Artificial Intelligence Applications in Power Systems: A Brief Tutorial

Drs. Fangxing (Fran) Li, Yilu Liu and Kevin Tomsovic

University of Tennessee

Early Years

Artificial Intelligence

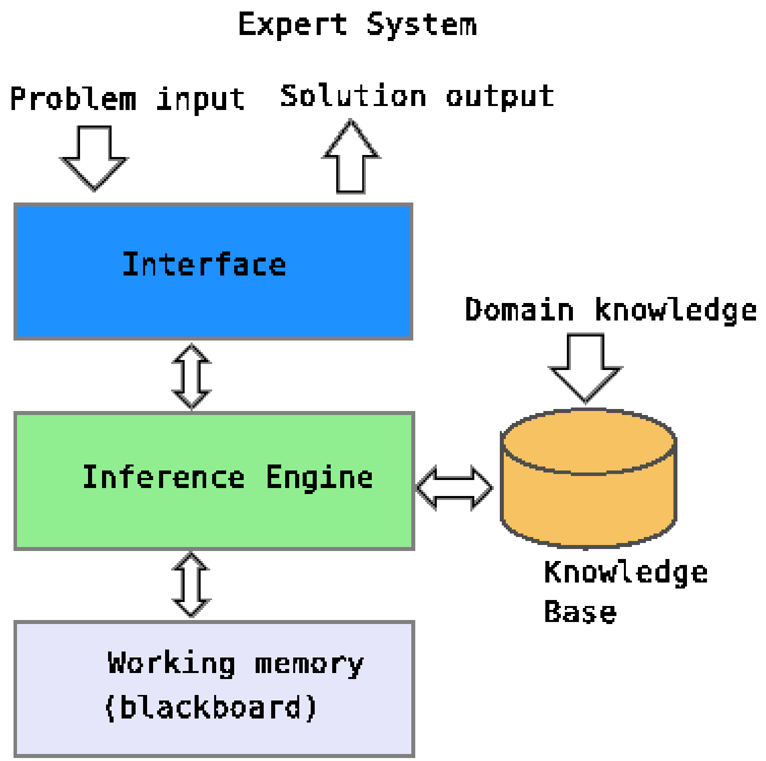
A field "of computer science concerned with designing intelligent computer systems, that is, systems that exhibit characteristics we associate with intelligence in human behavior - understanding language, learning, reasoning, solving problems, and so on. [Feigenbaum, Stanford]"

➔ Much of the development in the 1960s was on more general intelligence. This was found to be extremely difficult and led Feigenbaum to propose more specialized systems.

CURENT

Knowledge-Based Systems

- The general AI approach eventually morphed into the concept of knowledge-based systems and more specifically **Expert Systems**
  - MYCIN – Diagnostic system for blood disease infection
  - SID – Design aid for the VAX 9000 system



➔ These were very specialized systems with a general "engine"

Historical Development in Power Systems

- Earliest applications in nuclear power and security assessment (1970s)
- Numerous projects by mid 1980s
- Example application areas
  - Alarm processing
  - Diagnostics
  - Load forecasting
  - Operations
  - Security assessment

Knowledge Representation

- How to represent knowledge of some specific domain in an

efficient and clear manner?

  - Graphs/trees – example diagnostic trees
  - Rules – typically if-then statements
  - Objects

➔ Emphasis on relationships between data, not data types and not on algorithms

Knowledge Representation
Rules

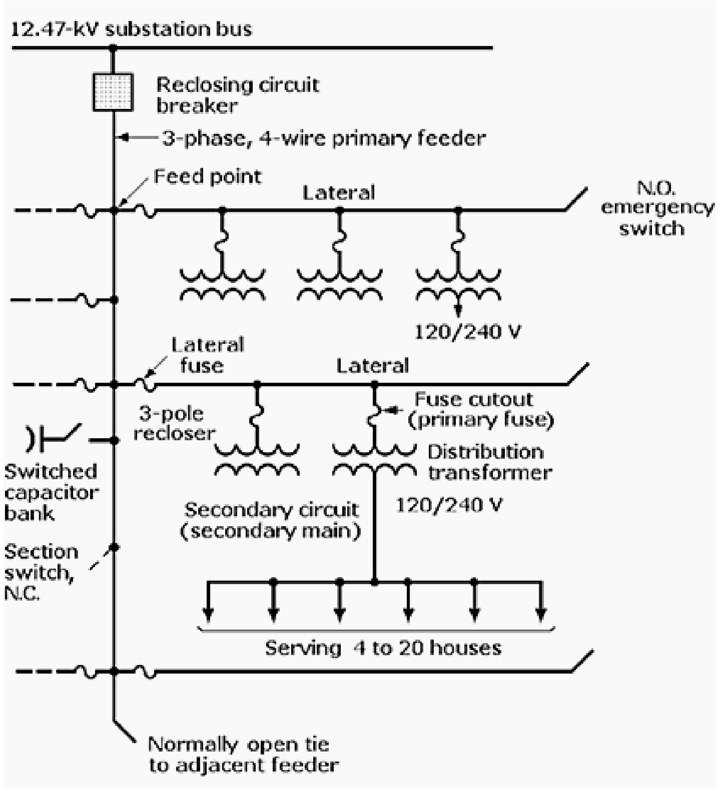
IF-THEN Structures

- Example:

  IF a feeder can be restored from a tie switch

  AND there is sufficient capacity on that feeder

  THEN restore from that tie switch



➔ Raises the question of then how are rules applied

Solution Mechanisms
Inference

Need generic approach to applying rules but must have

- Appropriate context
  - Example: Load priority rules when system overloaded
- Appropriate sequence
  - Example: Restoration after fault isolation

➔ This information must be encoded in the rules so one can use a general logical approach

Solution Mechanism
Rule-chaining – forward chaining (data-driven)

Conditions ➔ conclusions

- Example: Restoration

    Rule A: IF outage THEN search for restoration path

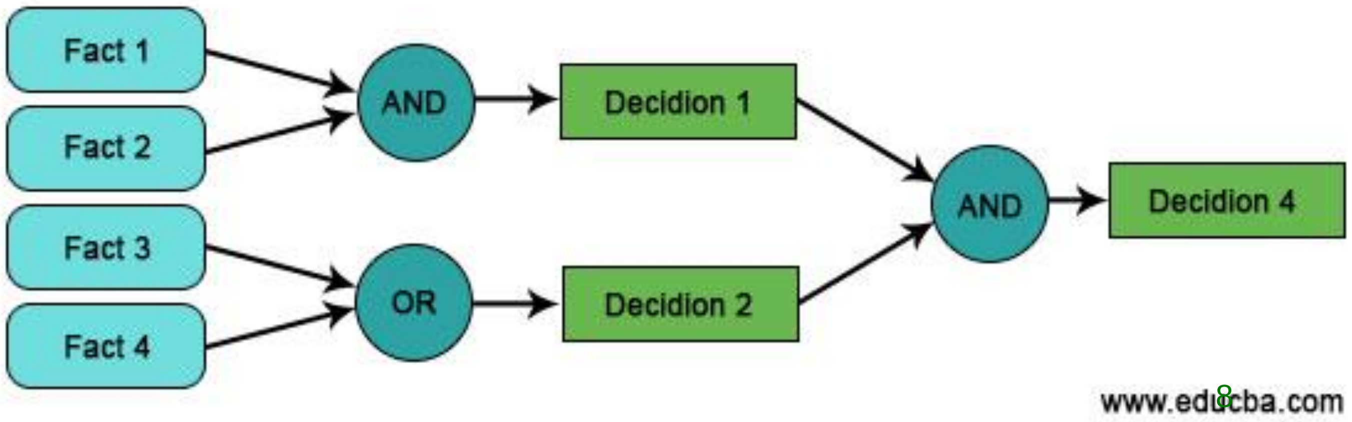
    Rule B: IF searching for restoration path

    THEN look for feeders with excess capacity

- Inference process

    Outage ➔ Apply rule A

    After applying rule A ➔ Apply rule B



CURENT

Solution Mechansims
Rule-chaining – backward-chaining (goal driven)

Conclusions ==> conditions

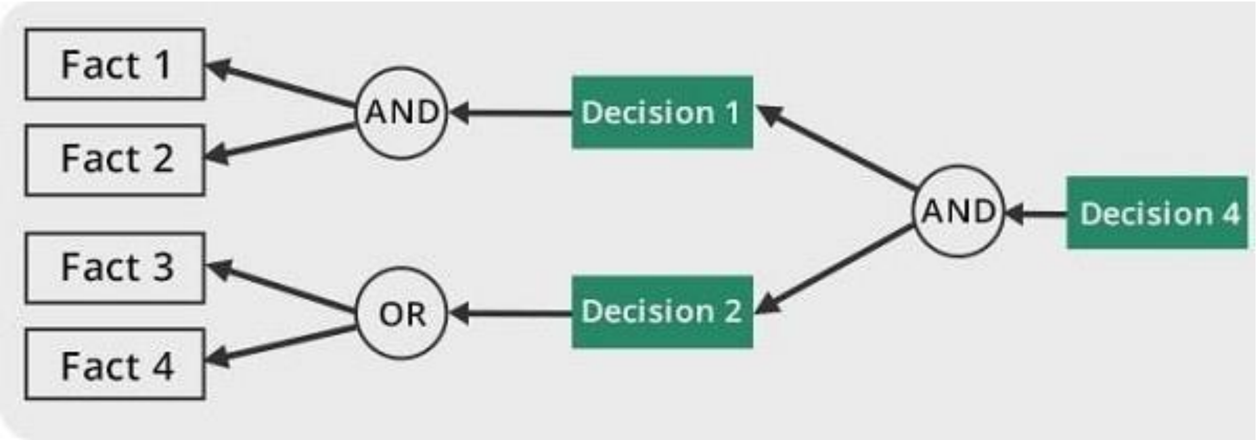
- Example: Transformer diagnostics

  Rule A: IF a thermal fault in transformer

  THEN [H2] will be elevated in oil

- Inference process

  Detect elevated [H2] ➔ thermal fault



CURENT

Solution Mechanisms
Rule-chaining – design criteria

Search from fewer to greater possibilities

- Example: Remedial control action

  IF outage A

  THEN control action C

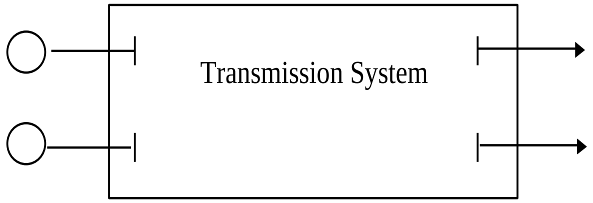
- Assume number of possible outages far fewer than possible control actions then you want to use Forward-chaining

VCES – Voltage Control Expert System

Power flow equation

$$P_i = V_i \sum_{j=1}^{n} V_j \left(B_{ij} \sin(\delta_i - \delta_j) + G_{ij} \cos(\delta_i - \delta_j) \right)$$

$$Q_i = V_i \sum_{j=1}^{n} V_j \left(G_{ij} \sin(\delta_i - \delta_j) - B_{ij} \cos(\delta_i - \delta_j) \right)$$



Voltage dependent loads

$$P_i = P_{io} \overline{V}_i^{a_i} \qquad Q_i = Q_{io} \overline{V}_i^{b_i}$$

Load voltage constraints

$$\overline{V}_i^{\min} \le \overline{V}_i \le \overline{V}_i^{\max}$$

Controls and Constraints

- Shunt capacitors

$$Q_i^{\min} \le Q_i \le Q_i^{\max}$$

- Transformer tap changers

$$t_i^{\min} \le t_i \le t_i^{\max}$$

- Generator voltages

$$\overline{V}_i^{\min} \le \overline{V}_i \le \overline{V}_i^{\max}$$

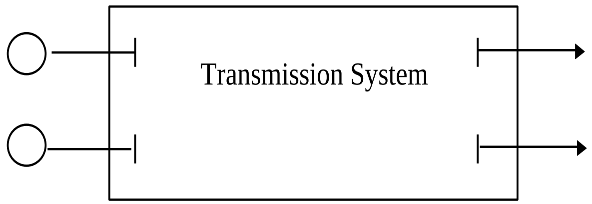
Heuristic (empirical) rules

- It is most efficient to apply VAR injections locally
- Position of local tap changer can be raised (lowered) to correct low (high) voltage
  - May cause other voltages to drop
- Generator bus voltages can be raised (lowered) to solve the low (high) load voltage problems

Justication for a Rule 1

Assume

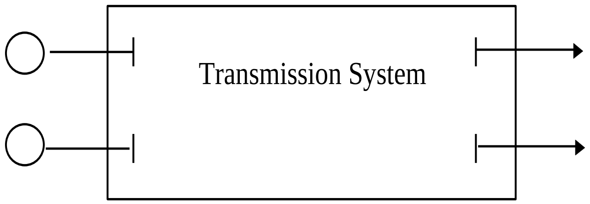
1) Load voltages near 1.0 p.u.
2) Transformer tap settings near 1.0
3) Line angles near 0
4) Lines lossless



If reactive compensation Q is applied at bus i then

$$\Delta\overline{V}_i \geq \Delta\overline{V}_j$$

Justication for a Rule 2



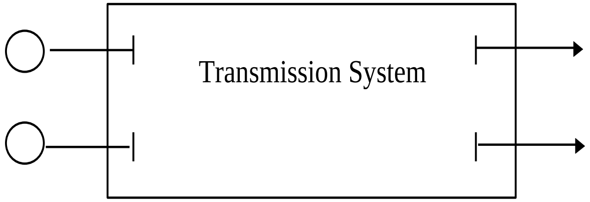
Assume 1) -4) hold

If the tap position t is raised at bus i then

$$\Delta \overline{V}_i \geq 0 \qquad \Delta \overline{V}_j \leq 0$$

Justication for a Rule 3

Assume 1) -4) hold



If the generator voltage is raised at bus *i* then

$$\Delta\overline{V}_j \geq 0$$

Compare to using Optimization (Linear Program)

$$\min_{Q,t,V} C^T \underline{x}$$

such that

$$\underline{x} \geq 0$$

$$A\underline{x} \geq \underline{b}$$

Comparison between LP and VCES

- VCES solves scenarios in a single iteration while LP requires multiple iterations
- For more severe problems, VCES uses fewer controls
- LP has problems with small unrealistic control adjustments and other hard to represent constraints
- VCES approach is generally faster
- VCES tends to provide a better voltage profile
- VCES performance improvement greater with severe problems
- VCES can explain performance

➔ But VCES doesn’t guarantee performance

Uncertainty/Subjectivity in Knowledge

Almost all expert systems have to deal with some more of uncertainty that rarely fits standard probabilistic approaches

- Representations
  - Subjective probability
  - Certainty factors (MYCIN)
  - Fuzzy logic (Zadeh)
    - Membership functions represent use of generic terms, e.g., small, medium and large

CURENT

Limitations of Knowledge-Based Systems

- Incremental improvements (adding new rules) may be difficult.
- Development often slowed by the process of extracting knowledge from human experts.
- Computational efficiency concerns for systems with 1000s of rules. Few expert systems can adapt logic to time constraints.
- It may be difficult to evaluate performance evaluation
- User acceptance of a new technology may be slow, unless coupled with explanation systems

➔ AI research (most) has moved on to learning from data

Terminology

- Artificial intelligence
- Backward-chaining/goal-oriented
- Certainty factors
- Expert systems
- Forward-chaining/data-driven
- Fuzzy sets and logic
- Heuristics
- Inference engine
- Intelligent systems
- Knowledge-based
- Knowledge representation
- Logic programming
- Model-based reasoning
- Objects/frames
- Rule-based
- Subjective probability

Deep Learning Applications in Power Systems

2023 CURENT Industry Conference

April 2023

Fangxing Fran Li, Ph.D., P.E.

James McConnell Professor, The University of Tennessee - Knoxville

Email: fli6@utk.edu

Overview of Artificial Intelligence, Machine Learning, and Deep Learning

Artificial Intelligence & Machine Learning

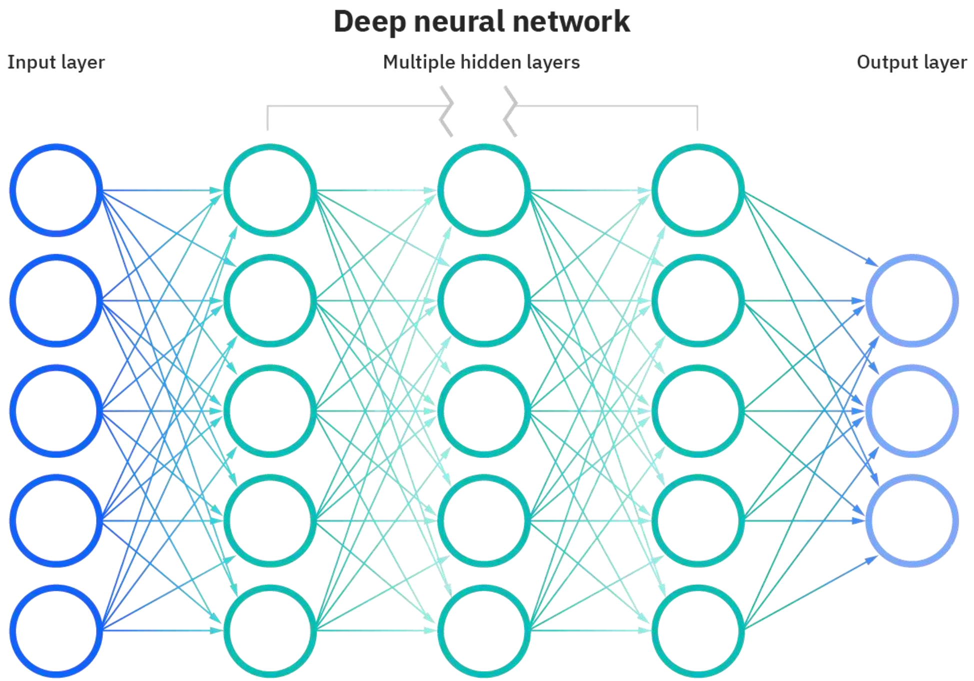
- Artificial Intelligence (AI) refers to the creation of intelligent machines that can perform tasks that typically require human intelligence, such as recognizing speech, making decisions, and understanding natural language. Main areas:
  - Expert systems (rule-based, knowledge-based)
  - Machine learning

- Machine Learning (ML):
  - Supervised learning (labelled data)
  - Unsupervised learning (data not labelled, for clustering)
  - Reinforcement learning
  - Semi-supervised learning

Deep Learning

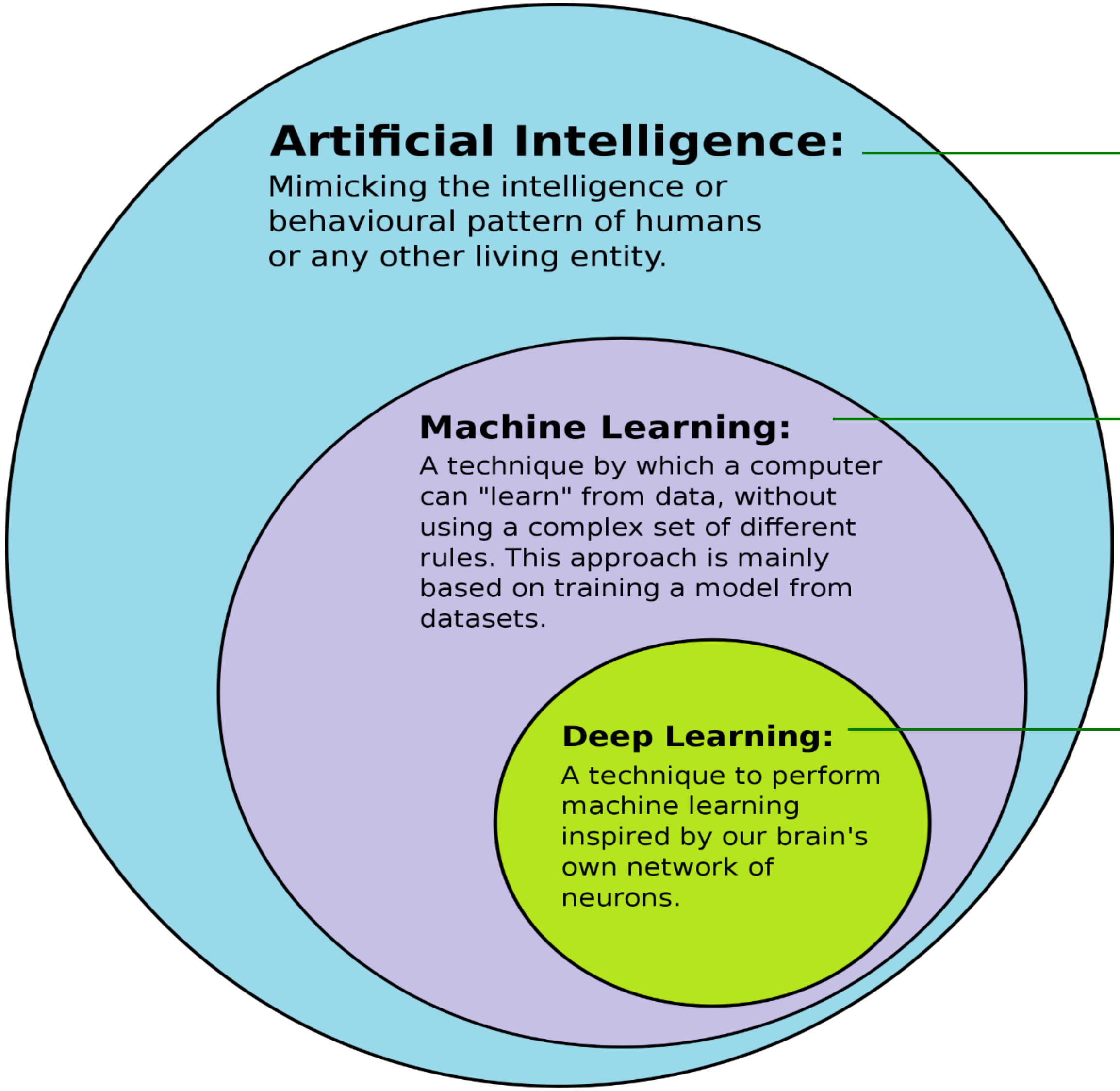
- Deep Learning (DL): typically involves deep neural networks (DNN) or similar architecture with multiple layers in the network.

- Deep learning structures:
  - Deep neural networks
  - Convolutional neural networks
  - Deep belief networks
  - Recurrent neural networks
  - Long short-term memory
  - Graph neural networks
  - Transformer
  - …



A typical DNN with multiple hidden layers.

Artificial Intelligence, Machine Learning, & Deep Learning



Expert systems
(Rule bases,
Knowledge bases)

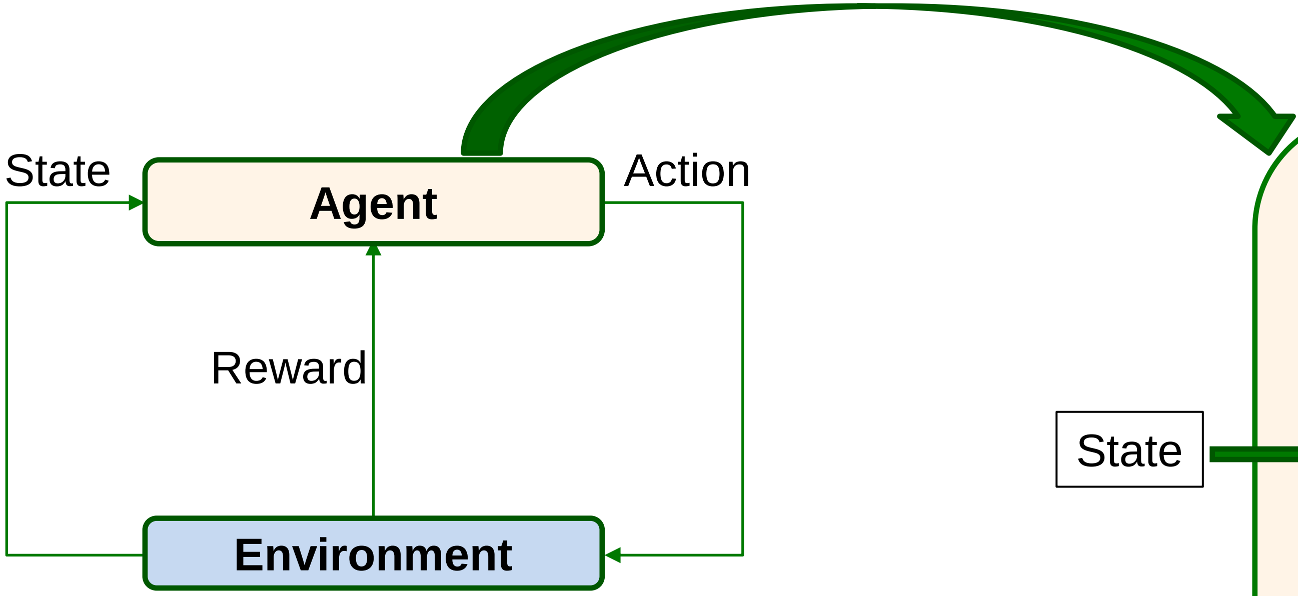
Supervised learning
Unsupervised learning
Reinforcement learning

DNN, CNN, DBN,
RNN, LSTM, GNN,
GAN, Transformer, etc.

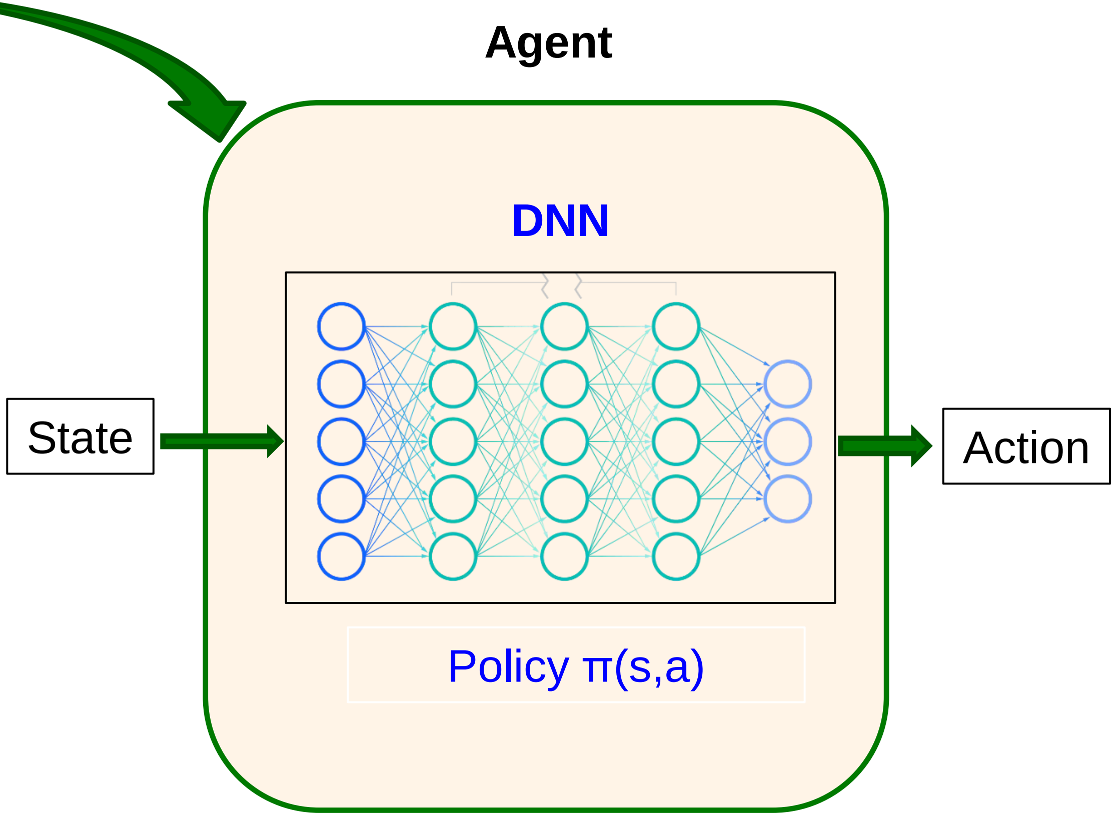
https://en.wikipedia.org/wiki/Deep_learning

Deep Reinforcement Learning

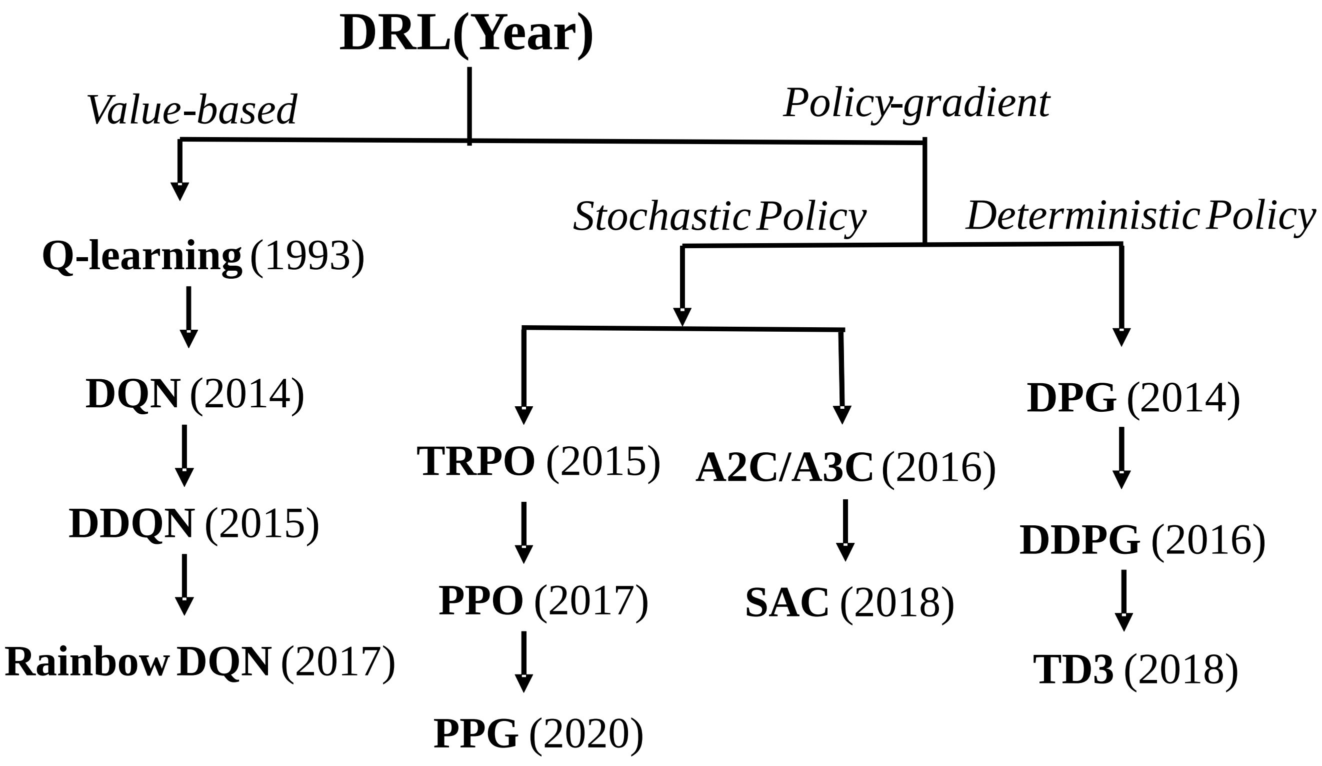
- Reinforcement Learning + Deep Neural Network ➔ DRL



RL: Essentially a trial-and-error process



Mythology of Reinforcement Learning



- **Value:** value of state-action pair (*s*, *a*)
- **Policy:** map state to action
  - **Stochastic policy**: probabilistic distribution of action
  - **Deterministic policy**: deterministic action

- DQN: deep *Q*-learning
- DDQN: Dual deep Q-learning
- *Rainbow DQN: 'Rainbow' deep Q-learning*
- TRPO: Trust region policy optimization
- PPO: Proximal policy optimization
- *PPG: Phasic policy gradient*
- A2C: Advantage Actor-Critic
- A3C: Asynchronous Advantage Actor-Critic
- *SAC: Soft Actor-Critic*
- DPG: Deterministic policy gradient (DPG)
- DDPG: Deep Deterministic Policy Gradient
- *TD3: Twin Delayed Deep Deterministic policy gradient (TD3)*

CURENT

AlphaGo - An Epic Achievement of Deep Learning

- Go, also known as Weiqi or Baduk
  - A strategic board game originated in China over 2500 years ago
  - Broadly considered by mathematicians and computer scientists as the most complex board game and the best testbed for artificial intelligence

- Big news in AI in 2016
  - **AlphaGo (by Google DeepMind)** beat World Champions in 2016, at least 30 years earlier than expected.
  - Based on deep reinforcement learning
  - **A featured cover article was published in *Nature* in 2016 by the AlphaGo team**

- Further improved:
  - AlphaGo Zero (2017), AlphaZero (2018), and MuZero (2019)

D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, et al. “Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, no. 7587, pp. 484-489, 2016.

Deep Learning: Motivation and Development

- Motivation and Development
  - The core idea behind deep learning: successive layers of representation.
  - Examples of representation: 1) image in its RGB matrix; 2) figure in its binary code.
  - The term “deep” refers to the multiple layers that are connected end to end to learn the data representations → automates the feature extraction.
  - The idea of multi-layer representation is based on the assumption that the data in the real-world can all be regarded as composition of features.
  - Current research works have explained how multiple-layer network in a hierarchical structure captures local features and gradually forms the high-level concept.
  - It is combined with reinforcement learning to form deep reinforcement learning (the key structure of AlphaGo).
- Why is “deep” so powerful?
  - Many more hidden layers allowing NN to learn more complex patterns with more complex representations.
  - Enabled by new hardware – GPU, TPU, etc.
  - Better software, better data management, etc.
- Majority of recent AI applications in power are based on deep learning.

Deep Learning Application in Power

- Deep CNN-based Contingency Screening with Uncertain Scenarios

Multi-scenario Security Screening

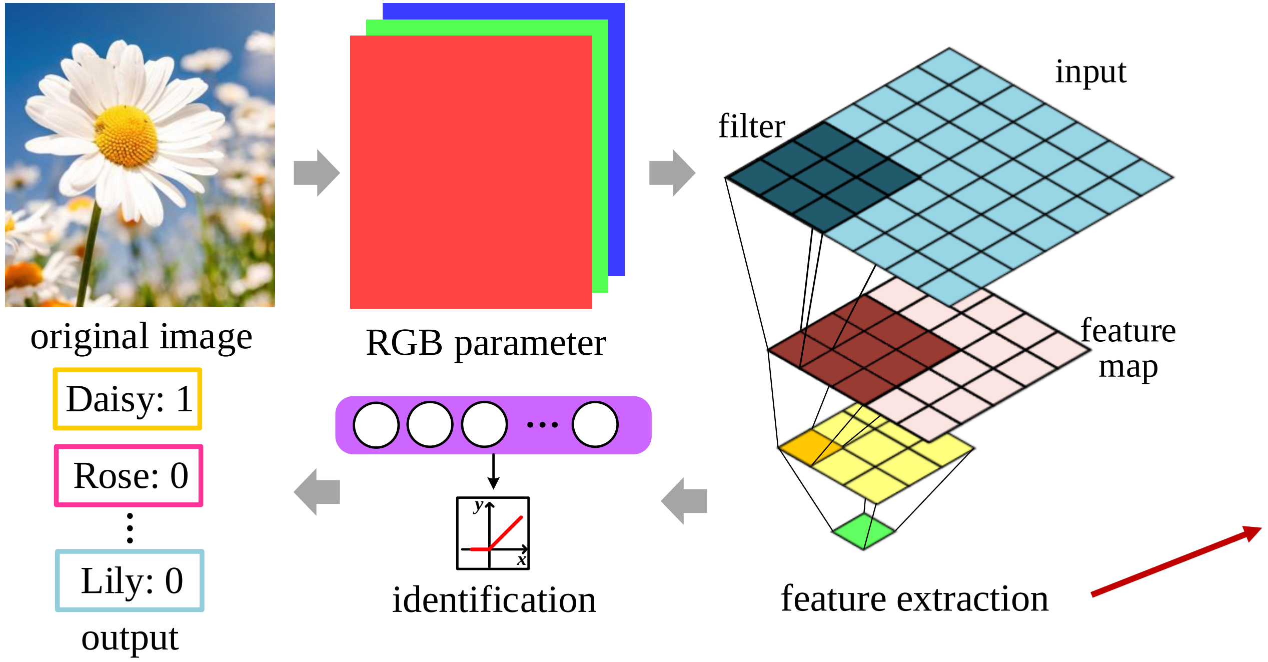
- The **increasing penetration of renewable energy** makes the traditional N-1 contingency screening highly challenging when a large number of **uncertain scenarios** need to be combined with **security screening.**

- The combination can be a very complicated search problem, e.g., in the scale of N^S where N is the number of (uncorrelated) wind plants and s is the number of scenarios.

- A data-driven method, similar to image-processing technique using **deep convolutional neural network (Deep CNN)** method, is proposed for **accelerating multi-scenario N-1 contingency screening**.

A Brief on Deep Convolutional Neural Network

- Deep convolutional neural network (deep CNN): an artificial neural network with multiple hidden layers.
- Strong automatic feature extraction ability in possessing data with grid-like structure, i.e., image data.
- With a hierarchical structure, it mimics the visual cortex of human.

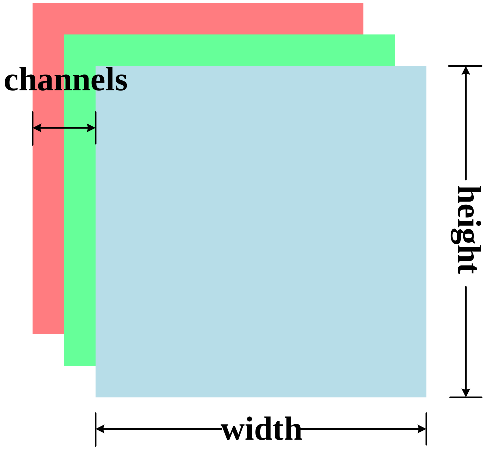


Convolution operation:

$$I'(i,j)=\sum_{u=1}^{c}\sum_{v=1}^{c} I(u,v)\omega(u,v)+b$$

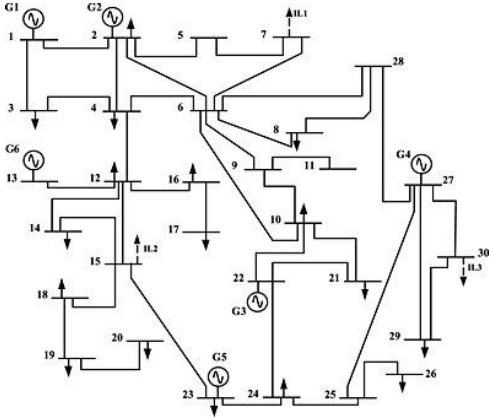
Analogy to Image Classification

Original image



Pixel matrices

Power system topology



$$\begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix}_{n\times n} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}_{n\times n}$$

$$\begin{bmatrix} P_{inj,1} & 0 & \cdots & 0 \\ 0 & P_{inj,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_{inj,n} \end{bmatrix}_{n\times n} \begin{bmatrix} Q_{inj,1} & 0 & \cdots & 0 \\ 0 & Q_{inj,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_{inj,n} \end{bmatrix}_{n\times n}$$

Bus admittance matrices and power injection matrices

Similarity between image processing data and power system raw data: grid-like structure and sparsity.

CURENT

Mapping Power System Data to Deep CNN Input

$$P_{inj,i} = v_i \sum_{j=1}^{n} v_j (g_{ij} \cos\theta_{ij} + b_{ij} \sin\theta_{ij})$$

$$Q_{inj,i} = v_i \sum_{j=1}^{n} v_j (g_{ij} \sin\theta_{ij} - b_{ij} \cos\theta_{ij})$$

System parameters

$$\boldsymbol{G} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix}_{n\times n} \quad \boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}_{n\times n}$$

$$\boldsymbol{P} = \begin{bmatrix} p_{inj,1} & 0 & \cdots & 0 \\ 0 & p_{inj,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{inj,n} \end{bmatrix}_{n\times n} \quad \boldsymbol{Q} = \begin{bmatrix} q_{inj,1} & 0 & \cdots & 0 \\ 0 & q_{inj,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_{inj,n} \end{bmatrix}_{n\times n}$$

Simplification

$$\boldsymbol{P} = \begin{bmatrix} p_{inj,1} & p_{inj,2} & \cdots, & p_{inj,n} \end{bmatrix}_{1\times n}$$

$$\boldsymbol{Q} = \begin{bmatrix} q_{inj,1} & q_{inj,2} & \cdots, & q_{inj,n} \end{bmatrix}_{1\times n}$$

$$\boldsymbol{B} = \begin{bmatrix} b_{11} & b_{22} & \cdots, & b_{nn} \end{bmatrix}_{1\times n}$$

$$\boldsymbol{V} = [v_1, v_2, \ldots, v_n]_{1xn}$$

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_n]_{1xn}$$

Output

CURENT

Deep CNN for Security Screening

Convolution operation: Aggregate local features with different weights

$$I_{new}(i,j)=\sum_{u=0}^{c-1}\sum_{v=0}^{c-1} I(i+u,j+v)\cdot\omega(u,v)+b$$

Activation function: Delinearize the affine transformation

$$I'_{new}(i,j)=\sigma(I'_{new}(i,j))$$

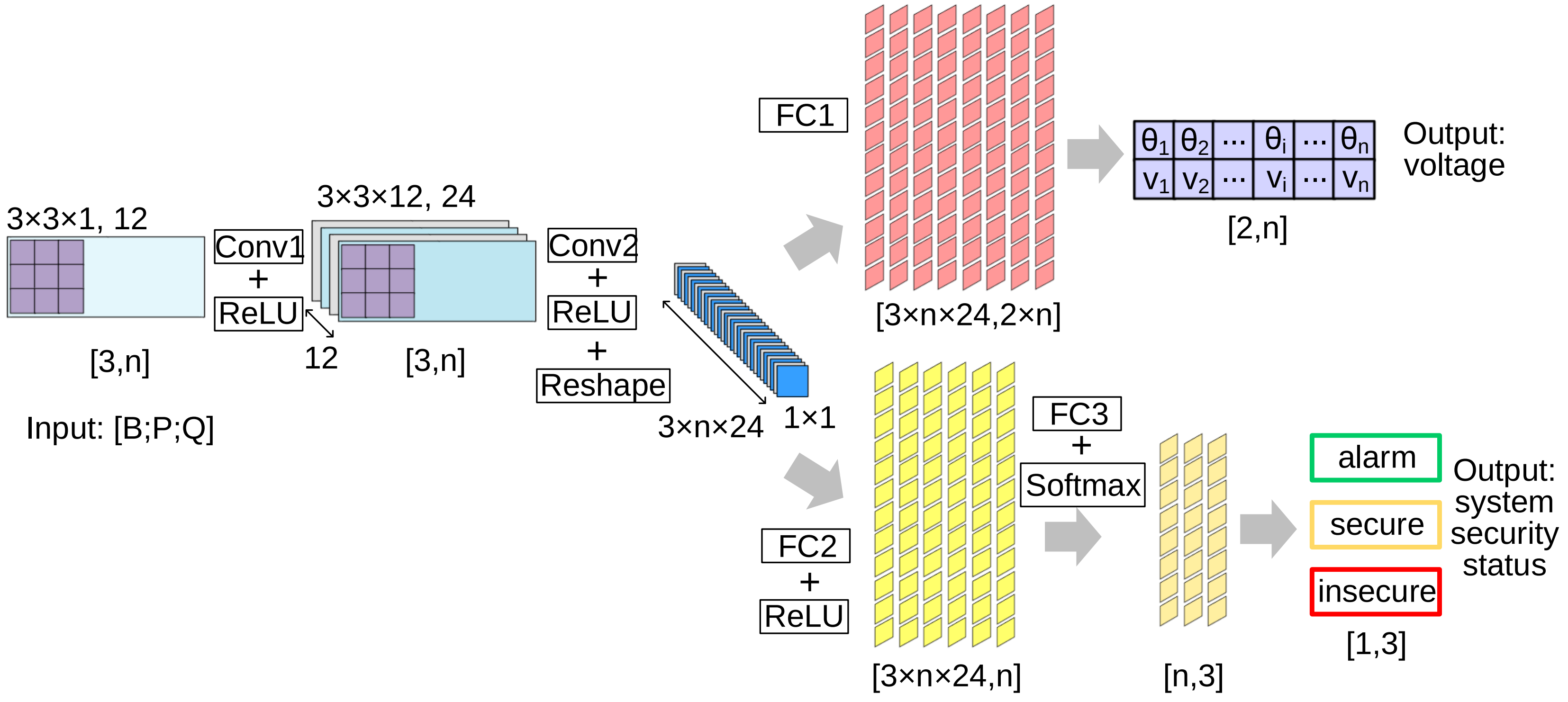
Loss function: Calculate the mean square error between the output and the actual values

$$L=\frac{1}{N_S}\sum_{s=1}^{N_S}(\frac{1}{n}(\sum_{i=1}^{n}(\theta_{i,s}^{*}-\theta_{i,s})^2+\sum_{i=1}^{n}(v_{i,s}^{*}-v_{i,s})^2)-y_s^{*}\log(y_s))$$

Chain rule: Update the weight and bias parameters

$$\omega_l^{(k+1)}=\omega_l^{(k)}-\eta\frac{\partial L}{\partial J_{N_L-1}^{(k)}}\cdot\frac{\partial J_{N_L-1}^{(k)}}{\partial J_{N_L-2}^{(k)}}\cdot\cdots\cdot\frac{\partial J_l^{(k)}}{\partial\omega_l^{(k)}}$$

Design of Deep CNN Structure



- Two convolutional layers, three fully-connected layers
- Input: B; P; Q (3 × n matrix)
- Output: voltage (2 × n matrix); system security status (1 × 3 vector)

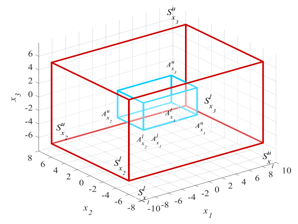
CURENT

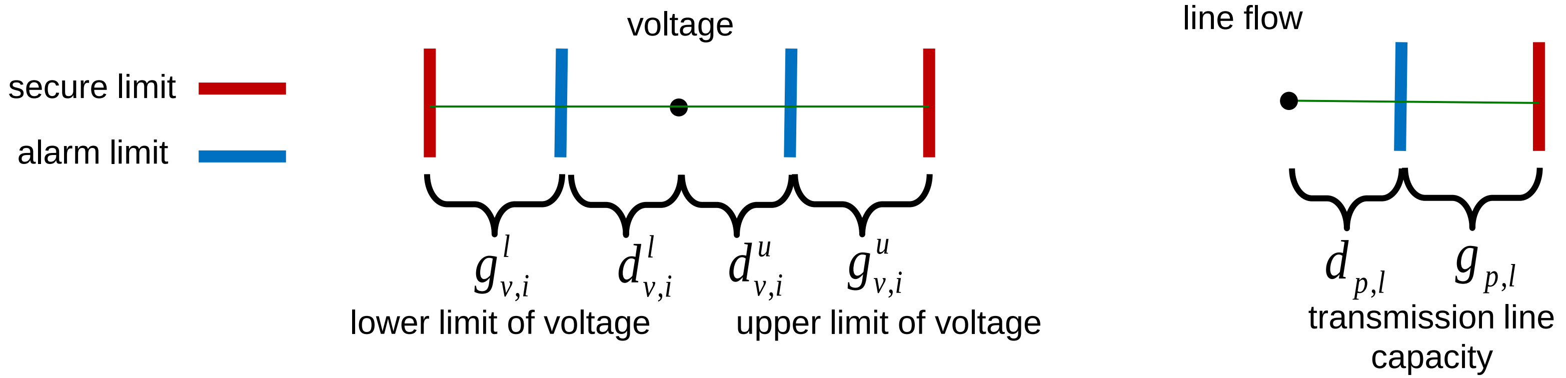
Security Index

- A security index is needed for evaluating system operation status based on power flow results. Here we use:

$$SI = \left[\sum_i \left(\frac{d_{v,i}^{u}}{g_{v,i}^{u}}\right)^{2m} + \sum_i \left(\frac{d_{v,i}^{l}}{g_{v,i}^{l}}\right)^{2m} + \sum_l \left(\frac{d_{p,l}}{g_{p,l}}\right)^{2m} \right]^{\frac{1}{2m}}$$

$$\begin{aligned} SI &= 0: && \textit{secure} \\ 0 < SI &\le 1: && \textit{alarm} \\ SI &> 1: && \textit{insecure} \end{aligned}$$





Case Studies (1)

AC Power Flow Results of Deep CNN

Case	No. of samples		Errors		Training time(s)	Classification Accuracy
	Training	Test	θ	v		
9	3292	1412	6.1e-3	7.2e-4	11.42	**97.24%**
30	4262	1066	1.5e-3	5.4e-4	23.06	**96.25%**
57	3360	1440	4.9e-3	1.6e-3	31.59	**99.24%**
118	3027	1298	7.5e-3	2.9e-4	57.88	**100.00%**
181(WECC)	2530	1085	5.7e-2	3.8e-3	65.04	**97.70%**
300	3445	1477	6.9e-2	2.3e-3	148.91	**99.05%**
1354 (Eu.)	3981	1707	1.1e-2	1.9e-3	1548.94	**96.84%**

- Tested on the IEEE 9, 30, 57, 118, and 300-bus systems, WECC 181-bus system, and European 1354-bus system
- Considered load uncertainty, renewable generation variation, and N-1 contingency in each ACPF case
- Load forecast error ~U(0.8,1.2), Wind speed forecast error ~N(0,0.05^2)

CURENT

Case Studies (2)

TEST TIME COMPARISON & ACCELERATION

Case	Test size	Test time (s) (deep CNN)	Test time (s) (model-based)	Acceleration ratio
9	1412	0.017	3.500	**206**
30	1066	0.016	3.303	**206**
57	1440	0.018	4.323	**240**
118	1298	0.021	4.905	**234**
181 (WECC)	1085	0.025	4.655	**186**
300	1477	0.044	10.15	**231**
1354 (Eu.)	1707	0.264	34.13	**129**

- Software: TensorFlow
- Hardware: Nvidia GeForce GTX 1080 Ti Graphic Card with 11 GB memory and 1.582 GHz core clock

Case Studies (3)

TEST TIME COMPARISON OF DCNN VS. ANN

Case	Errors (ANN)		Errors (deep CNN)		Classification Accuracy (ANN)	Classification Accuracy (deep CNN)
	θ	v	θ	v		
9	2.0e-2	2.3e-3	6.1e-3	7.2e-4	**91.64%**	**97.24%**
30	9.0e-3	2.4e-3	1.5e-3	5.4e-4	**87.43%**	**96.25%**
57	2.7e-2	9.0e-3	4.9e-3	1.6e-3	**92.43%**	**99.24%**
118	2.7e-2	9.0e-4	7.5e-3	2.9e-4	**98.54%**	**100.00%**
181	1.8e-1	1.3e-2	5.7e-2	3.8e-3	**75.94%**	**97.70%**
300	2.0e-1	5.7e-3	6.9e-2	2.3e-3	**78.00%**	**99.05%**
1354	--	--	1.1e-2	1.9e-3	**--**	**96.84%**

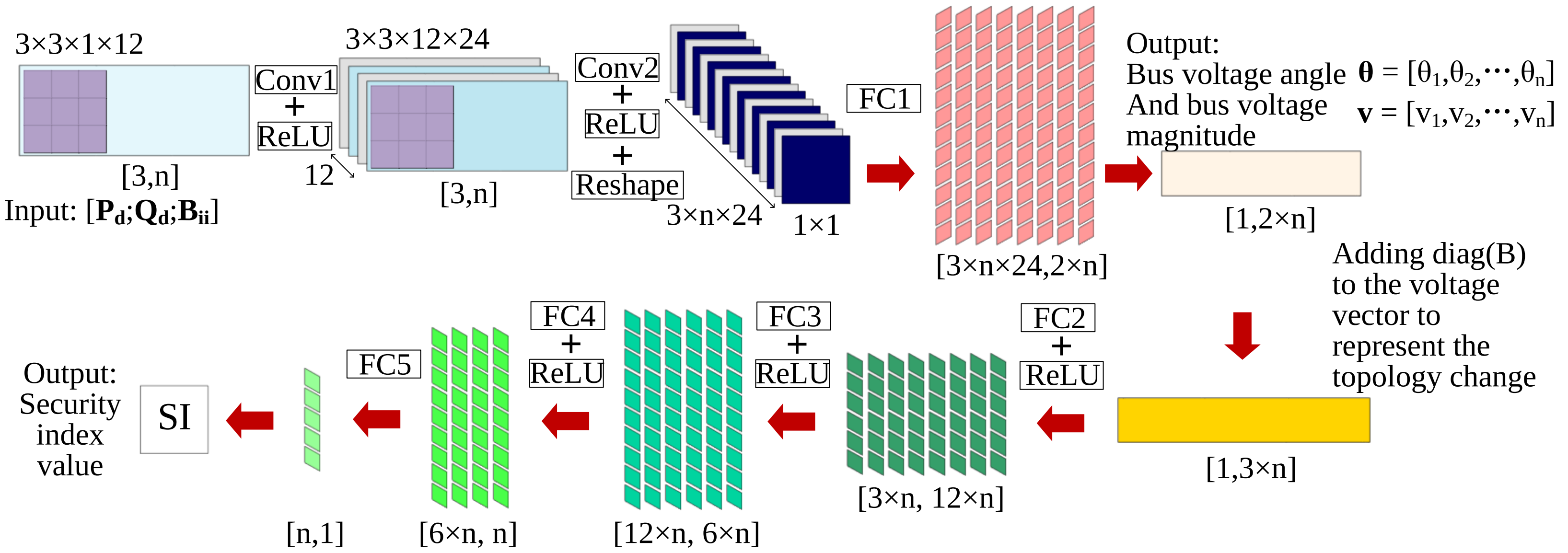
- The results on 1354-bus system is not available for ANN due to out-of-memory (OOM) issue.
- Deep CNN is more efficient in feature extraction and computation.

Deep CNN-based Contingency Screening with Uncertain Scenarios

- **Summary**
  - A power system can be modeled with matrices similar to the models of images such that some AI-based image processing technique can be utilized.
  - The deep CNN is constructed as a classifier to evaluate system security status based on power system raw data.
  - Compared with the conventional model-based method, the proposed deep CNN has high computational efficiency (achieving over 100x speedup), while maintaining considerable classification accuracy (98.05% accuracy in average), which can be a promising tool for future real-time applications.

Y. Du, F. Li, J. Li, T. Zheng, "Achieving 100x Acceleration for N-1 Contingency Screening with Uncertain Scenarios using Deep Convolutional Neural Network," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3303-3305, July 2019.

Extension: Deep CNN for Cascading Failure Assessment



- DCNN is combined with Depth-First Search

- Two convolutional layers and five fully-connected layers
- Loss function: Mean Square Error

$$L = \frac{1}{N_S}\sum_{s=1}^{N_S}\left(\frac{1}{n}\sum_{i=1}^{n}(\theta_{i,s}^{*}-\theta_{i,s})^2+\frac{1}{n}\sum_{i=1}^{n}(v_{i,s}^{*}-v_{i,s})^2+(SI_s^{*}-SI_s)^2\right)+\frac{\alpha}{2}\boldsymbol{\omega}^T\boldsymbol{\omega}$$

Case	Training time(s)	Test time (s) (deep CNN)	Test time (s) (model-based)	Acceleration ratio
57-bus	906	0.16	225.85	**1,412**
1354-bus	24,692	5.7	8,185	**1,436**

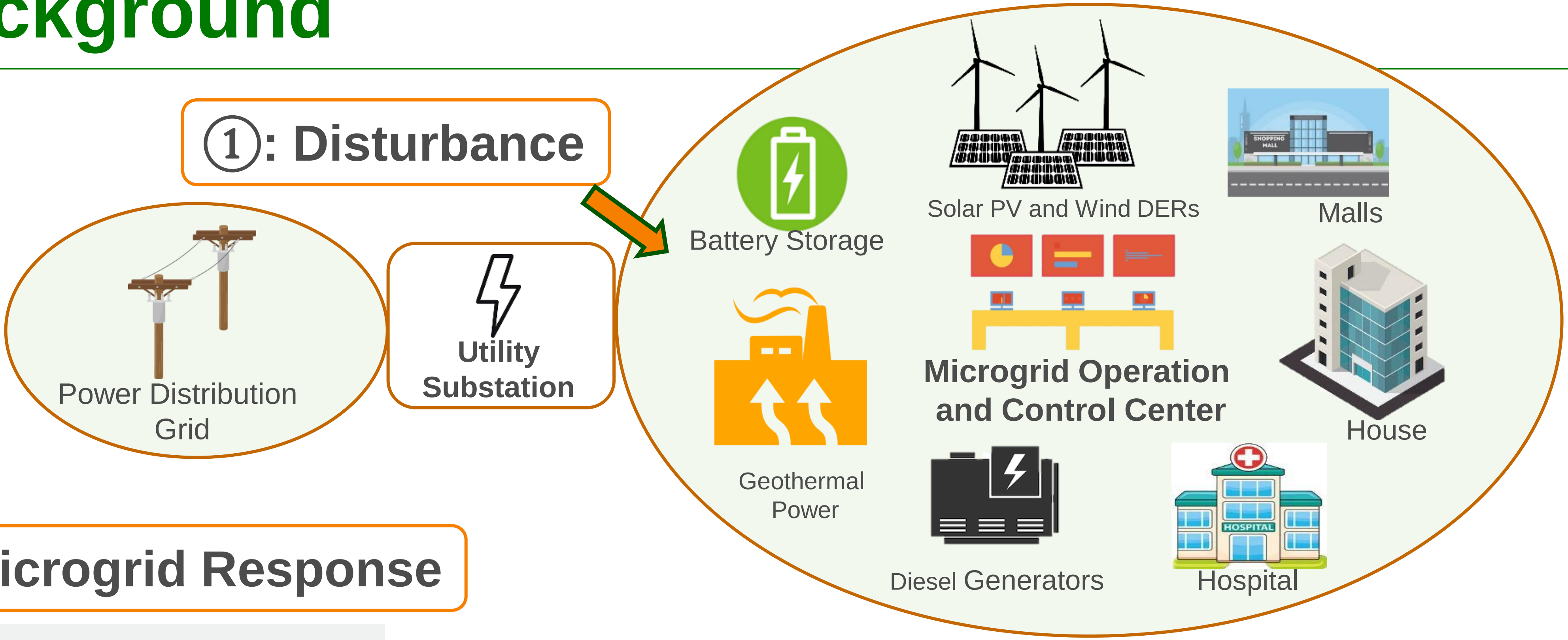
Y. Du, F. Li, et al, "Fast Cascading Outage Screening based on Deep Convolutional Neural Network and Depth-First Search," *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 2704-2715, July 2020.

Deep Learning Application in Power:

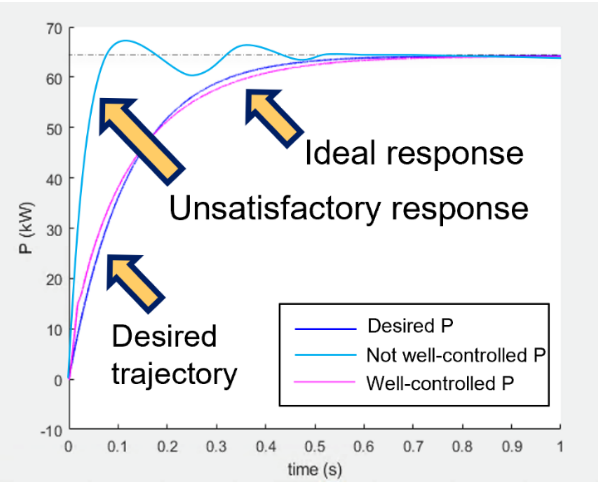
- Physics-informed DRL for Inverter PQ Control in Microgrid

Background

①: Disturbance



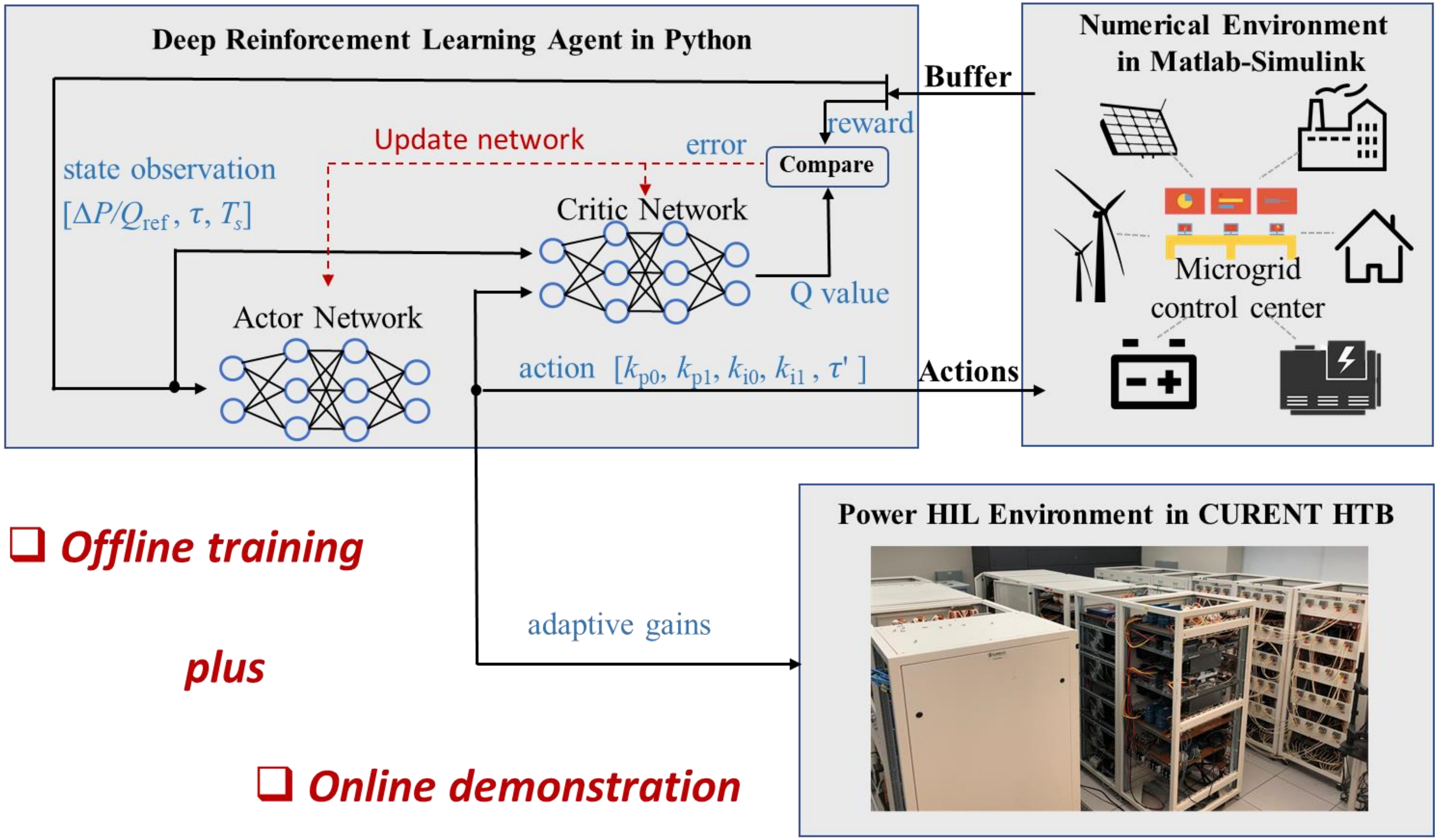
② Microgrid Response



Key Idea: In the event of a disturbance, keep the actual response following the desired trajectory by adaptively adjusting the control gains [1].

Li, H., Li, F., Xu, Y., Rizy, D.T. and Kueck, J.D., “Adaptive voltage control with distributed energy resources: Algorithm, theoretical analysis, simulation, and field test verification,” *IEEE Transactions on Power Systems*, *25*(3), pp.1638-1647.

Physics-informed DRL for Microgrid Control



- ❑ ***Offline training***

plus

- ❑ ***Online demonstration***

Diagram of implementing model-free reinforcement learning in microgrid PQ control

- ❑ Model-based analysis reduce learning space from **function space** to **real space**

$$\begin{cases} k_p(t) = \boxed{k_{p0}} + \boxed{k_{p1}} e^{-t/\tau'} \\ k_i(t) = \boxed{k_{i0}} + \boxed{k_{i1}} e^{-t/\tau'} \end{cases}$$

$$k_p(t), k_i(t) \in \boldsymbol{f}(t)$$

$$k_{p0}, k_{p1}, k_{i0}, k_{i1} \in \boldsymbol{R}$$

CURENT

Test Microgrid and Training Results

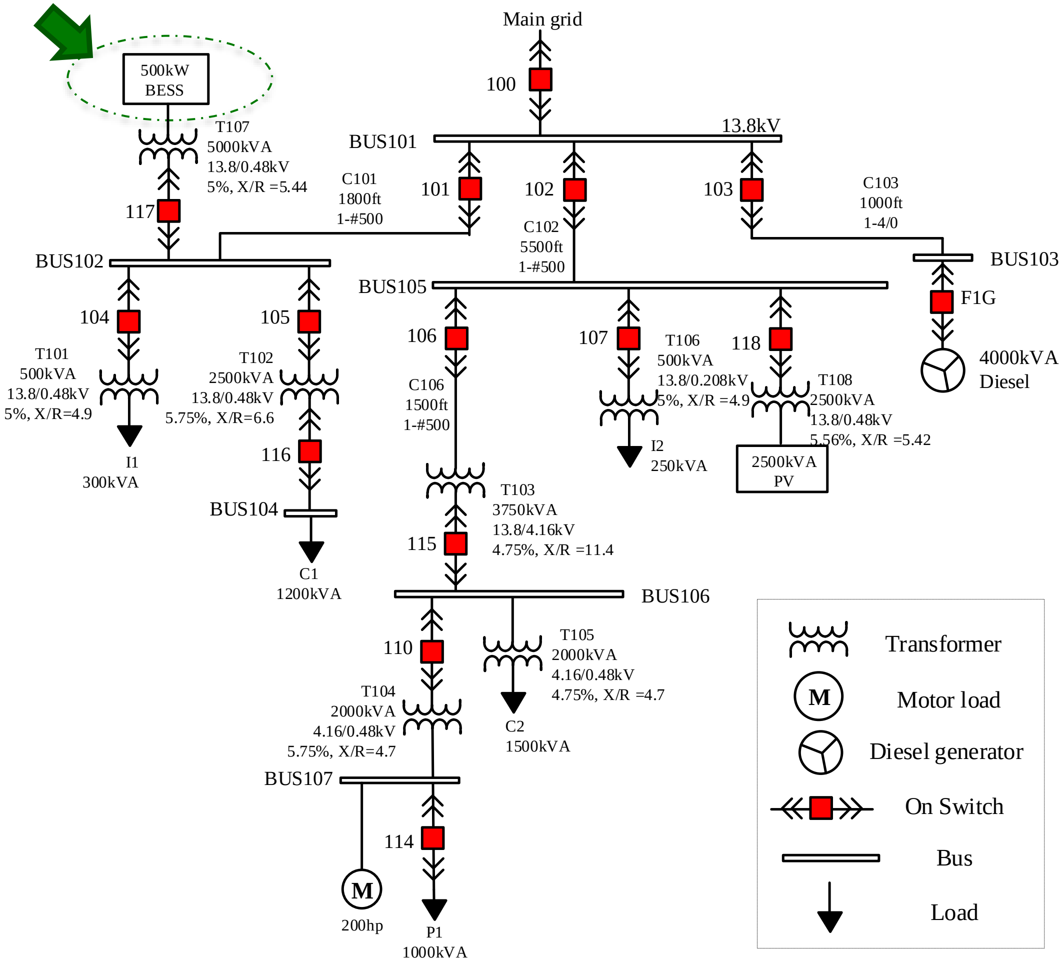
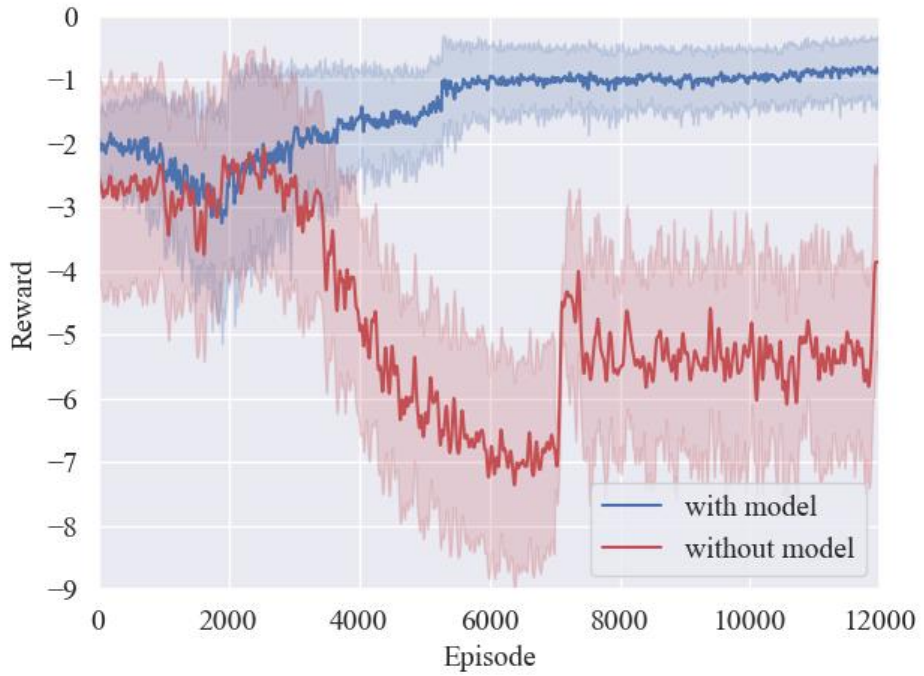
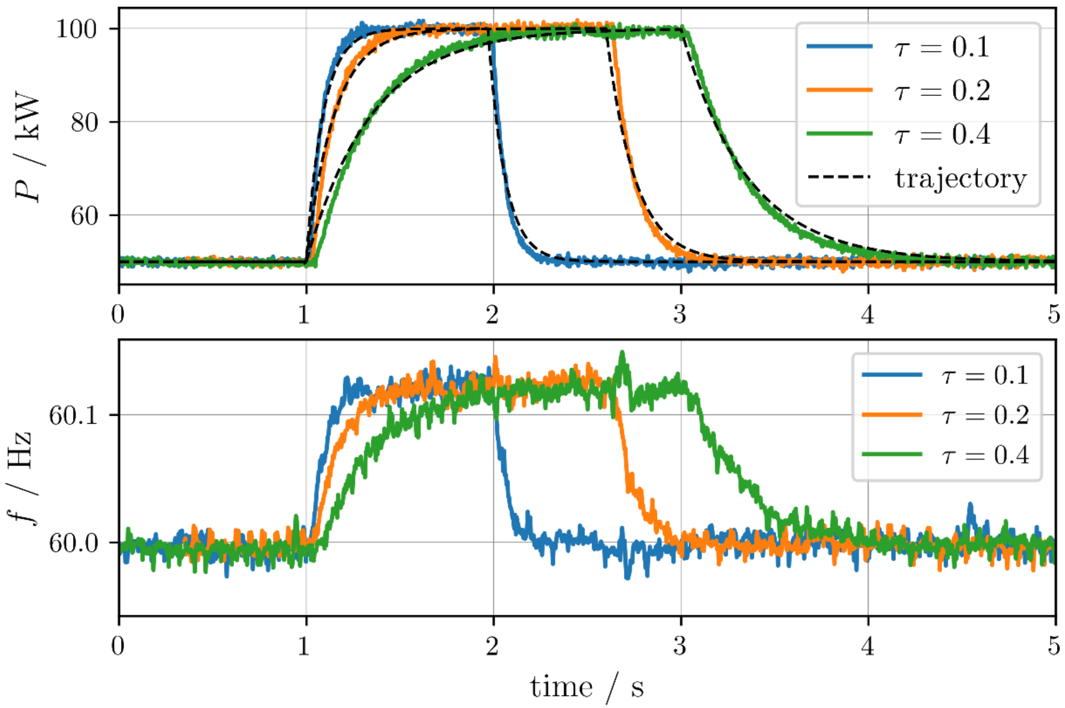


Diagram of modified Banshee microgrid

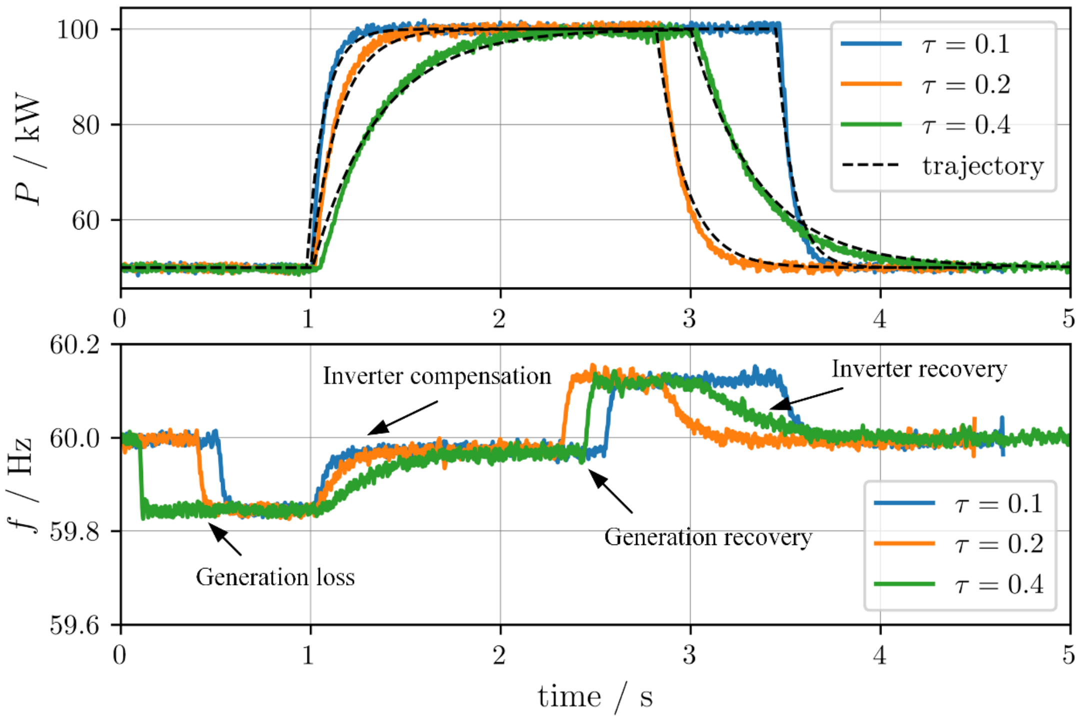


Reward curve with and without model-based analysis

Validation in CURENT Hardware Testbed



Scheduling reference change

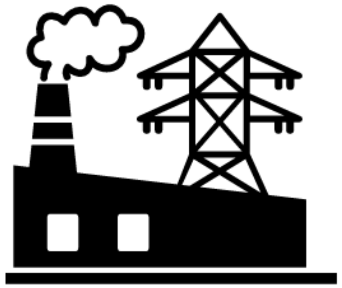


Generation reduction & recovery

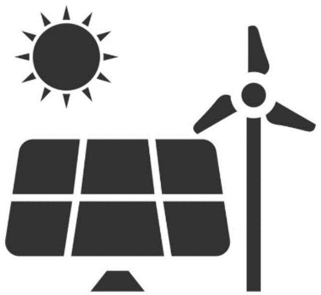
- ❑ Inverters can be freely assigned any time constant and respond either slow or fast.
- ❑ The proposed physics-informed DRL algorithm is validated under scheduling reference change and generation reduction and recovery.

Vision of Future Directions of AI/DL in Power

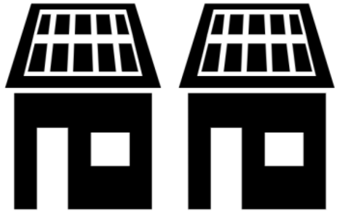
DL for Different Power System Applications



Power transmission system
- ✓ Cascading outage screening
- ✓ Voltage stability assessment
- ✓ Bulk system restoration support

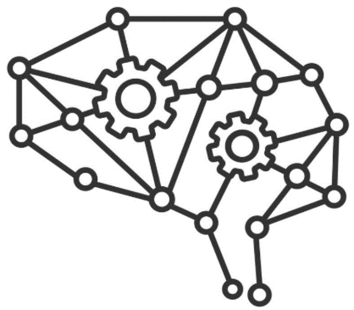


Power distribution system
- ✓ Intelligent multi-MGs energy management
- ✓ Resilient multi-MGs defense
- ✓ Microgrid PQ control
- ✓ Line outage cause classification



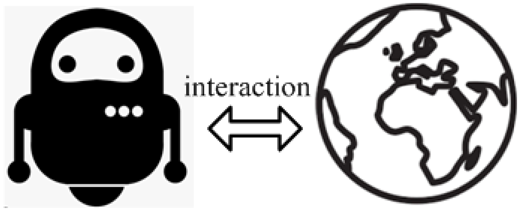
Loads and consumers
- ✓ Multi-zone residential HVAC control
- ✓ Robust Load restoration

Deep learning (DL)



DL dominates

DL+ reinforcement learning (RL) = Deep RL (DRL)



DRL dominates

When we apply DL in power, what do we expect to achieve?

Advanced AI/DL methods have been explored in almost everywhere in the field of power systems such as load/RES forecasting, power system operation and planning, optimal control, etc.

Simplify computation

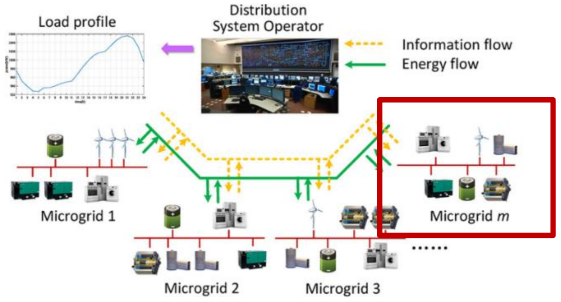
- Highly complex/non-convex issues: hard to solve using conventional optimization.
- Time-consuming calculation: dynamic simulations, two stage robust optimization.

Completely/partly model free

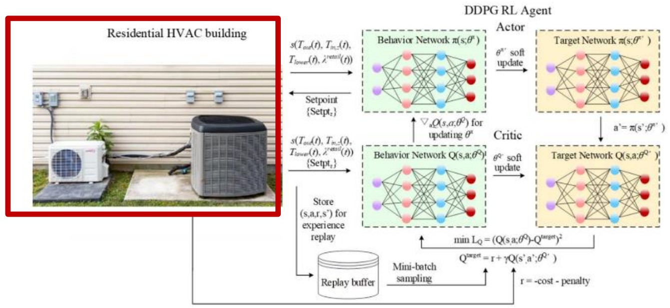
- Intractable modeling: hard to model the issue, residential load behavior.
- Privacy requirements: no rights to access the model

Data processing

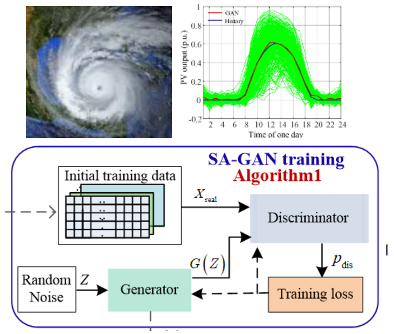
- Forecasting: automatic data generation with high accuracy
- Data filtering: organize good data for decision-making
- Unsupervised feature extraction: no need for manual data analysis



Replace non-convex MG models with DNN



Protect customer privacy of the residential side



Automatically exact high-dimensional data features

Challenge of Current AI: Transfer Gap

- The transfer gap: Are the lab-based AIs suitable for real-world power system conditions?

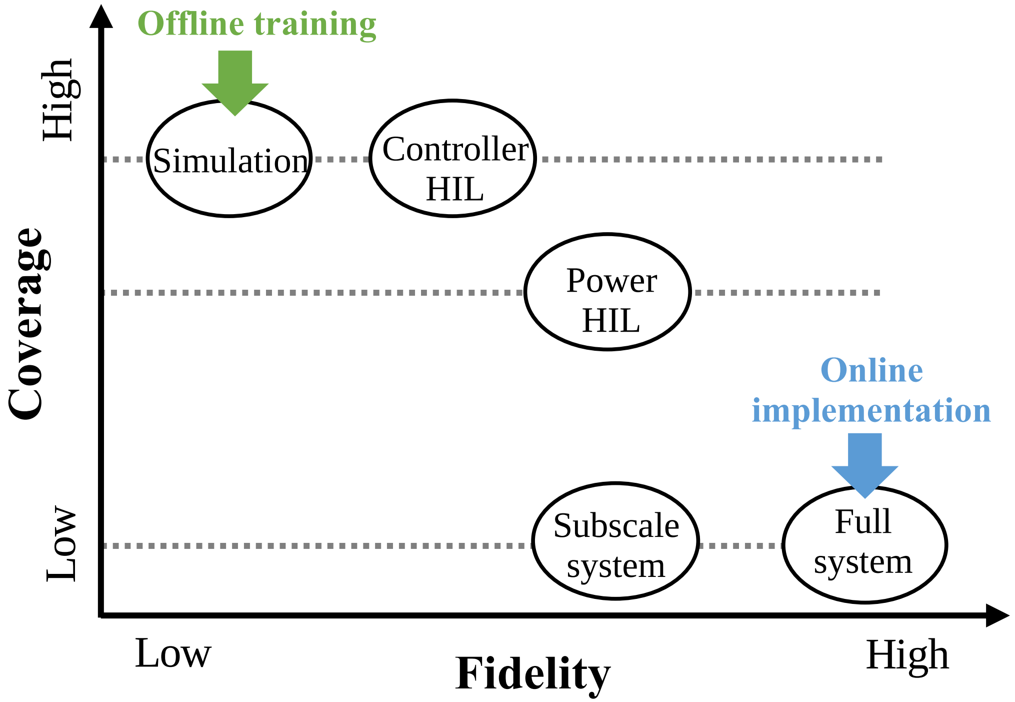
Training/Learning

the gap

Real-world application

Difficulty to involve real-world stuffs

- Data scarcity of abnormal operation conditions
- Over simplified power grid simulation (environment)



Test bed types [1] and options for model-free RL environment

[1] “IEEE Standard for the Testing of Microgrid Controllers,” IEEE 2030.8-2018.

Challenge of Current AI Methods: Action Security

- Requirements of reliable on-line actions

- **Security in training/exploration**

  RL agent needs **sufficient exploration** of the environment. Sometimes, the explored actions are harmful for the system.

- **Security in action**

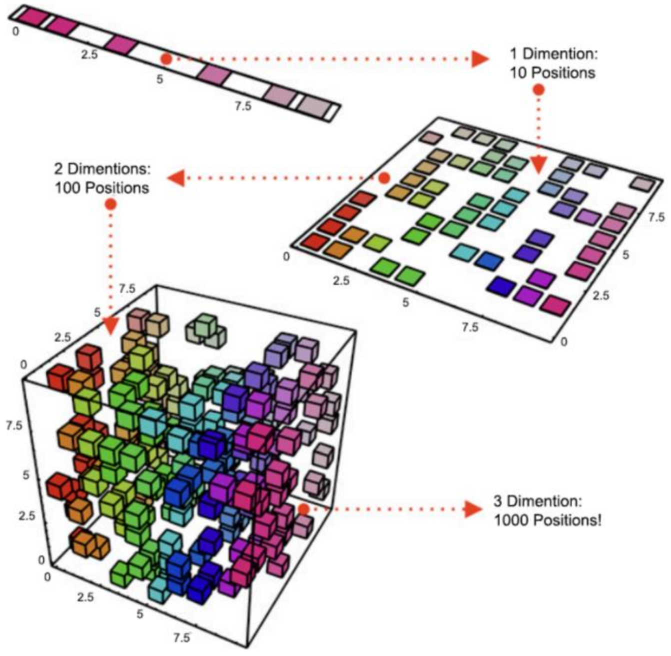
  Model-based controllers can pass **the security test** through eigenvalue analysis or the Lyapunov function before implementation, but RL agent cannot.

- **Efficient training can help**
  Case to case design benefit the targeted issue/systems.
  Enhance the security feature in database and environment.
- **Develop specialized hardware-software test bed**
  With protection schemes that can tolerate random exploration to some degree .
- **Integrate domain knowledge**
  Consider physical operational constraints and stability criteria, and use constrained RL and safe RL.
  Employ physics-constrained and physics-informed neural network.

CURENT

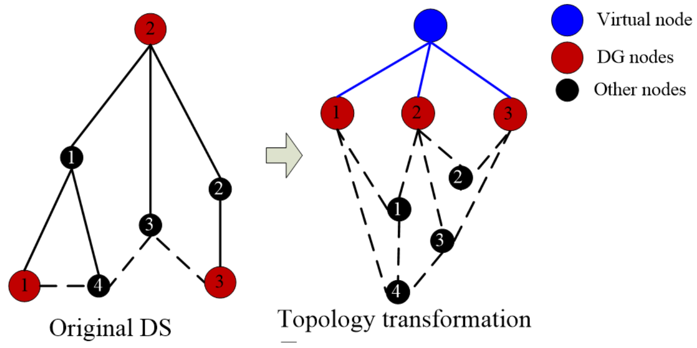
Challenge of Current AI : Scalability & Explainability

- The curse of dimensionality



Curse of dimensionality

- Simplify power system models for AI (physics-informed AI)
  - The expansion of state space and action space will result in an exponential increase in control complexity, thereby increasing the difficulty of exploration and training
  - Solutions:
    - ✓ Increase the capability of existing RL models
    - ✓ Reduce the complexity with domain knowledge (MG topology)



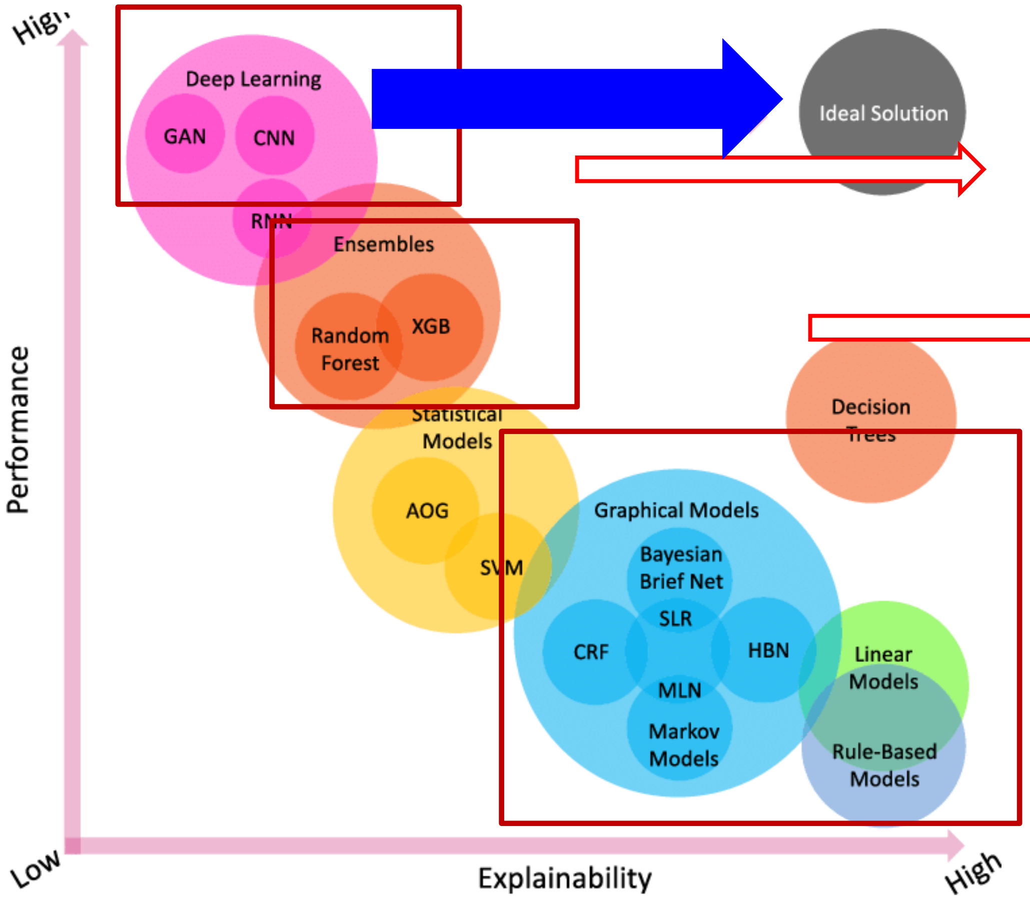
Reduces the action space from exponential to polynomial.

- Physics-informed AI may improve the explainability of AI-based solutions.

J. Zhao, F. Li, S. Mukherjee, C. Sticht, "Deep Reinforcement Learning based Model-free On-line Dynamic Multi-Microgrid Formation to Enhance Resilience," IEEE Transactions on Smart Grid, vol. 13, no. 4, pp. 2557-2567, July 2022.

Perspective of Future AI Development

- Making AI-based approach more understandable and explainable



Deep learning

Ensemble learning

Conventional ML

The ideal solution should have both high explainability and high performance.

Explainable/interpretable AI (XAI)

Model explainability vs. model performance for widely used AI techniques [1]

[1] G. Yang, Q. Ye, J. Xia, “Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond”, *Information Fusion*, vol. 77, pp 29-52, 2022.

AI for Power Systems

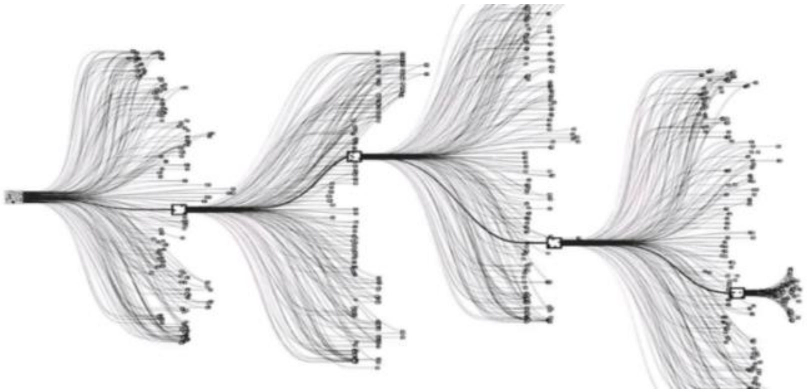
Advanced AI techniques are tools to help us.
How well they perform largely depends on how we use them.

Current concerns

- Transfer gap
- Action security
- Scalability
- Explainability

- ✓ **Efficient off-line training**
- ✓ **Reliable on-line actions**
- ✓ **Better understanding**

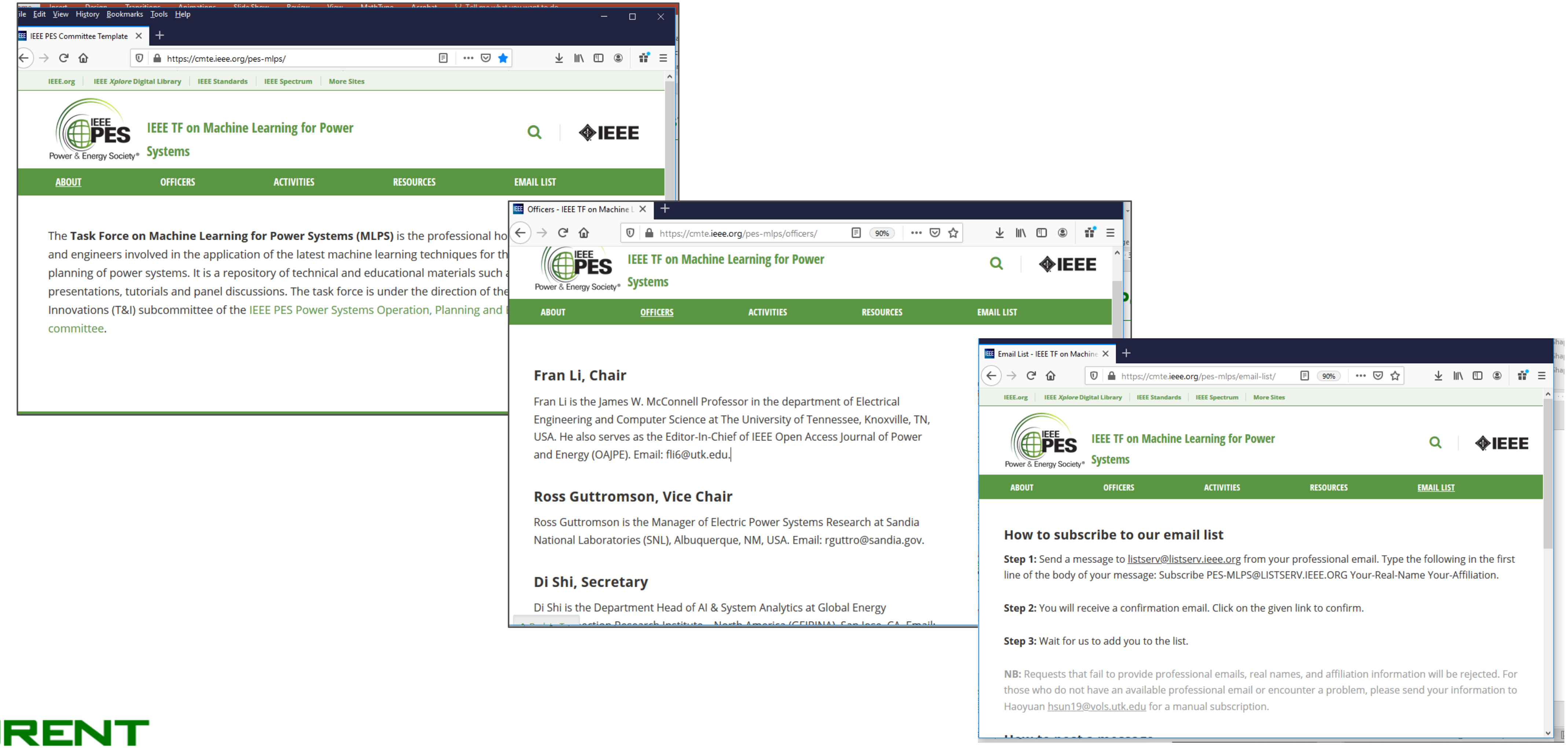
AI & power systems
Promising further development



PES Activities of ML for Power

IEEE WG on Machine Learning for Power Systems

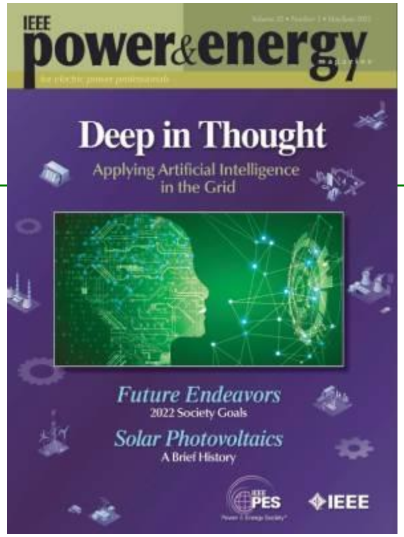
https://cmte.ieee.org/pes-mlps/

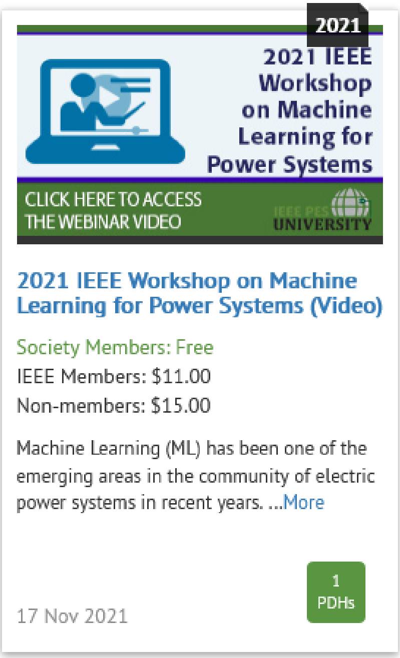


CURENT

Ongoing & Past Activities

- *IEEE Power and Energy Magazine* special issue on Data Analytics and AI (May-June issue), Co-guest editor: Fran Li
- *IEEE Workshop on Machine Learning for Power Systems* (November 17th, 2021): 200+ online participants, Slides and videos available at PES Resource Centers. Organized by Fran Li





- IEEE PES GM22 Supersession (8 speakers)
- IEEE PES GM22 Panel Session (6 speakers)
- IEEE PES GM21 Panel Session (5 speakers)
- IEEE PES GM20 Panel Sessions, two sessions (10 speakers in total)

CURENT

Related Recent Publications

1. **F. Li**, "Successful Applications and Future Challenges of Machine Learning for Power Systems—A Summary of Recent Activities by the IEEE WG on Machine Learning for Power Systems," *IEEE Electrification Magazine*, vol. 10, no. 4, pp. 90-91 & 96, December 2022.
2. Yan Du (*), **F. Li**, Kuldeep Kurte, Jeffery Munk, and Helia Zandi, "Demonstration of Intelligent HVAC Load Management with Deep Reinforcement Learning: Real-World Experience of Machine Learning in Demand Control," *IEEE Power and Energy Magazine*, vol. 20, no. 3, pp. 42-53, May-June 2022.
3. Jin Zhao (*), **F. Li**, Xi Chen, and Qiuwei Wu, "Deep Learning based Model-free Robust Load Restoration to Enhance Bulk System Resilience with Wind Power Penetration," *IEEE Transactions on Power Systems*, vol. 37, no. 3, pp. 1969-1978, May 2022.
4. Hang Shuai (*), **F. Li**, Hector Pulgar-Painemal, and Yaosuo Xue, "Branching Dueling Q-Network Based Online Scheduling of a Microgrid With Distributed Energy Storage Systems," *IEEE Transactions on Smart Grid*, vol. 12 no. 6, pp. 5479-5482, November 2021.
5. Xiao Kou (*), Yan Du (*), **F. Li**, Hector Pulgar-Painemal, Helia Zandi, Jin Dong, and Mohammed Olama, "Model-Based and Data-Driven HVAC Control Strategies for Residential Demand Response," *IEEE Open Access Journal of Power and Energy*, vol. 8, pp. 186-197, May 2021.
6. Yan Du (*), **F. Li**, Jeffery Munk, Kuldeep Kurte, Olivera Kotevska, Kadir Amasyali, and Helia Zandi, "Multi-task Deep Reinforcement Learning for Intelligent Multi-zone Residential HVAC Control" *Electric Power Systems Research (Elsevier)*, vol. 192, article 106959, March 2021.
7. Yan Du (*), Helia Zandi, Olivera Kotevska, Kuldeep Kurte, Jeffery Munk, Kadir Amasyali, Evan Mckee, and **F. Li**, "Intelligent Multi-zone Residential HVAC Control Strategy Based on Deep Reinforcement Learning," *Applied Energy (Elsevier)*, vol. 281, article 116117, Jan. 2021.
8. Yan Du (*), **F. Li**, Tongxin Zheng, and Jiang Li, "Fast Cascading Outage Screening based on Deep Convolutional Neural Network and Depth-First Search," *IEEE Transactions on Power Systems*, vol. 35, no.4, pp. 2704-2715, July 2020.
9. Yan Du (*) and **F. Li**, "Intelligent Multi-Microgrid Energy Management Based on Deep Neural Network and Model-free Reinforcement Learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1066-1076, March 2020.
10. Yan Du (*), **F. Li**, Jiang Li, Tongxin Zheng, "Achieving 100x Acceleration for N-1 Contingency Screening with Uncertain Scenarios using Deep Convolutional Neural Network," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3303-3305, July 2019.
11. **F. Li** and Yan Du (*), "From AlphaGo to Power System AI: What Engineers Can Learn from Solving the Most Complex Board Game," *IEEE Power and Energy Magazine*, vol. 16, issue 2, pp. 76-84, March-April 2018.

What's new in the past a few months in AI?

Large language model (LLM) and beyond LLM

Google

Meta

Microsoft

……

AI is changing the world.

ChatGPT

Midjourney Gen5 model

GPT4

LLaMA

Auto GPT

……

Visual ChatGPT

Microsoft Copilot

CURENT

Key features

- **Large scale model**
  - Huge amount of data
  - Deep architecture
  - High computing power
  - Human-like language understanding.

- **Multi-modal input and output**
  - Process and learn from multiple modalities
  - Feature fusion
  - Cross-modal retrieval
  - Contextual understanding.

- **Automatic learning**
  - Continuous learning
  - Automatic feedback loops
  - Reinforcement/Transfer learning
  - Online learning.

Open question:

How will the large-scale AI model change the power community?

CURENT

Acknowledgements

The presenter acknowledges the support of relevant project sponsors and my group members (Yan Du, Evan Mckee, Jin Zhao, Buxin She, and Hang Shuai) for their contributions to the materials used in the slides.

Thank you!

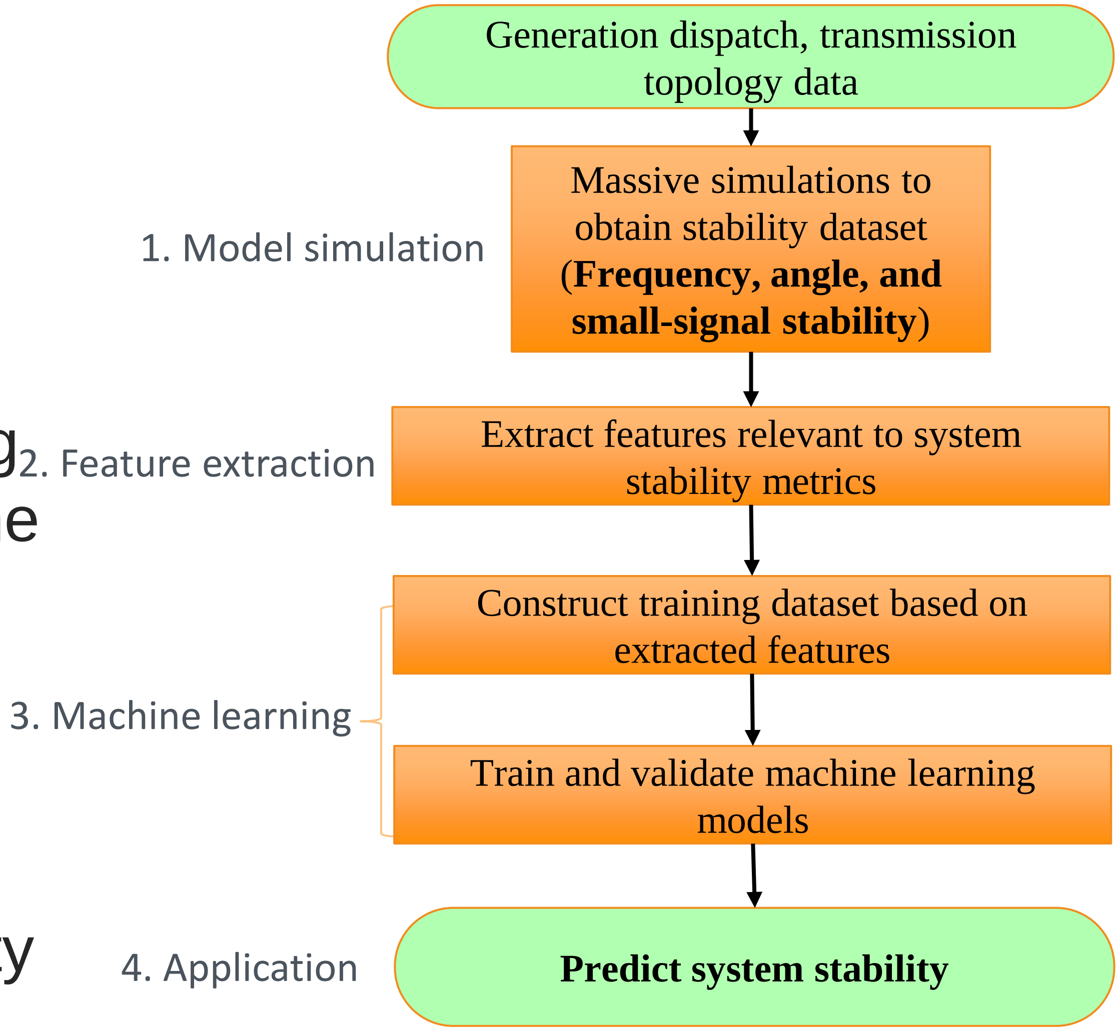
Q&A

Machine Learning for Fast Stability Screening Using Power Dispatch as Inputs

Presented by Yilu Liu
Work performed by Annie Zhao, Jenny Dong, Cici Jia
University of Tennessee and Oak Ridge National Laboratory
Liu@utk.edu

Machine Learning Based Stability Assessment

- **Objective**
  - Fast screening of the system stability (angle, frequency and small-signal stability) at dispatch planning stage.
- **Challenge**
  - Save time to run full dynamic modeling simulations both near real time & offline
  - Non-linear correlation of stability margins with operation variables (dispatch)
- **Approach**
  - Use machine learning to obtain stability margin based on simulations covering 15 min or hourly dispatch.



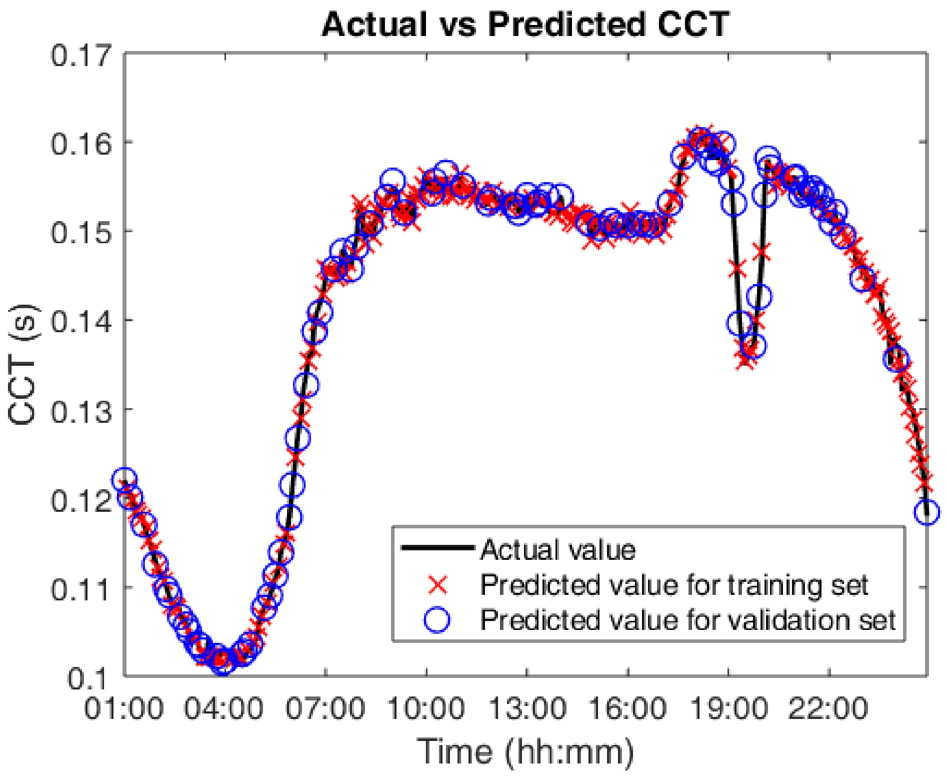
Framework of AI-based system stability prediction

Background - Transient Stability Metrics Used

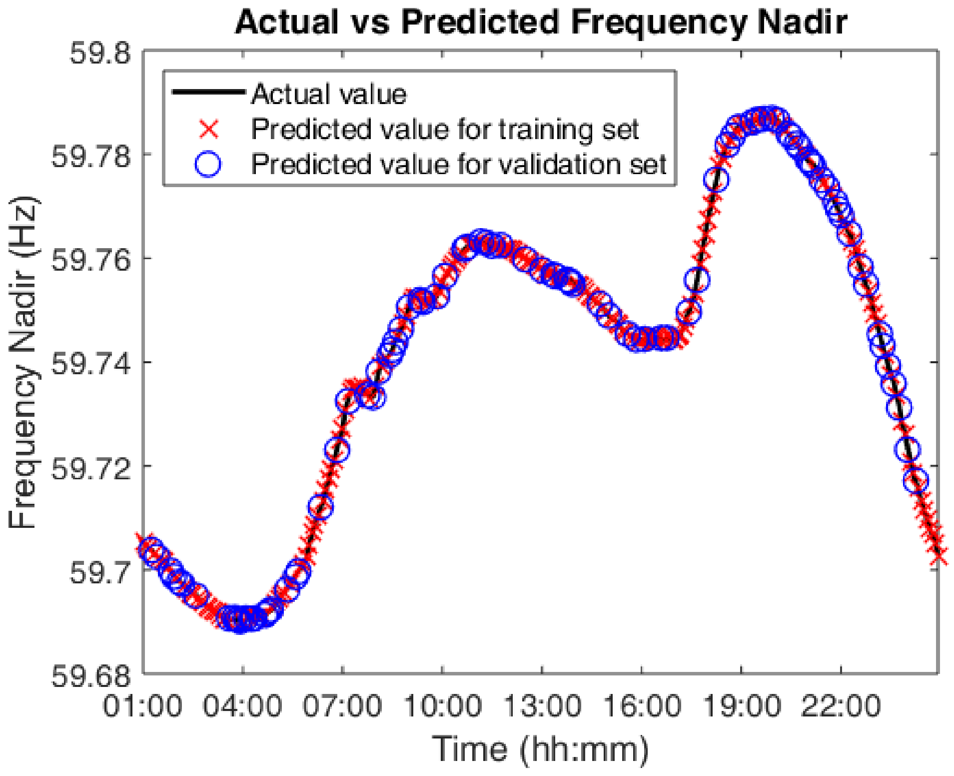
- Angle stability
  - Rotor angle stability refers to the ability of the synchronous machine of an interconnected power system to keep synchronism after being subjected to a disturbances. The maximum allowable value of the fault-clearing time for the system to remain stable are known as critical clearing time (CCT). **A larger CCT value generally indicates higher angle stability margin.**
- Frequency stability
  - Frequency stability refers to the ability of a power system to maintain a steady frequency following a severe system upset resulting in an imbalance between generation and load. After the largest generation loss contingency, the frequency nadir should be maintained above a certain level, for example, the under-frequency-load-shedding threshold. **A higher frequency nadir generally indicates better frequency stability.**
- Small-signal stability
  - Small-signal stability is about the stability of the power system when subject to small disturbances. **A larger damping ratio is more stable.**

Machine-learning-based Angle and Frequency Stability Assessment

CCT prediction



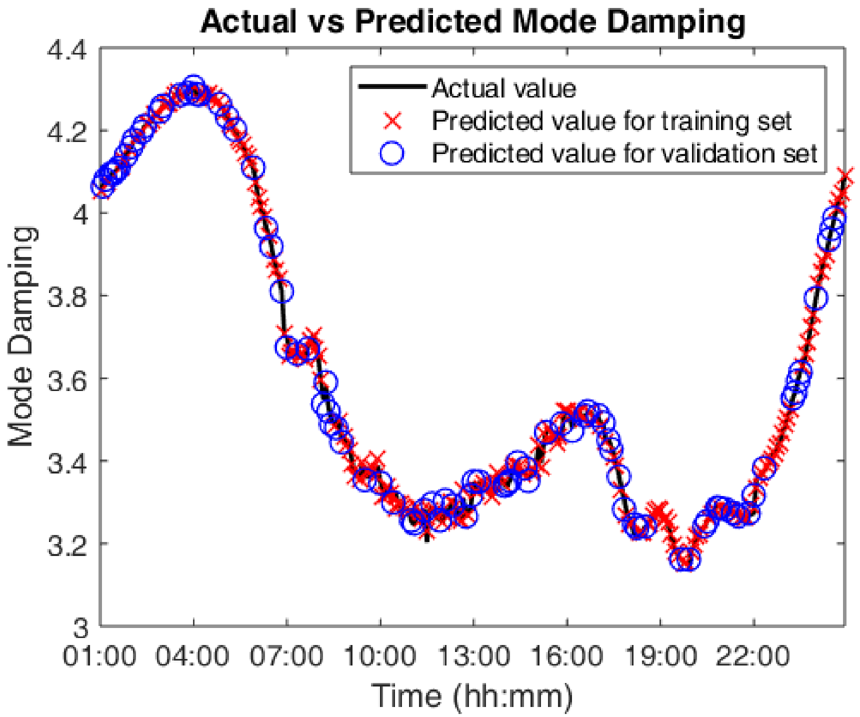
Frequency nadir prediction



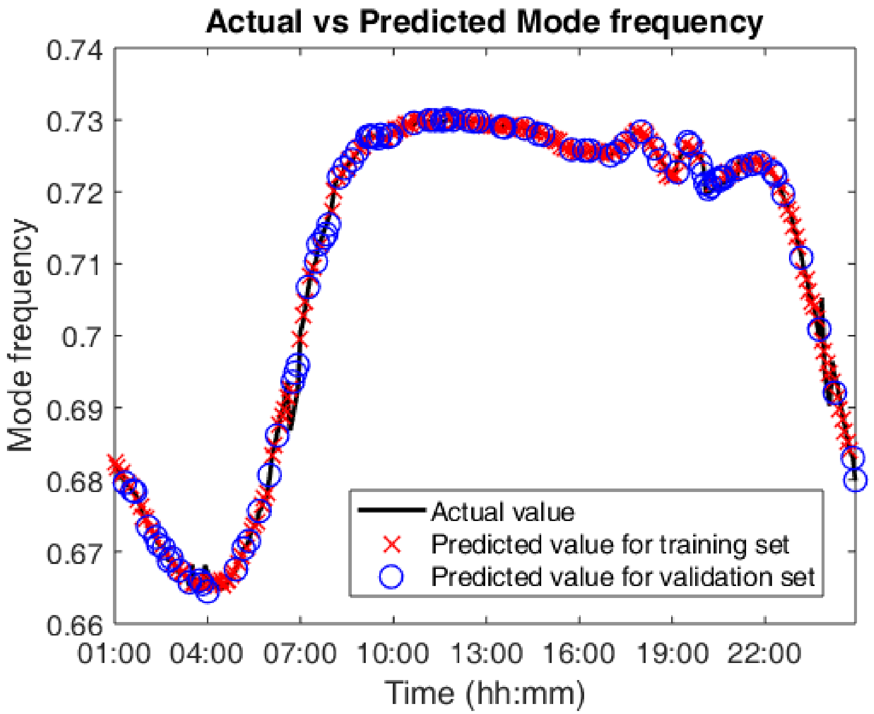
Machine-learning-based Small-signal Stability Assessment

Oscillation damping and oscillation frequency prediction

Damping ratio

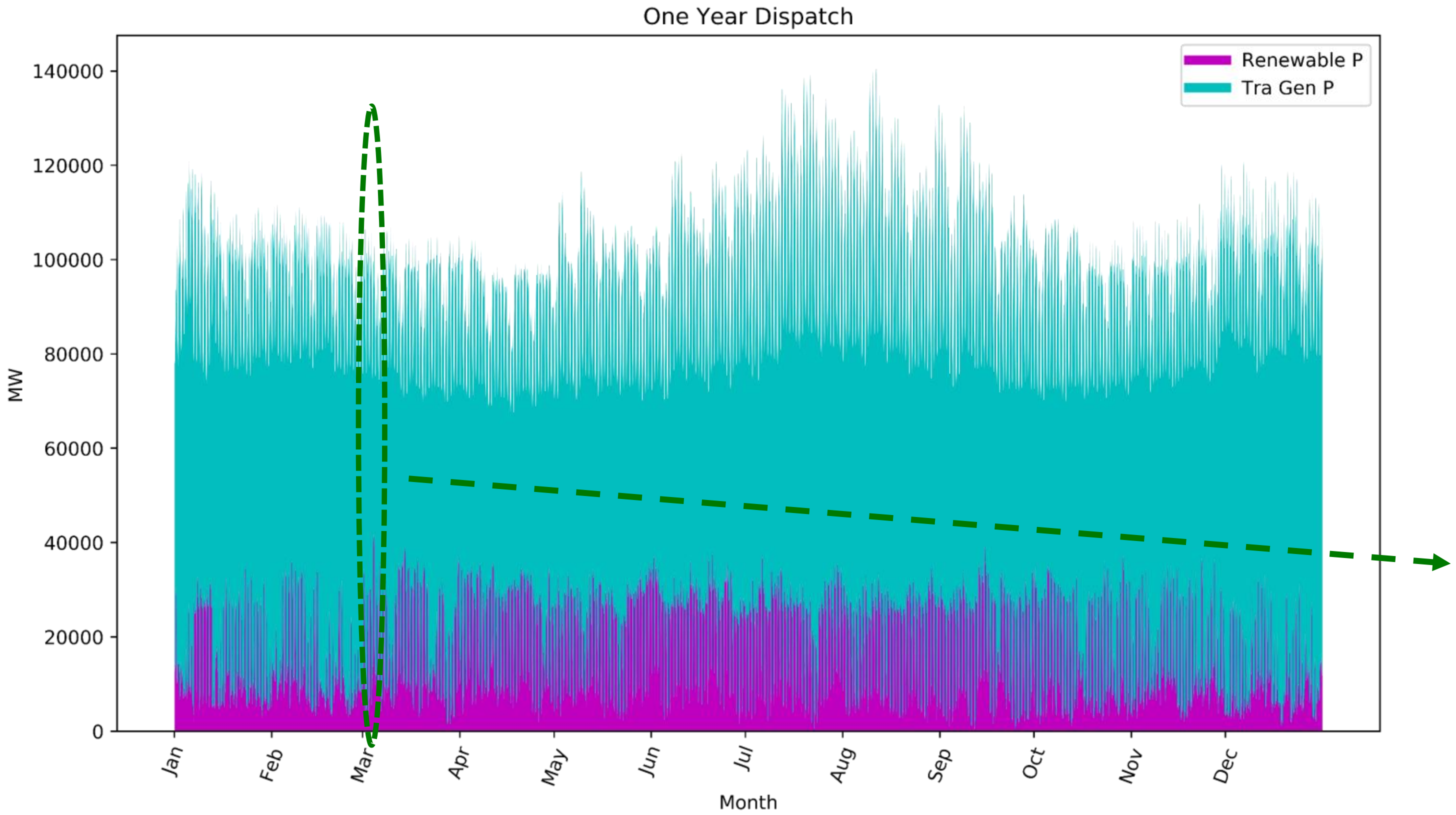


Frequency

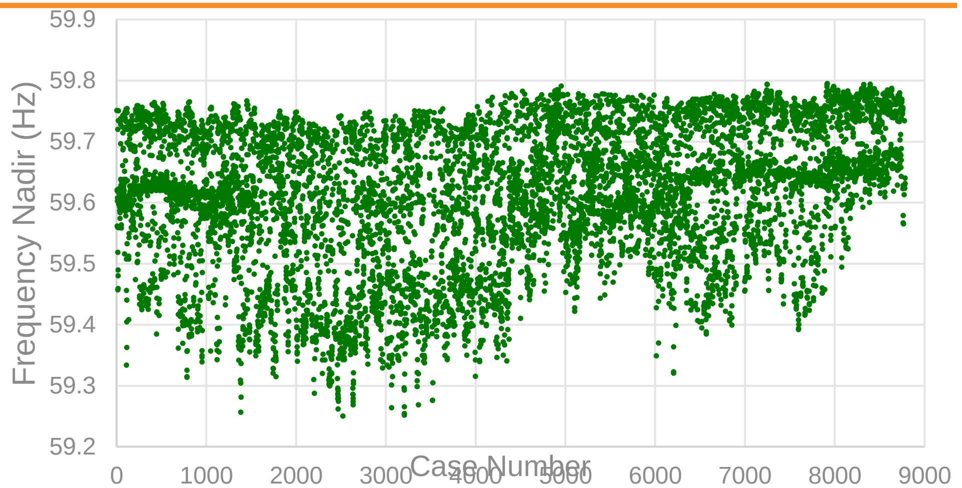


240 bus reduced WECC system results

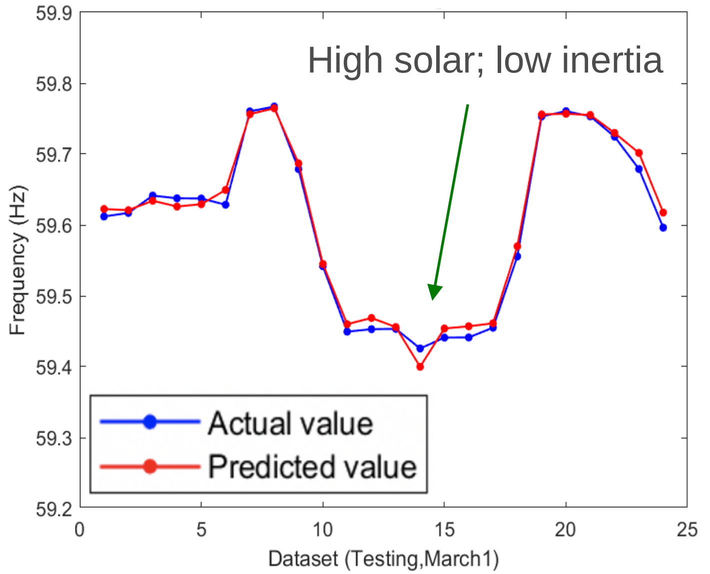
- Frequency stability (frequency nadir) prediction



One year (8000+ power flow scenarios) for 240-bus reduced WECC system



Frequency nadirs for the largest generation loss event

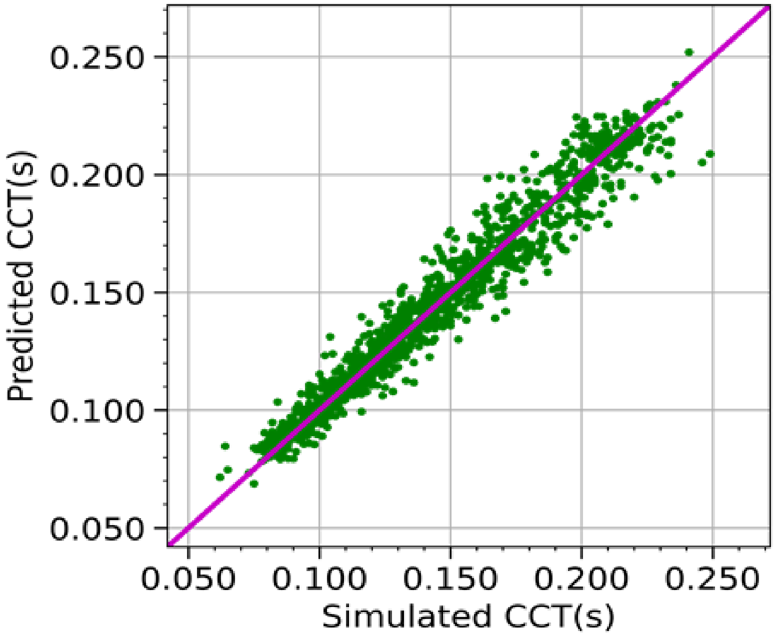


Predict the frequency nadir for the largest generation loss event (March 1)

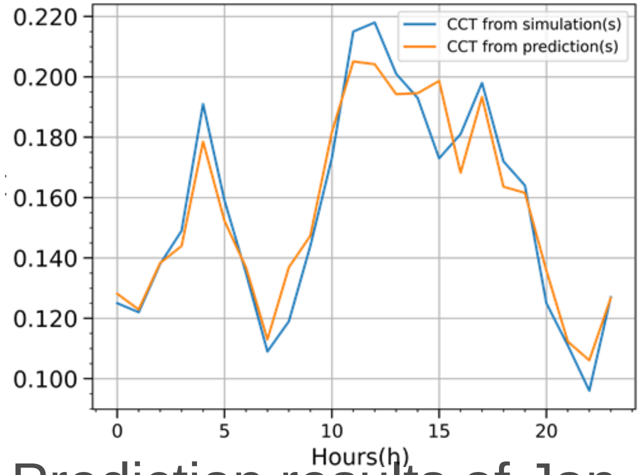
71

240 bus reduced WECC system results

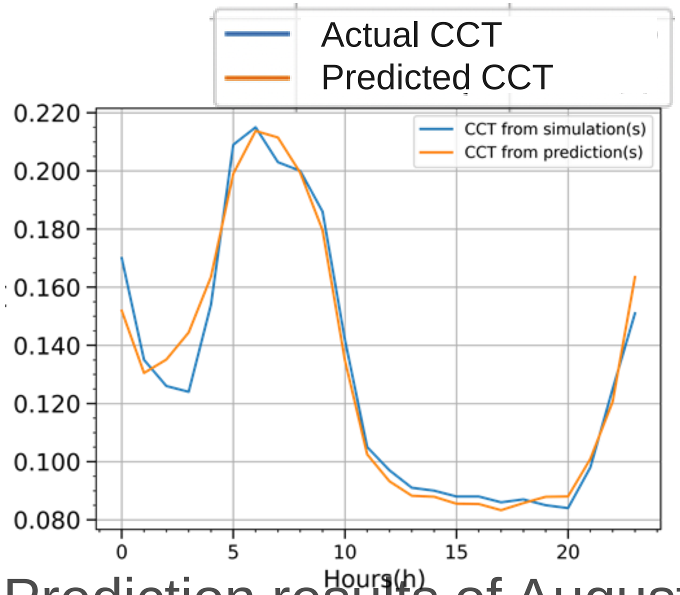
- Angle stability (CCT) prediction
  - Mean absolute error: 0.006s



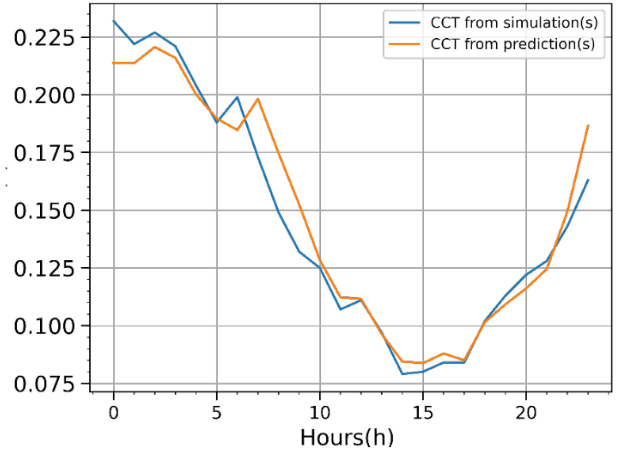
CCT prediction results for multiple days



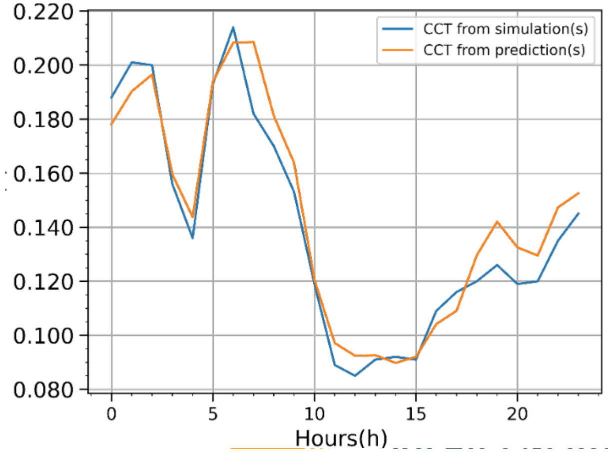
Prediction results of Jan. 20



Prediction results of August. 20

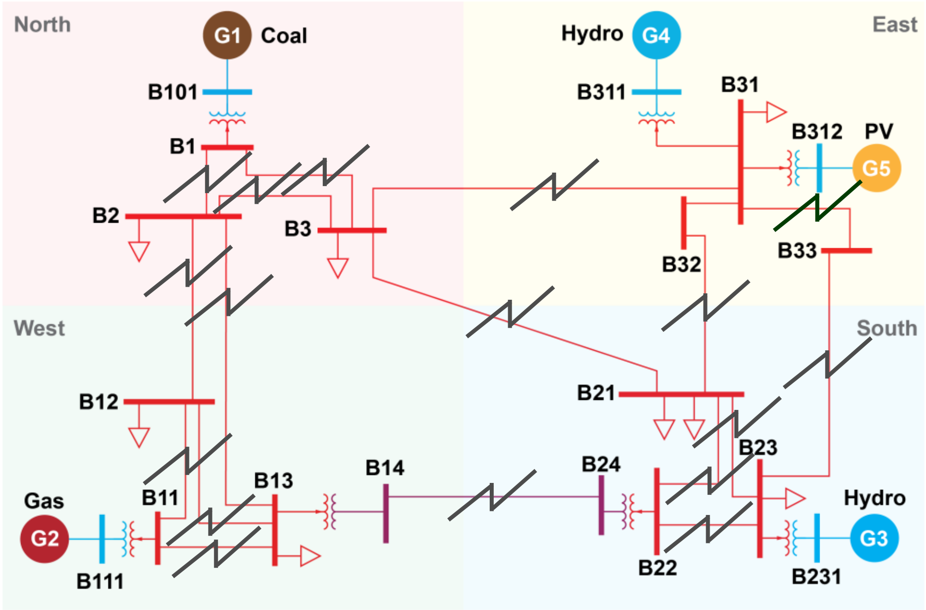


Prediction results of Sept. 9

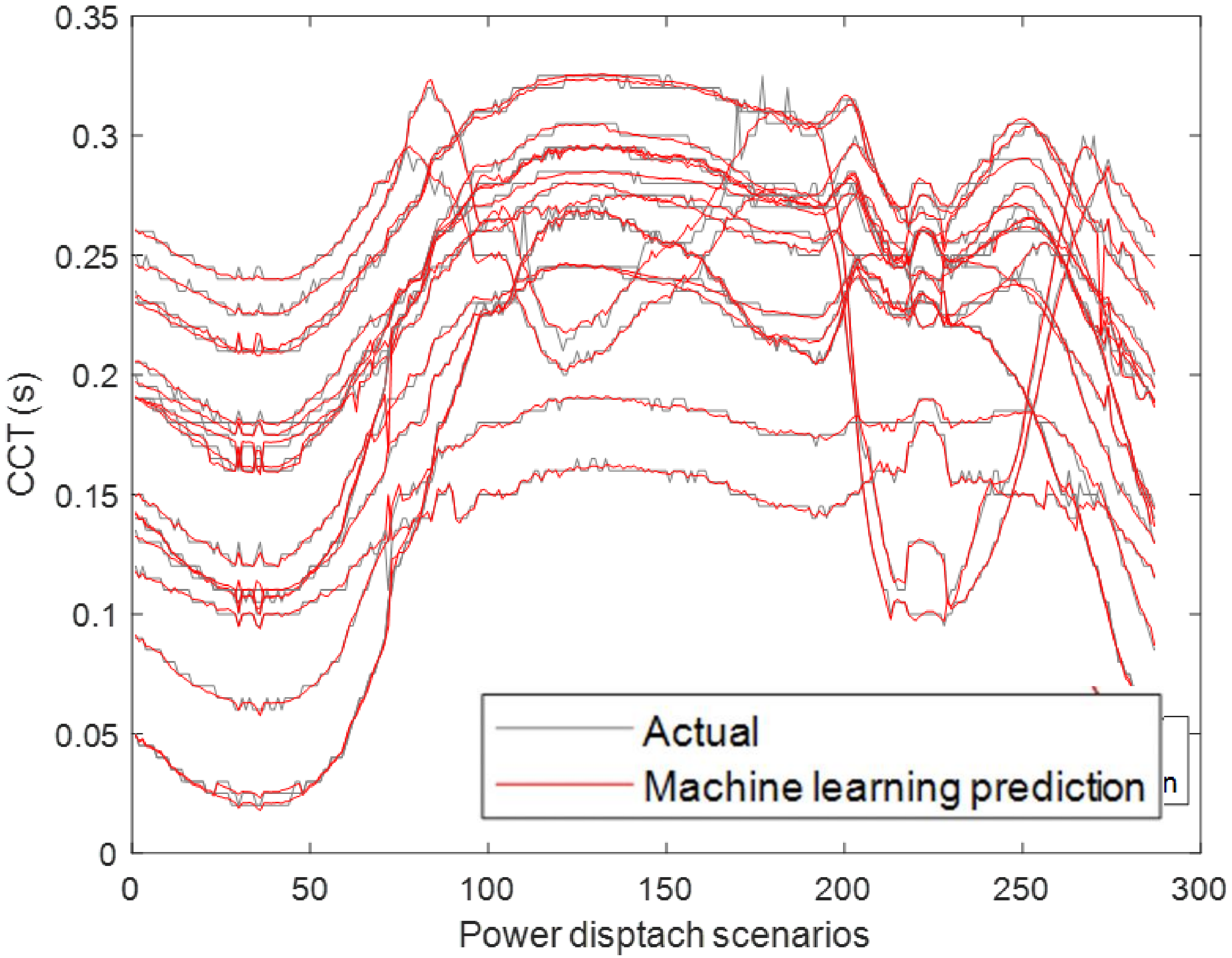


Prediction results of August. 5

Angle Stability Prediction Considering Topology Changes



N-1 line outages in the 18-bus system
Each line represents one N-1 scenario



AI-based CCT prediction considering topology change

ML Accuracy and computation time comparison

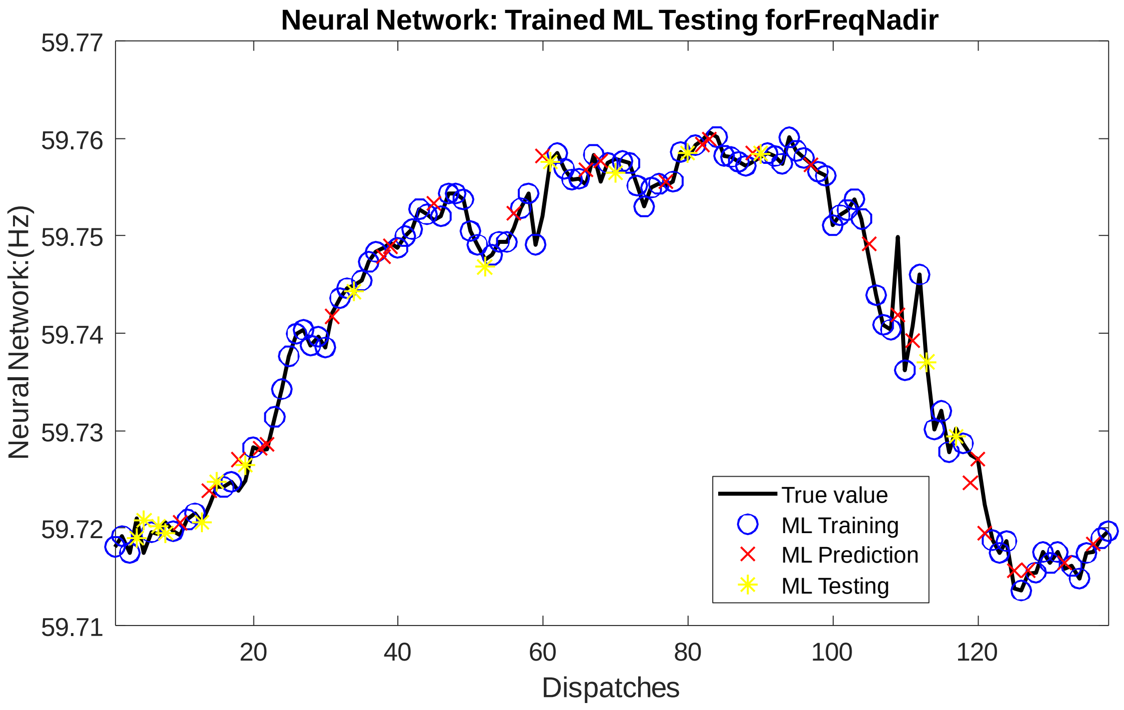
- Using the ML trained models, stability assessment time decreases significantly
- Minimal compromise on accuracy.

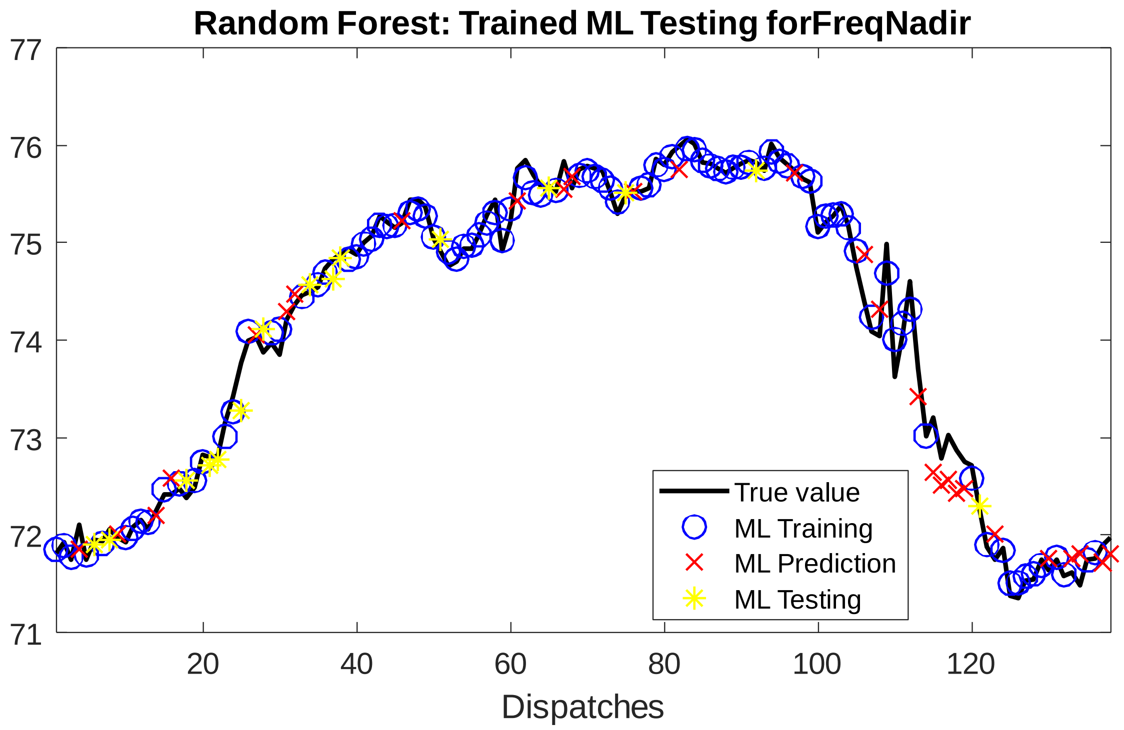
Accuracy of machine learning based stability assessment

<table>
<tr><th rowspan="2">Stability</th><th colspan="2">Estimation accuracy</th><th colspan="2">Time for stability assessment (86 dispatch scenarios)</th></tr>
<tr><th>Random forests</th><th>Neural network</th><th>Time domain simulation</th><th>Machine learning based</th></tr>
<tr><td>Frequency</td><td>98.30%</td><td>99.72%</td><td>~1 h</td><td rowspan="3">~0.18 ms (with trained model)</td></tr>
<tr><td>Angle</td><td>98.44%</td><td>99.29%</td><td>~16 h</td></tr>
<tr><td>Small-Signal</td><td>98.61%</td><td>98.59%</td><td>~1 h</td></tr>
</table>

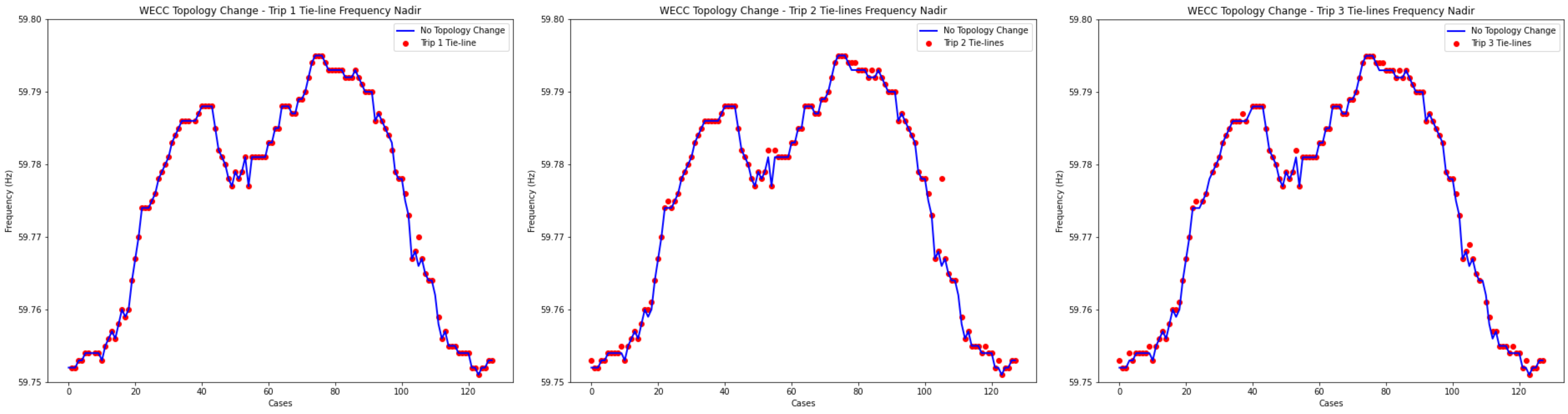
AI Agent for Frequency Stability Assessment: Results on Full WECC System

- Predicted frequency nadir using the developed AI agent



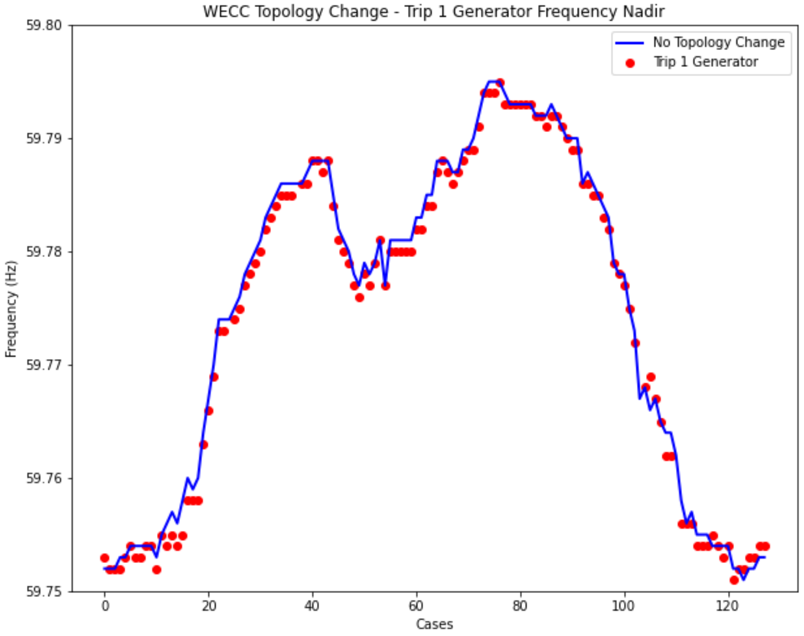


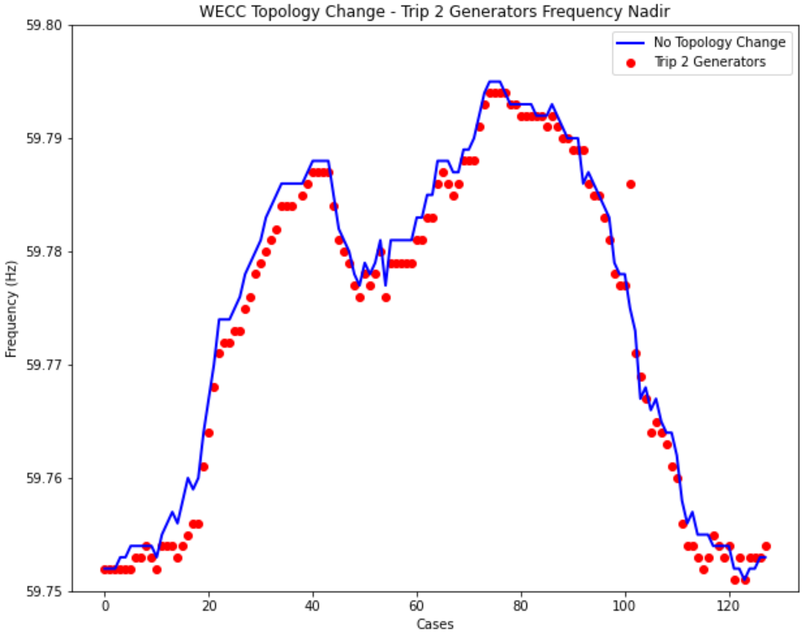
Topology Change on WECC Systems – Trip Tie-lines

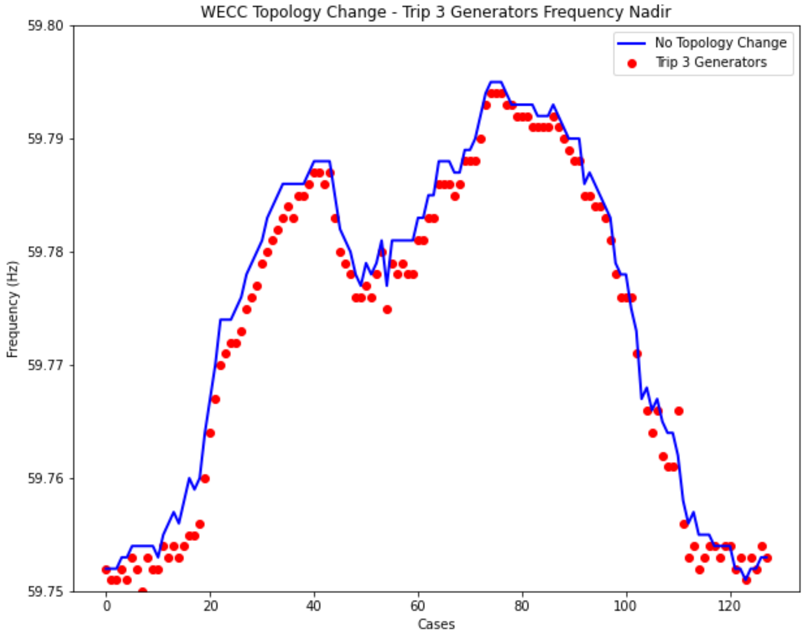


- Trip tie-lines did not lose any generators
- Total generation MW no obvious change resulted in no obvious change in frequency nadir

Topology Change on WECC Systems – Trip Generators



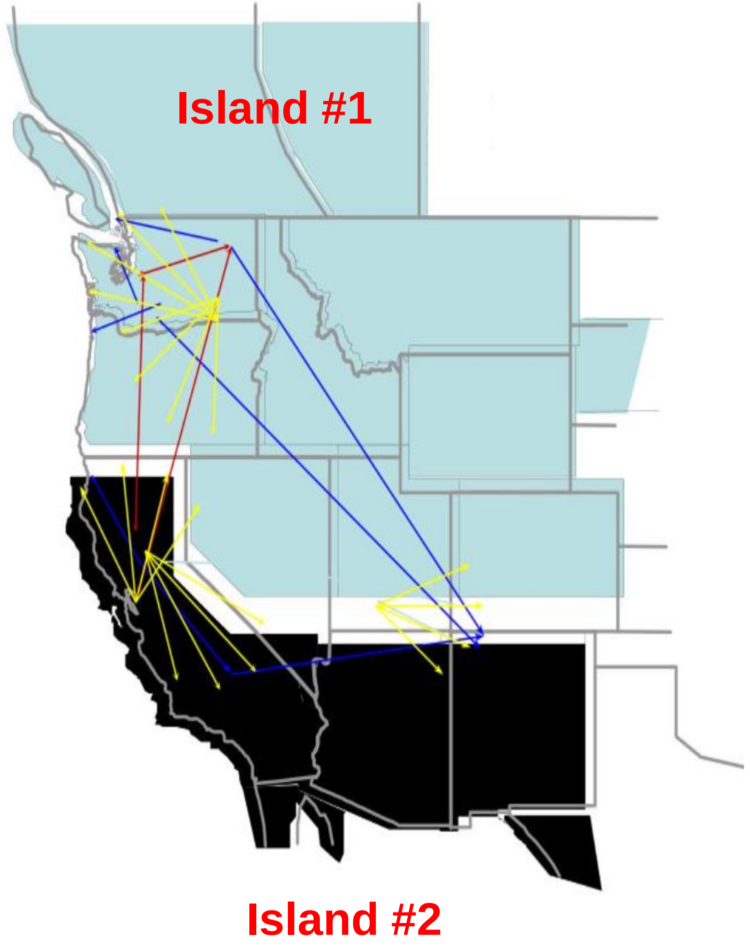




- Trip generators caused total generation MW decrease
- Lower frequency nadir were observed as more generators were lost

AI-based MW Imbalance Prediction after WECC-1 RAS Action

- Study system: The reduced WECC system model
  - ❑ 8,000+ power flow scenarios
  - ❑ Developed by NREL, MSU.
- WECC-1 RAS (Remedial Action Scheme)
  - ❑ Monitor 500kV transmission system within California, Origen, and Washington, etc.
  - ❑ When certain criterion is met, e.g., loss three important tie-lines, it leads to a controlled separation of the WECC system into two islands.
- AI-based method is used to suggest min MW adjustment in the two islands to keep frequency within range.

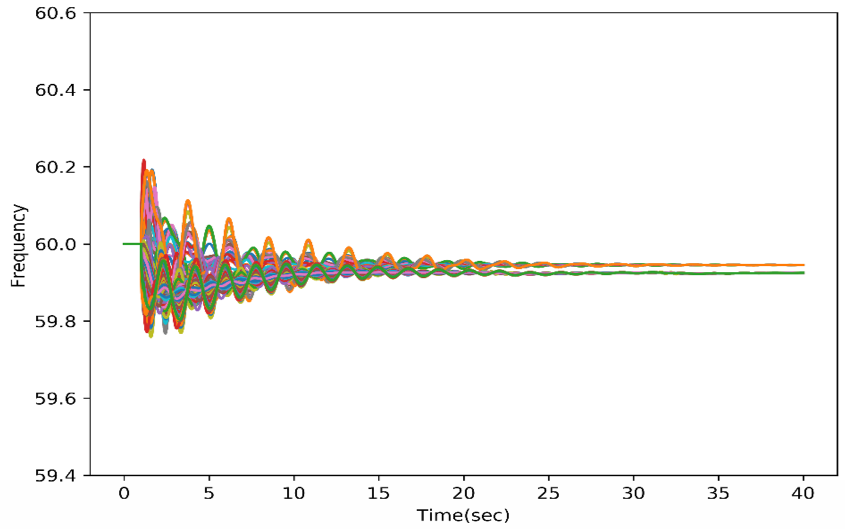


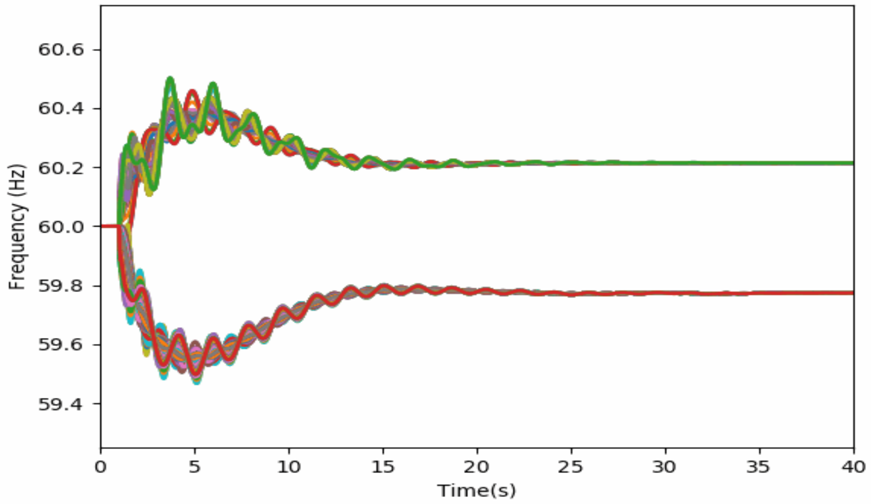
CURENT

AI-based MW Imbalance Prediction after WECC-1 RAS Action

- **Training dataset generation**:
  - ❑ Create the separation for each of the 8000+ dispatch for one year.
  - ❑ Start increasing MW in one area and decreasing MW in the other area based on the tie lines flow
  - ❑ Iteratively tune the MW amount according to the frequency requirement
  - ❑ Obtain the optimal load increase amount and load decrease amount to maintain within 59.5Hz and 60.5Hz range.

Note here load increase is actually generation drop.





AI-based MW Imbalance Prediction after WECC-1 RAS Action

- Neural network to predict optimal MW increase or decrease in two islanded systems after RAS to maintain frequency (= 59.5Hz one area and = 60.5Hz in the other area).
- Training dataset (68%) and test data set (20%)

Total generation (1)
Total load power (1)
Total system inertia (1)
Generator power output (146)
Load power (139)
Generator's inertia contribution (146)

→

Neural Network
(6 hidden layers
TensorFlow Library)

→

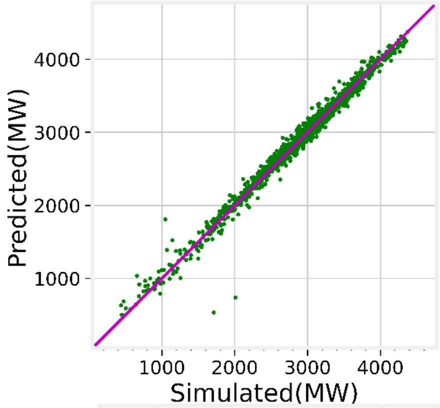
Gen decrease (1)
Load Decrease (1)

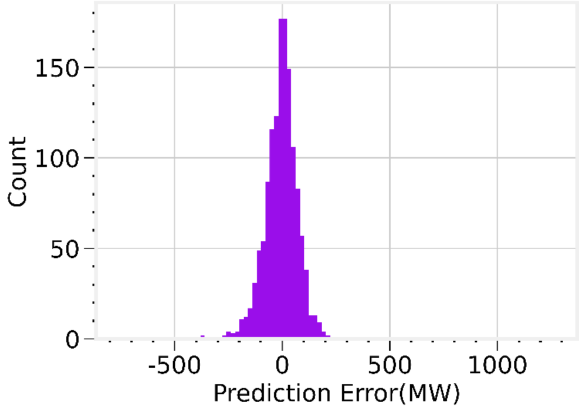
Input and output of Neural Network

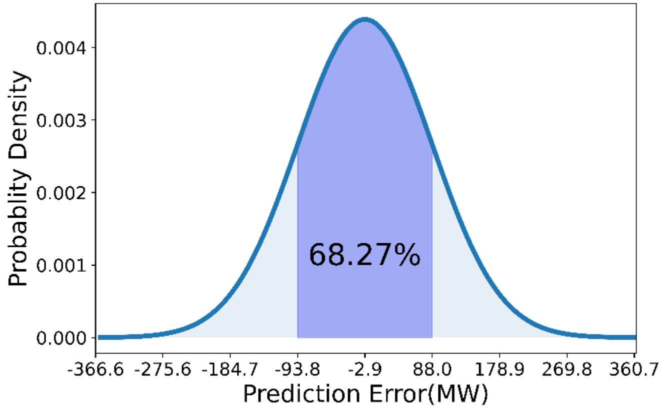
AI-based MW Imbalance Prediction after WECC-1 RAS Action

- Testing results (1350 samples) of the trained model

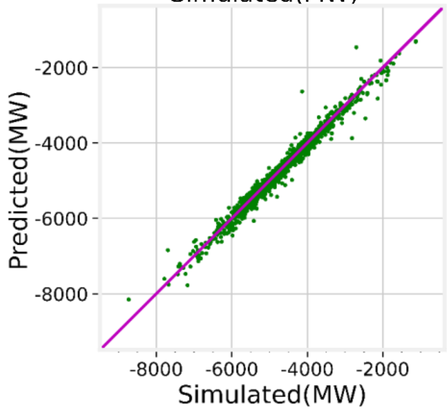
Load Increase

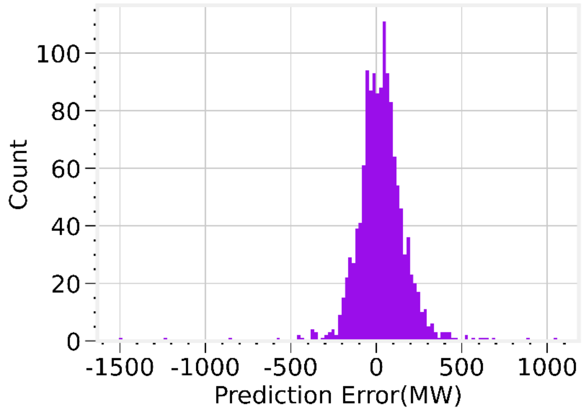


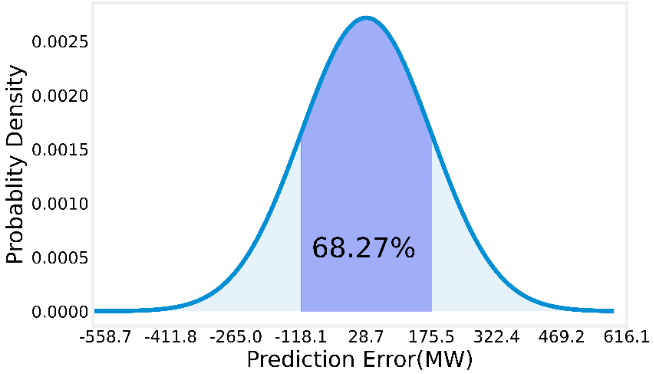




Load Decrease







	MAE (Mean Absolute Error)	RMSE (Root of Mean Squared Error)
Load Increase	58.765 MW	90.900 MW
Load Decrease	99.856 MW	146.839 MW

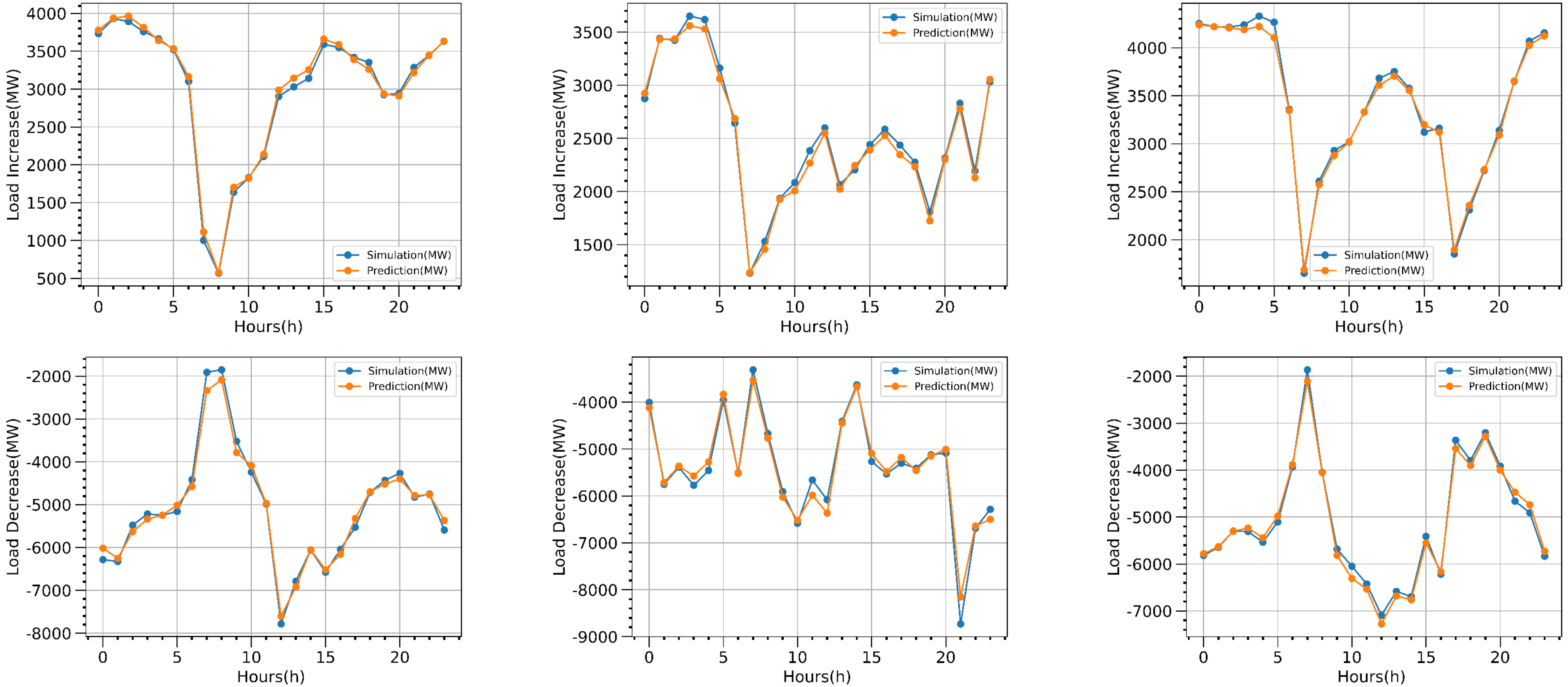
CURENT

AI-based MW Imbalance Prediction after WECC-1 RAS Action

- Daily results confirm that the model can predict the gen decrease and load decrease amount accurately.

Gen decrease

Load Decrease



Apr. 6th **May 27th** **Oct 5th**

Next Step

- **Pilot test at utilities and ISOs**

Looking for partner utilities to start a pilot implementation project in 2021. DOE will provide lab support under the Virtual Operator Assistance Program (VOA). Contact us if interested.