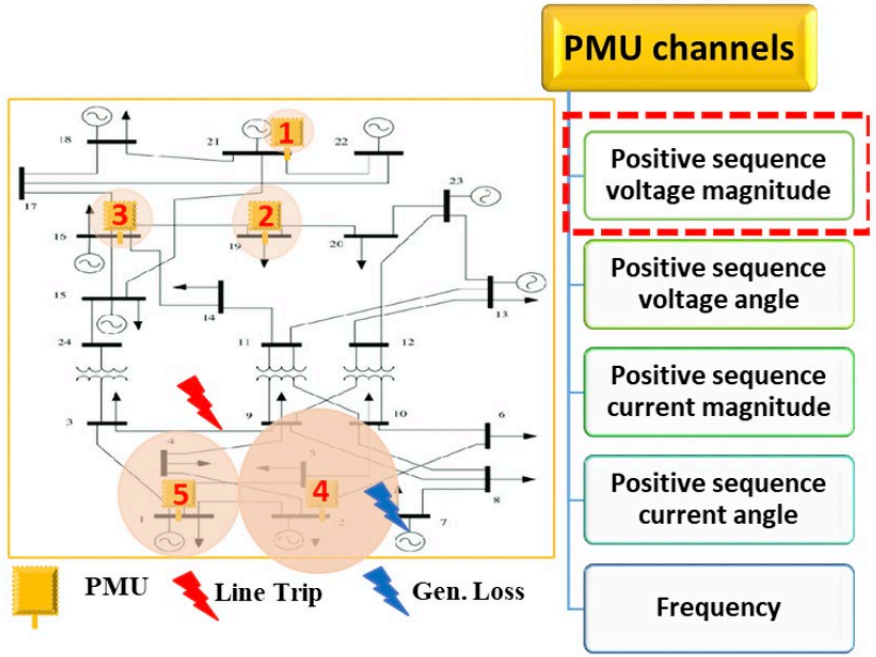


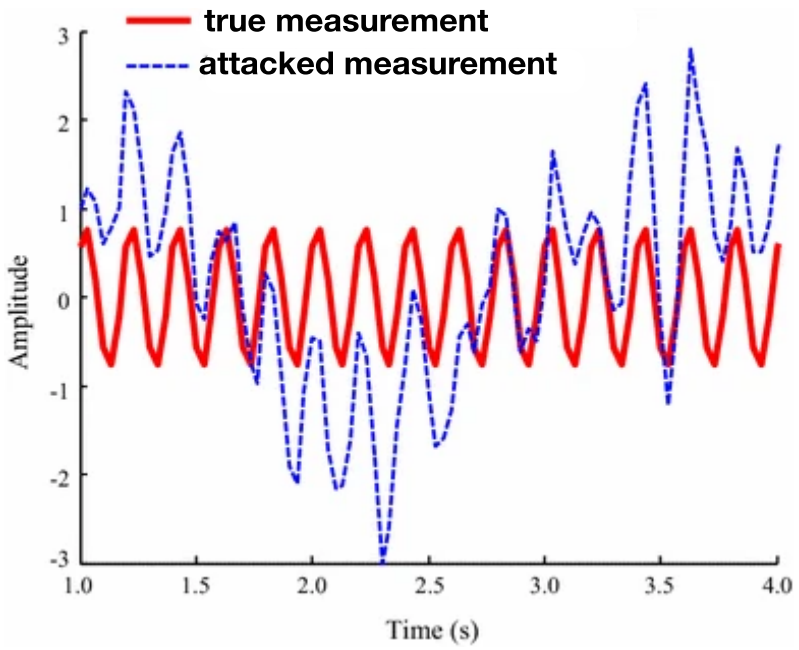
Can we identify physically realizable attacks

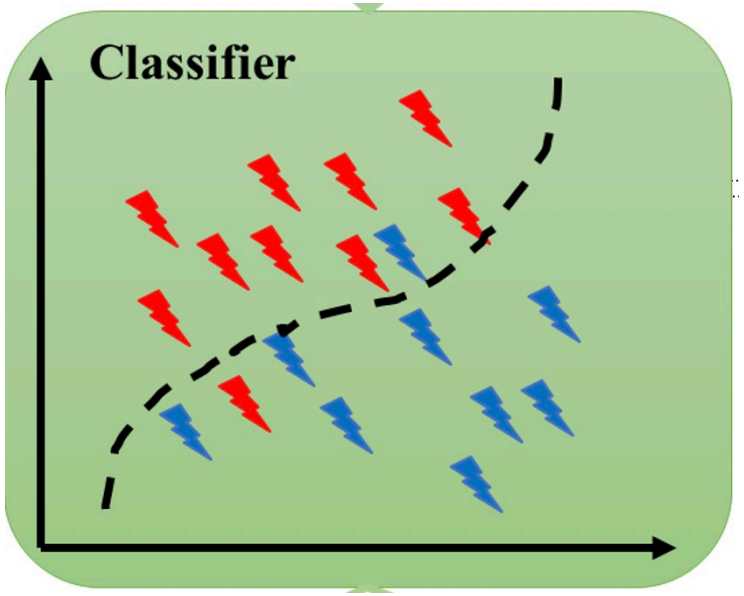
(e.g., event-mimicking) ?

Yes! By identifying key event features that are easy to synthesize by changing measurements!

Task 5 (a): Learn Event Signatures from Measurements

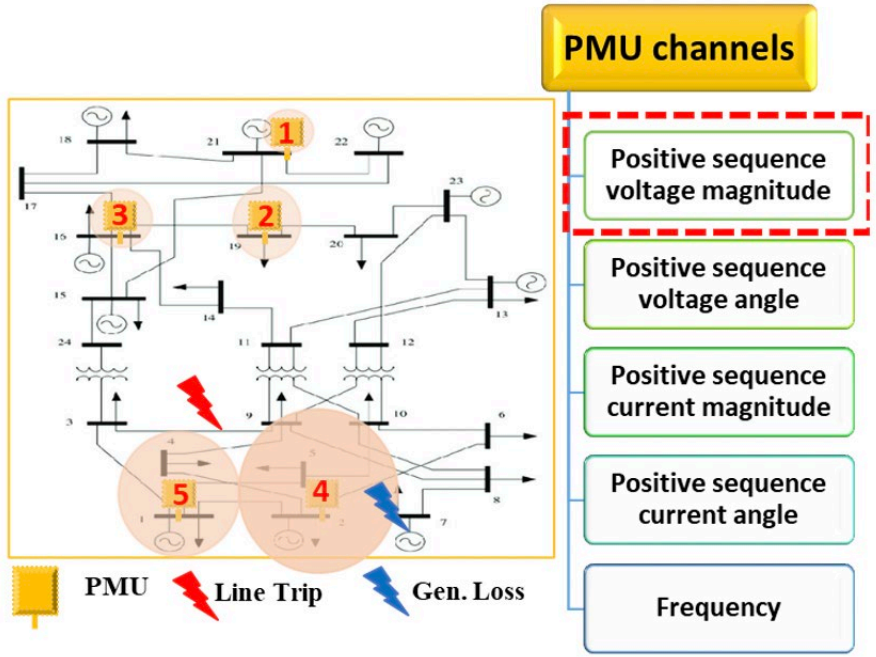


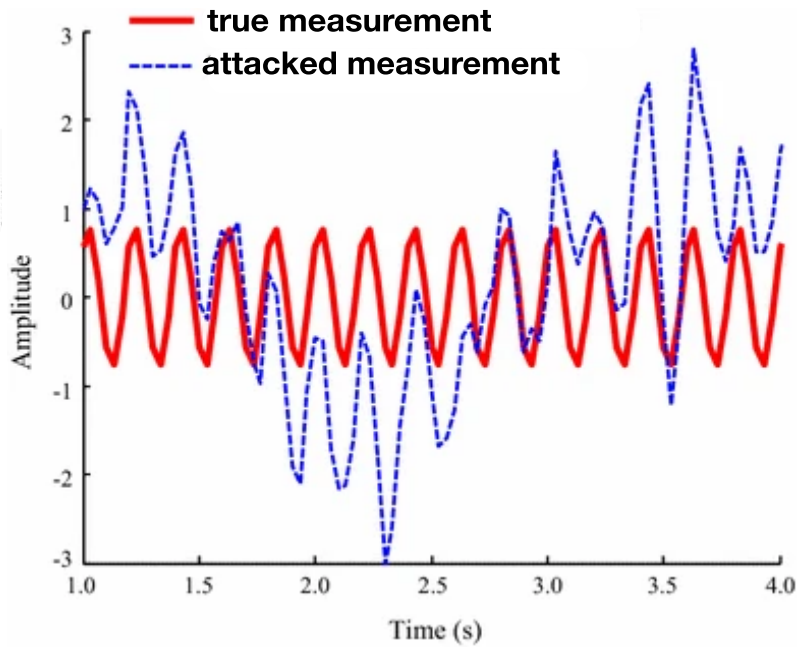


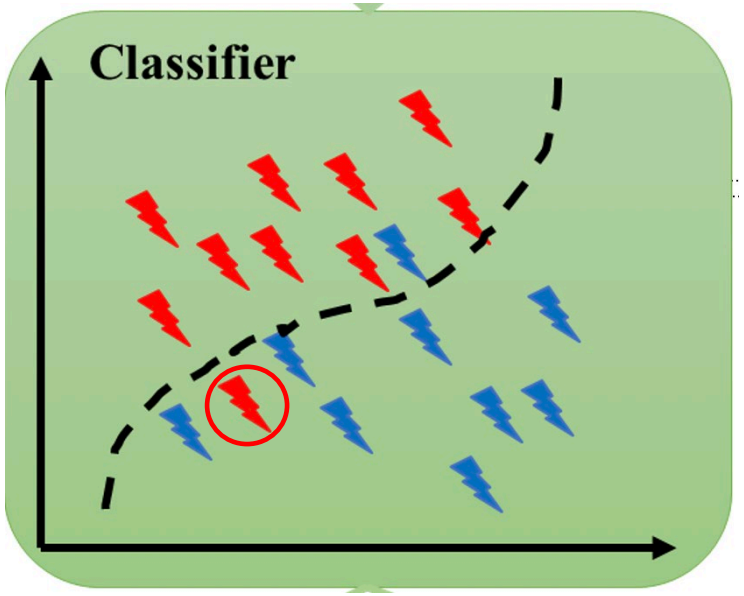


Yes! By identifying key event features that are easy to synthesize by changing measurements!

Task 5 (a): Learn Event Signatures from Measurements







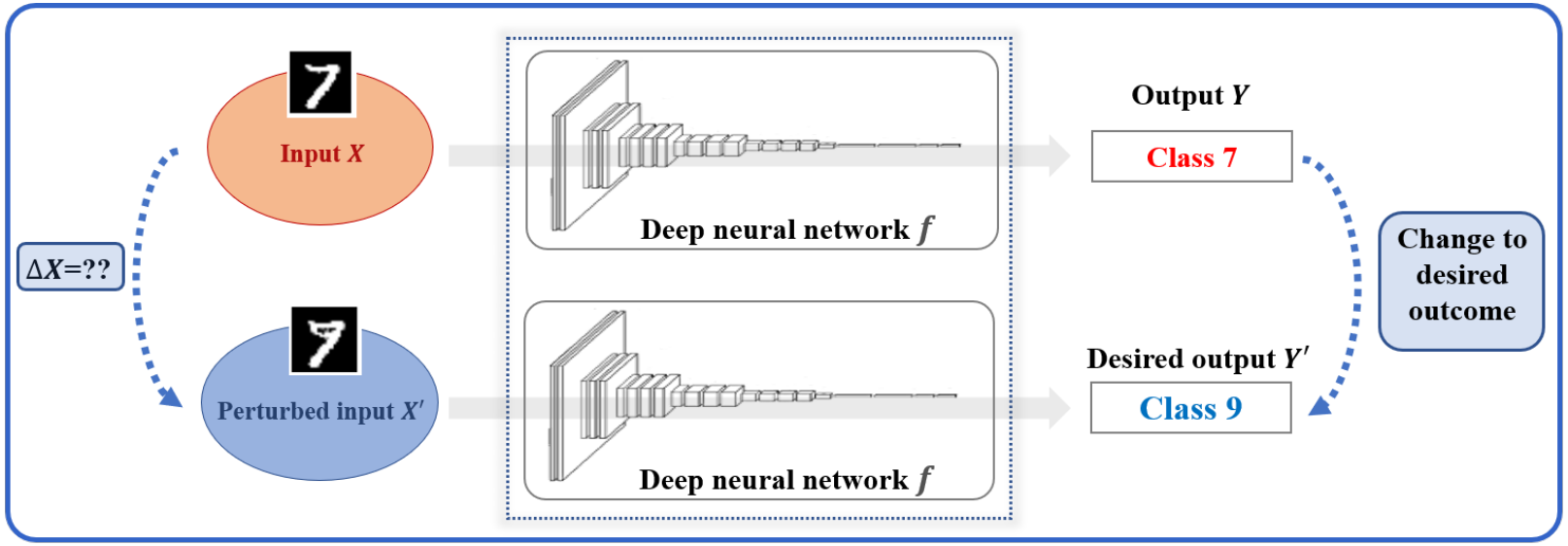
Challenge: Adding white noise or some arbitrary mode is not sufficient

Work in progress:

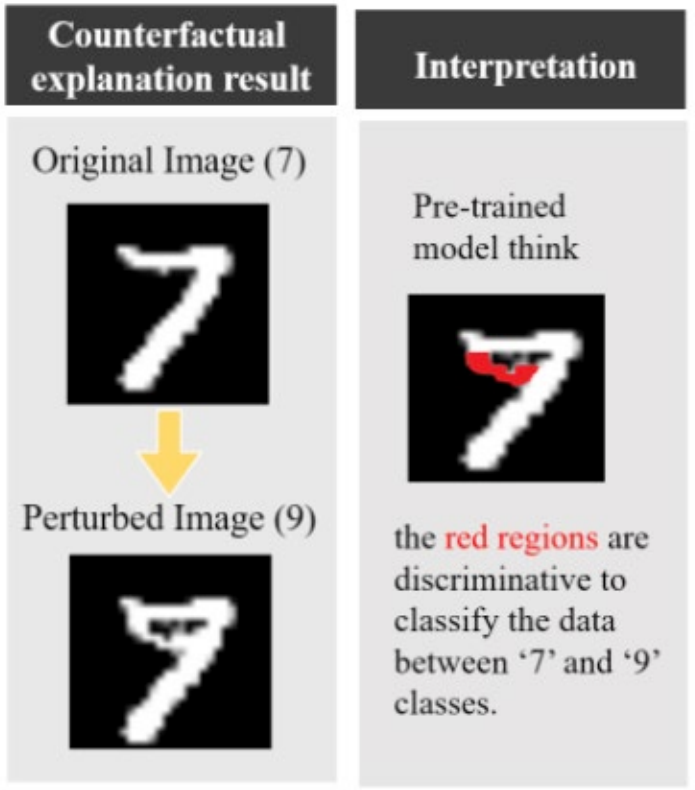
- Extend existing binary classifier to multi-class classifier to include attacks
- Identify the key set of features that can change normative data to mimic an event
- Integrate new synthesized attacks to the existing database

Task 5 (b): Interpretable Models for Attack Generation

- Counterfactual machine learning models:



Framework of counterfactual explanation*



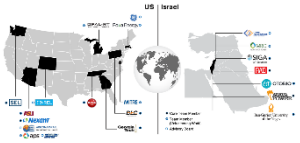
Counterfactual models for attacks on power system attacks:

- Determine *minimal set of features with large attack impact*
- Features should be *realizable by perturbing measurements*

*[Online] Available: https://da2so.github.io/2020-09-14-Counterfactual_Explanation_Based_on_Gradual_Construction_for_Deep_Networks/

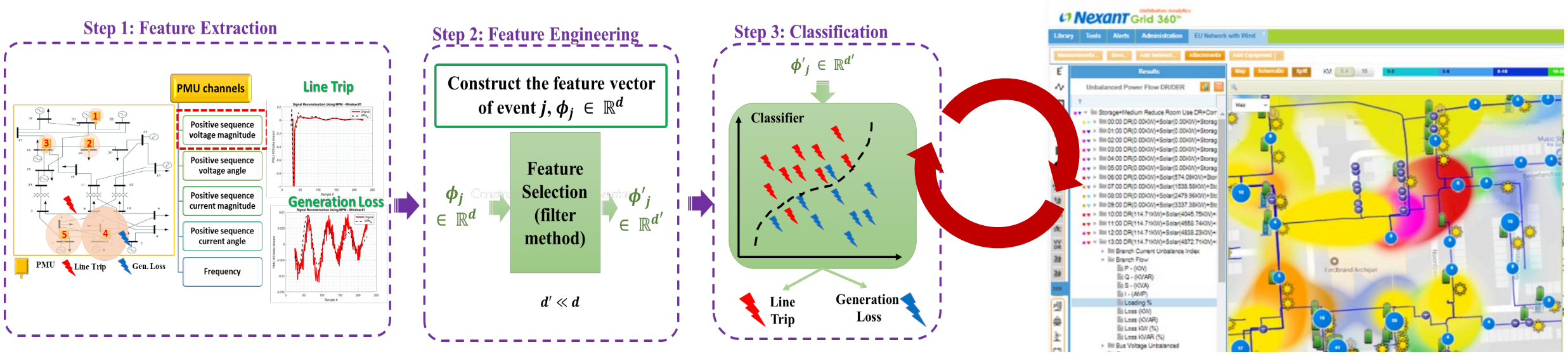
[2] A. Pinceti, O. Kosut and L. Sankar, "Data-Driven Generation of Synthetic Load Datasets Preserving Spatio-Temporal Features," *PESGM*-2019, pp. 1-5,

Summary of Task 5



Task 5	Status (Work in progress)	Work to be done
• Task 5(a): mimic attacks by tampering data • Task 5(b): interpretable models for attack generation	• Extend Binary to multi-class classifier • Evaluate the ML detector performance for attacks realized by adding noise.	• Fully automated mimicking attacks using data alone (e.g., GAN based attacks) (Q2) • Integrate new synthesized attacks to the database. (Q10-Q12) • Work with Resource Innovations on Commercialization

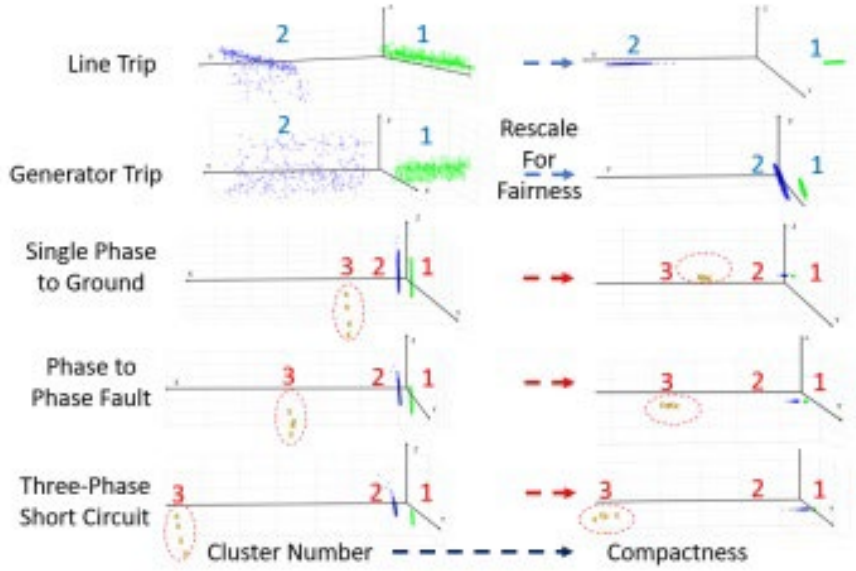
Commercialization Task



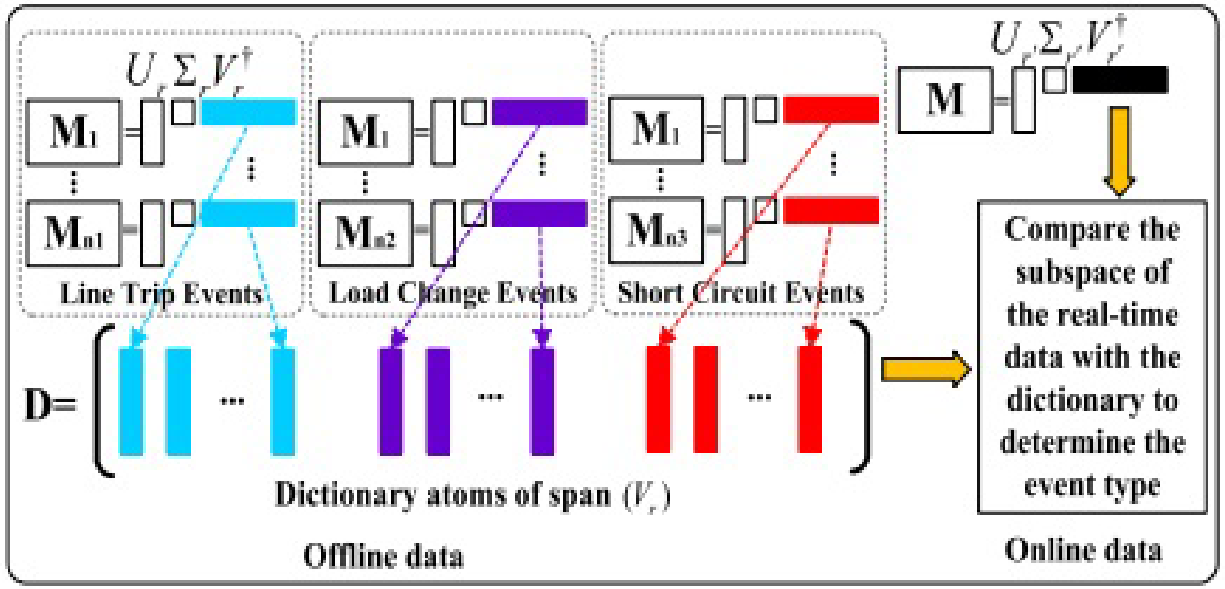
In collaboration with industry partner *Resource Innovations* (John Dirkman)*:*

- Implement end-to-end python package to synthesize mimicking attacks
- Overlay the python package on Nexant Grid 360
- Evaluate attacks for enhanced visualization

Task 5 (a): Learn Event Signatures from Measurements



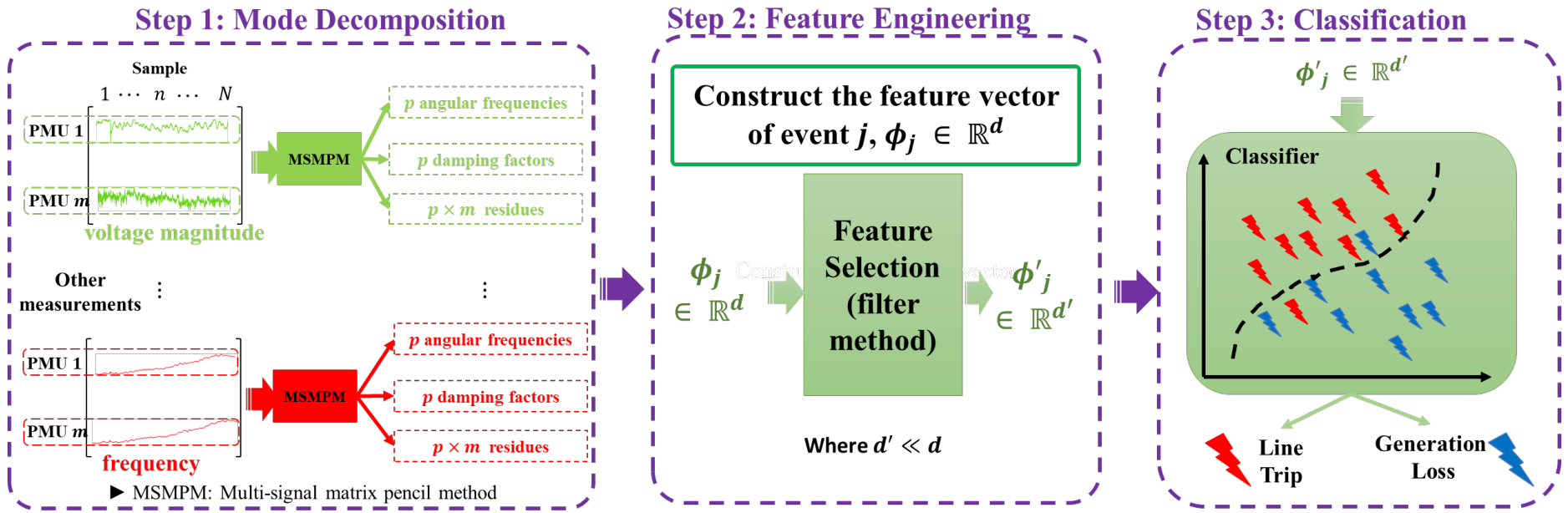
H. Li et. Al (2019), "**An Unsupervised Learning Framework for Event Detection, Type Identification and Localization Using PMUs Without Any Historical Labels**," *PESGM 2019*.

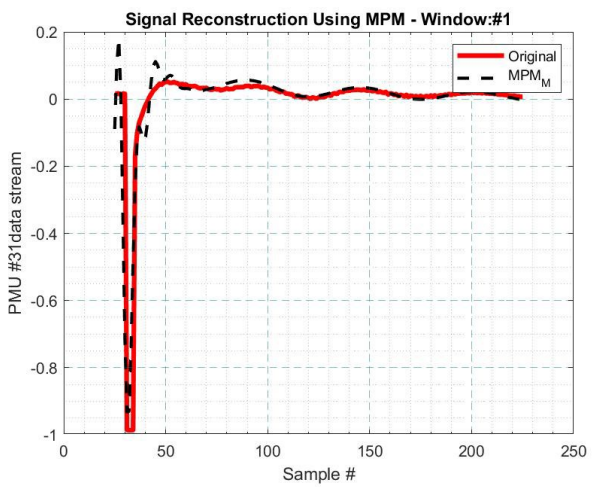


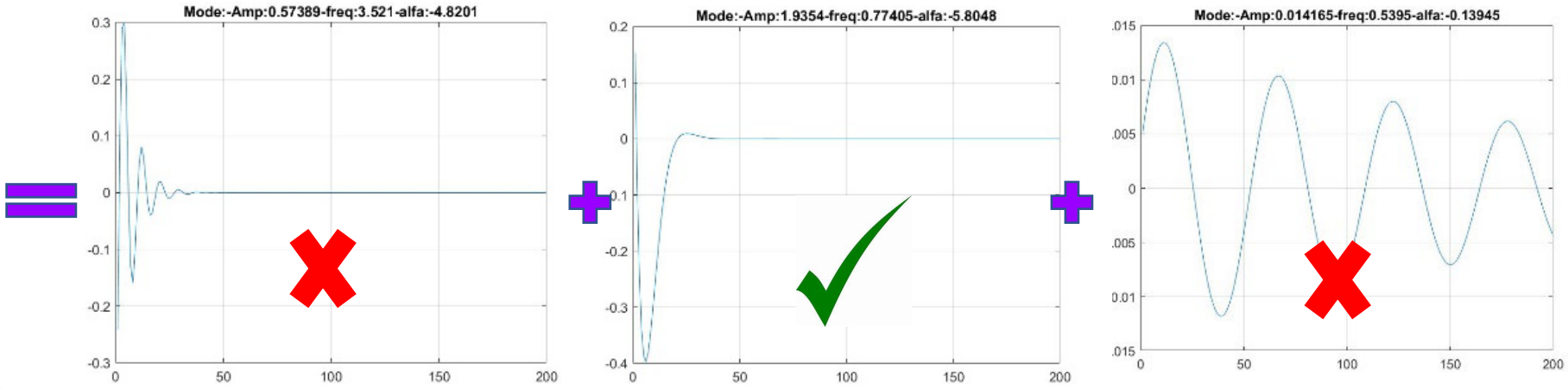
W. Li, M. Wang and J. H. Chow, "**Real-Time Event Identification Through Low-Dimensional Subspace Characterization of High-Dimensional Synchrophasor Data**," *in IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4937-4947, Sept. 2018.

Back up slides

Task 5 (a): Learn Event Signatures from Measurements

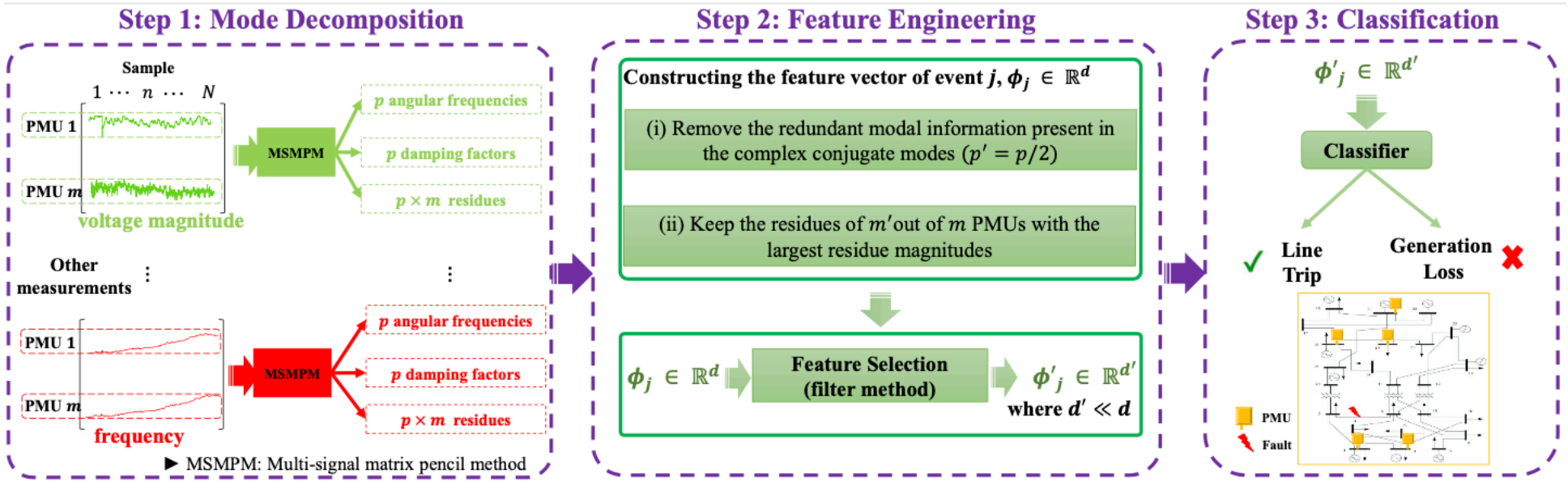






Task 5 (a): Learn Event Signatures from Measurements

- Can we identify physically realizable attacks (e.g., event-mimicking)?

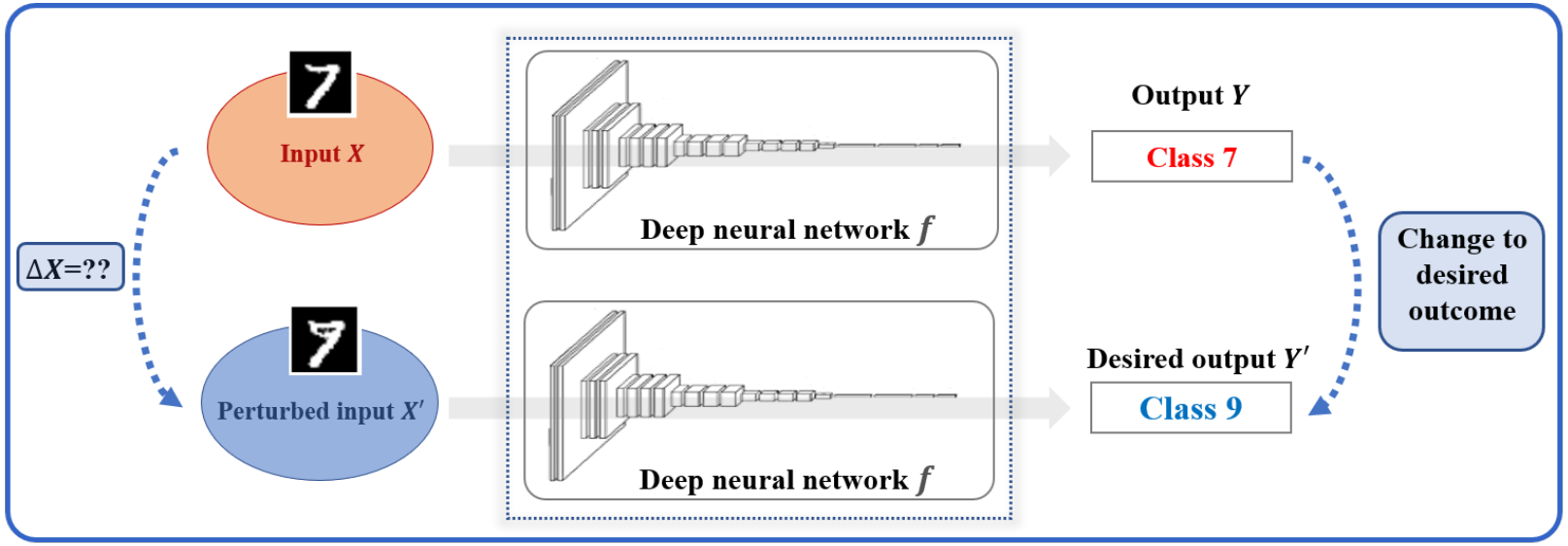


Yes! By identifying key event features that are easy to synthesize by changing measurements!

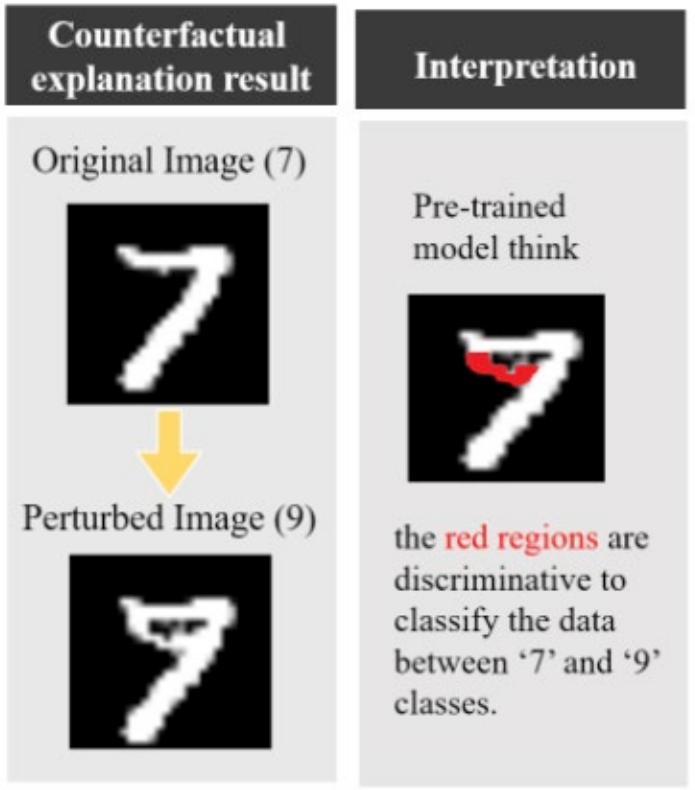
[3] N. Tahipourbazargani et.al (2022) A Machine learning framework for event identification via modal analysis of PMU data, under review, IEEE PES.

Task 5 (b): Interpretable Models for Attack Generation

- Counterfactual machine learning models:



Framework of counterfactual explanation*



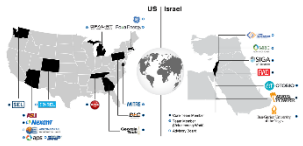
Counterfactual models for attacks on power system attacks:

- Determine *minimal set of features with large attack impact*
- Features should be *realizable by perturbing measurements*

*[Online] Available: https://da2so.github.io/2020-09-14-Counterfactual_Explanation_Based_on_Gradual_Construction_for_Deep_Networks/

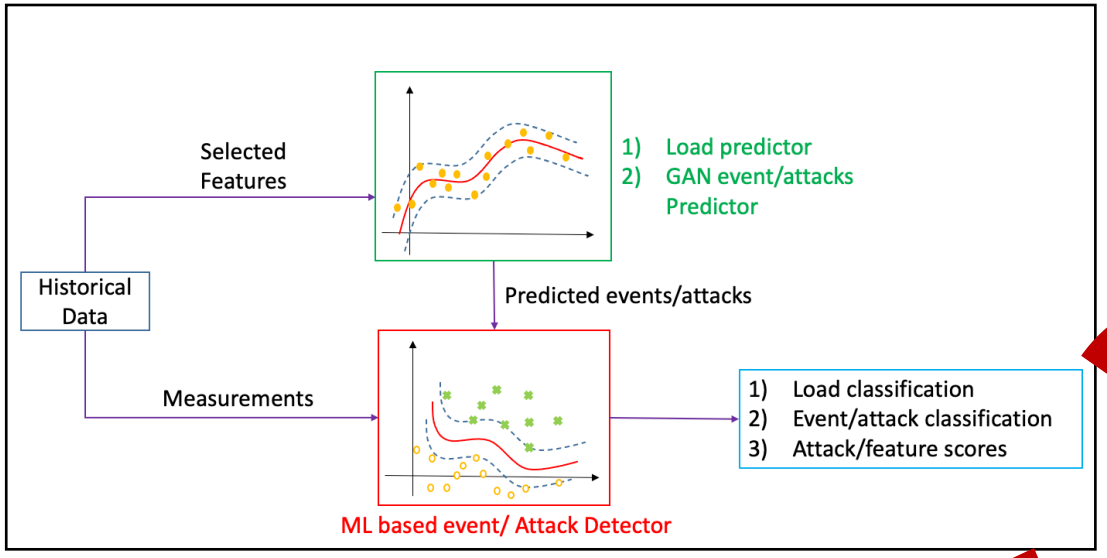
[2] A. Pinceti, O. Kosut and L. Sankar, "Data-Driven Generation of Synthetic Load Datasets Preserving Spatio-Temporal Features," *PESGM*-2019, pp. 1-5,

Ensuring Cybersecurity of Grid Operations



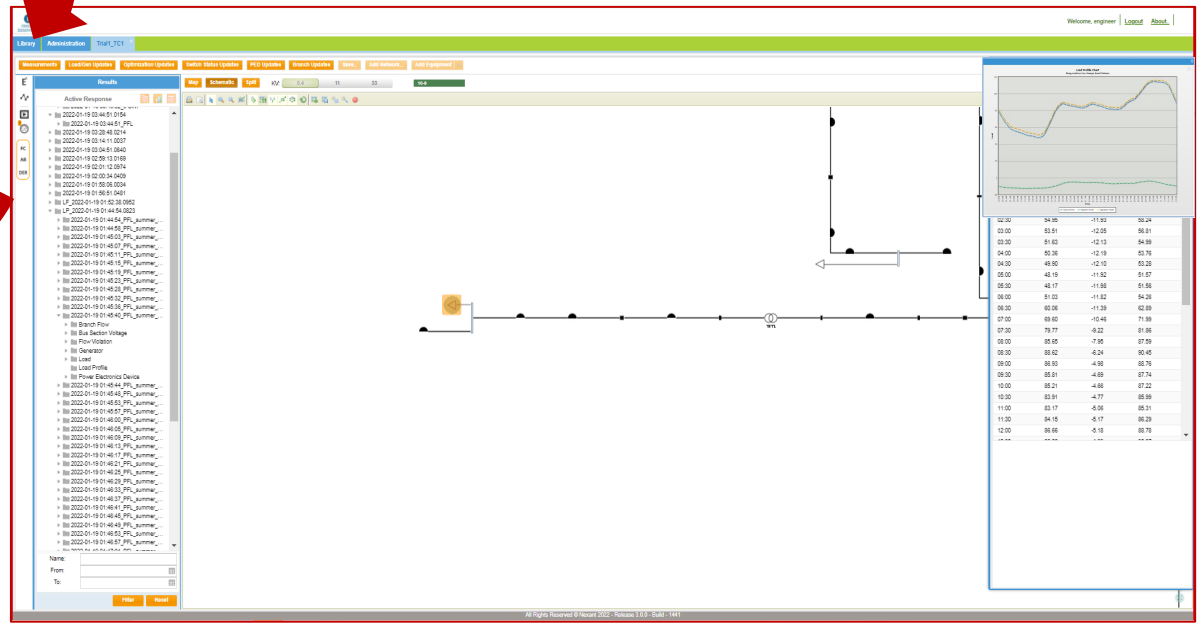
	Details	Status
Task 5 (attack generation)	• synthesize “intelligent” attacks that mimic “events” by tampering measurements.	• completed feature extraction • analyzing features realizable by altering measurements.
Task 8 (attack detection)	• develop ML and data-driven “robust” detectors that detect intelligent attacks.	• In two quarters.
Commercialization	• seamlessly integrate ML detector to Nexant Grid360 tool.	• **pilot study**: test our prior load-altering attacks and detectors using ”smart-meter” data. • **towards product**: in four quarters.

Commercialization – Detection to Anomaly Visualization



Intelligent and interpretable attack/event detector

Resource Innovations: Load Anomaly Visualization



- Things to argue for in terms of attacks:
  - Where can an attack happen?
  - Within the EMS control center?
  - Replay attack at a concentrator/aggregator?
  - We know that at least 3 PMUs have to be attacked tohave any effect (reference: Gyorgy Dan, …co-authors) (Nima)
  - Are we changing load data? Or measurements that affect load data?
  - They get direct load measurements (as injections)
    - Attack: how many load measurements should we change and how can it be realistic?
    - Depends on application – where data is coming from. Hope to get this info from John
    - Are there other mechanisms to verify if the load measurements have changed? To ask John

Task 5 (a): Learn Event Signatures from Measurements