DE-FE0031763: Deep Analysis Net with Causal Embedding for Coal Fired Power Plant Fault Detection and Diagnosis (DANCE4CFDD)

PI: Feng Xue

May 18, 2021

Acknowledgement:
This material is based upon work supported by the Department of Energy, National Energy Technology Laboratory under Award Number DE-FE0031763

Outline

- Project overview
- Progress update
  - Algorithm and system development
  - Experiments with public datasets
  - Experiments with plant data
- Next steps

Project objectives

Develop deep analysis net with causal embedding for coal-fired power plant fault detection and diagnosis (DANCE4CFDD), a novel end-to-end trainable artificial intelligence (AI)-based multivariate time series learning system for flexible and scalable coal power plant fault detection and root cause analysis

Program Team

GE Research

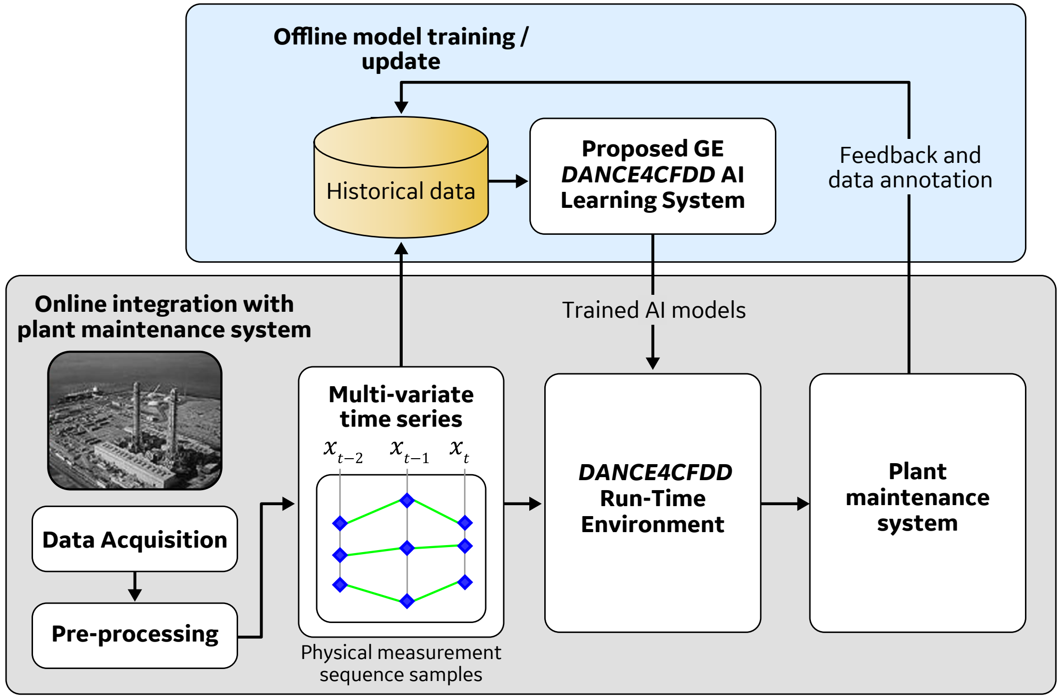
Decades of Developing AI and Machine Learning Technologies for Broad Range of Asset Health Management

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Coal-Fired Power Plant Physical Asset Domain Expertise; Boiler Section Experts; Data Management

Southern Company

Coal-Fired Power Plant Historical Data, Unit Simulator, Subject Matter Expertise, and Host Site for Technology Demonstration



Expected outcomes

- Achieve TRL 5 technology maturity of end-to-end trainable AI learning system for fault detection and root cause analysis
- Validate AI learning system with data from a coal-fired power plant
- Demonstrate advantages comparing with state-of-the-art technologies
- Publicize anonymized training data

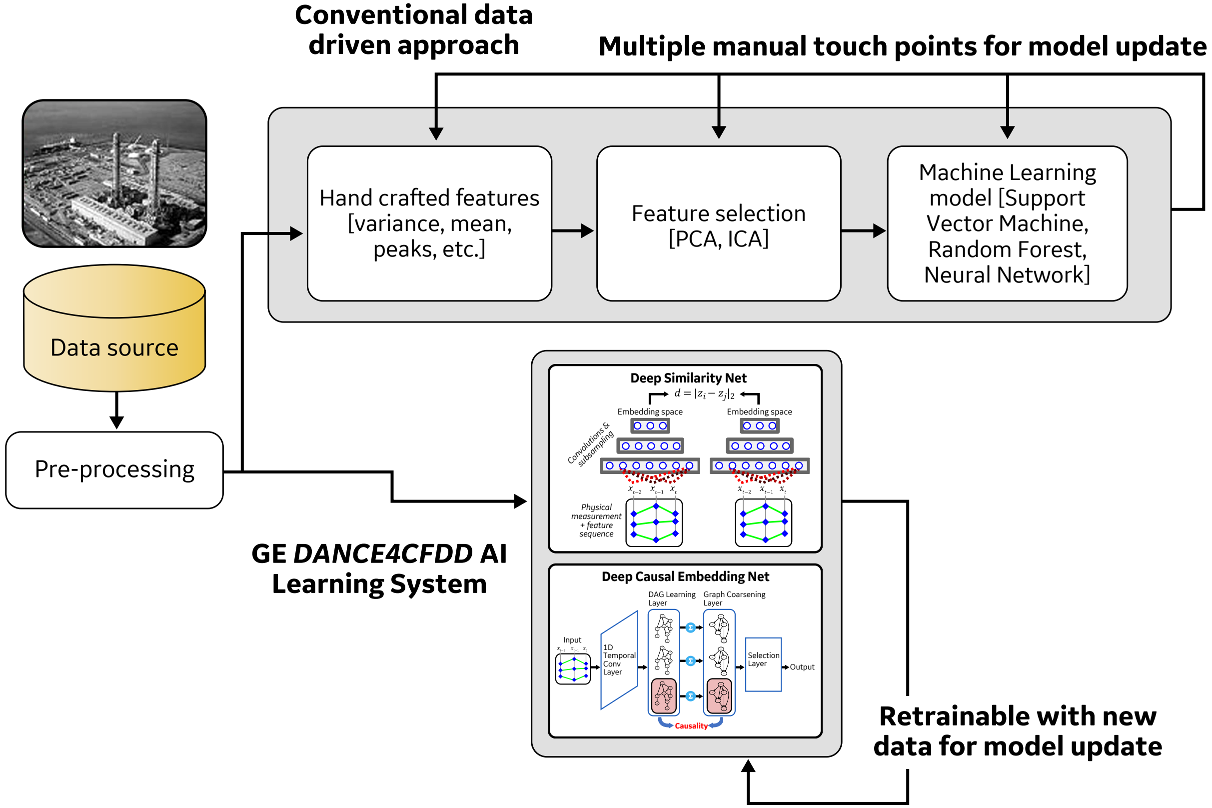
Project objectives

Motivation: Existing asset health management solutions have limited adaptability. Therefore, there is a critical need for a well-designed end-to-end solution that is applicable to a wide range of applications by leveraging large amounts of historical normal operation data from existing coal plants.

Technology Innovation: a novel end-to-end trainable artificial intelligence (AI)-based multivariate time series learning system for flexible and scalable coal power plant fault detection and root cause analysis.

Anticipated benefits:

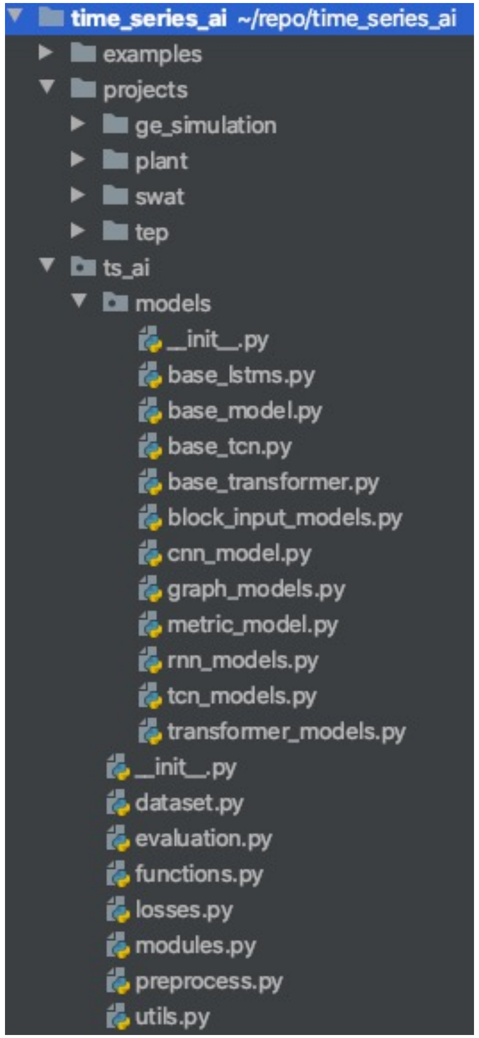
- Applicability to a broad range of asset types and plant configurations for improving coal-fired power plant reliability
- High scalability—reduce development time by eliminating the need for manual and time-consuming domain expert feature engineering



Progress summary

- Streamlined AI anomaly detection – end to end trainable
- Approach for few faulty case learning: how to effectively leverage a small sample of faulty data
- Anomaly interpretation for diagnosis
- Application to a range of plant use cases

Current state of DANCE4CFDD system



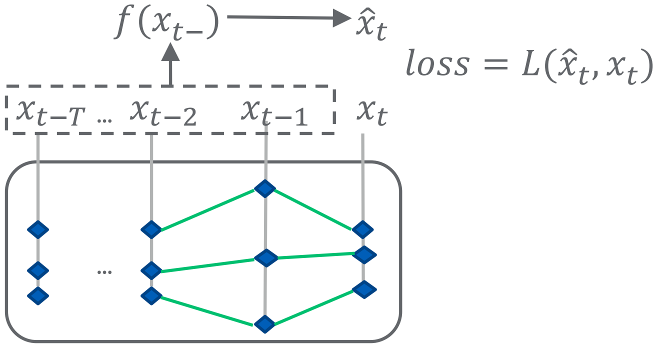
Carry research studies:
- Simulation data study
- Plant data study
- Secure Water Treatment (SWaT) dataset
- Tennessee Eastman Process (TEP) benchmark dataset

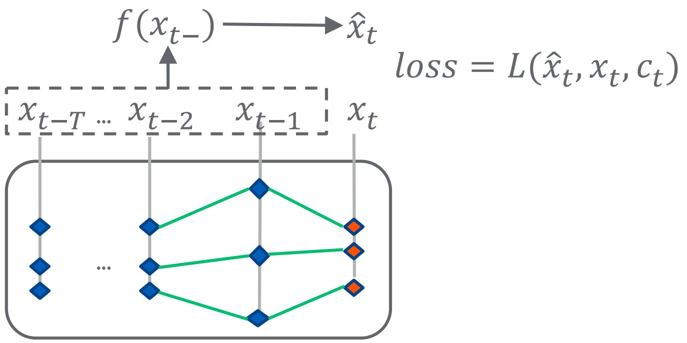
Core model libraries

Supporting functions:
- Time series data construction for training and testing
- Performance evaluation functions
- Other utility functions

Research focus

Industry needs	**Approaches**
• Majority of data is under normal operation, only small number of faulty events to learn from	• Unsupervised learning: only normal operation data is used
	• Supervised learning: both normal operation data and faulty data are available • Few faulty case learning: normal operation data with a small sample of faulty data



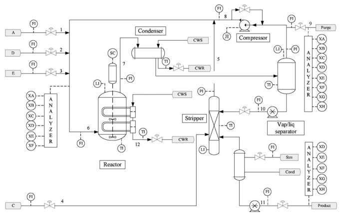


Datasets

- Two public datasets:
  - Secure Water Treatment (SWaT) testbed dataset*

- 7 days of normal operation
- 4 days of 41 episodes of attack
- 36 attacks are physical, treated as faults
- 1/second sampling
- 24 sensor + 27 actuator tags

  - Tennessee Eastman Process (TEP) dataset ^



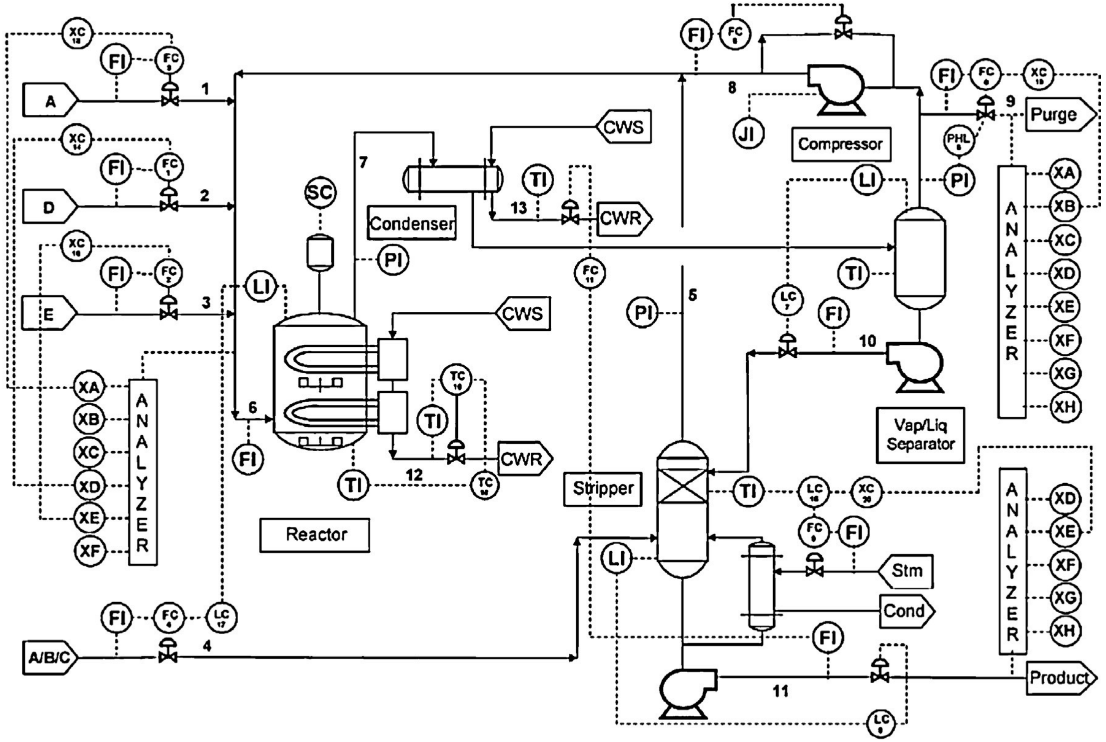
- Simulation data
- 41 sensor + 11 actuator tags
- 20 faults

- Southern Plant Barry Unit 5 data

- 16 months, 3994 tags, 1/min time series data
- NERC GADS report data (reliability events)

* https://itrust.sutd.edu.sg/itrust-labs_datasets/dataset_info/
^ https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/6C3JR1

Brief on TEP dataset



Variable	Description	Variable	Description
XMEAS(1)	A feed (stream 1)	XMEAS(22)	Condenser cooling water outlet temp
XMEAS(2)	D feed (stream 2)	XMEAS(23)	Composition of A (stream 6)
XMEAS(3)	E feed (stream 3)	XMEAS(24)	Composition of B (stream 6)
XMEAS(4)	Total feed (stream 4)	XMEAS(25)	Composition of C (stream 6)
XMEAS(5)	Recycle flow (stream 8)	XMEAS(26)	Composition of D (stream 6)
XMEAS(6)	Reactor feed rate (stream 6)	XMEAS(27)	Composition of E (stream 6)
XMEAS(7)	Reactor pressure	XMEAS(28)	Composition of F (stream 6)
XMEAS(8)	Reactor level	XMEAS(29)	Composition of A (stream 9)
XMEAS(9)	Reactor temp	XMEAS(30)	Composition of B (stream 9)
XMEAS(10)	Purge rate (stream 9)	XMEAS(31)	Composition of C (stream 9)
XMEAS(11)	Separator temp	XMEAS(32)	Composition of D (stream 9)
XMEAS(12)	Separator level	XMEAS(33)	Composition of E (stream 9)
XMEAS(13)	Separator pressure	XMEAS(34)	Composition of F (stream 9)
XMEAS(14)	Separator underflow (stream 10)	XMEAS(35)	Composition of G (stream 9)
XMEAS(15)	Stripper level	XMEAS(36)	Composition of H (stream 9)
XMEAS(16)	Stripper pressure	XMEAS(37)	Composition of D (stream 11)
XMEAS(17)	Stripper underflow (stream 11)	XMEAS(38)	Composition of E (stream 11)
XMEAS(18)	Stripper temperature	XMEAS(39)	Composition of F (stream 11)
XMEAS(19)	Stripper steam flow	XMEAS(40)	Composition of G (stream 11)
XMEAS(20)	Compressor work	XMEAS(41)	Composition of H (stream 11)
XMEAS(21)	Reactor cooling water outlet temp		

Variable	Description	Variable	Description
XMV(1)	D feed flow (stream 2)	XMV(7)	Separator pot liquid flow (stream 10)
XMV(2)	E feed flow (stream 3)	XMV(8)	Stripper liquid product flow
XMV(3)	A feed flow (stream 1)	XMV(9)	Stripper steam valve
XMV(4)	Total feed flow (stream 4)	XMV(10)	Reactor cooling water flow
XMV(5)	Compressor recycle valve	XMV(11)	Condenser cooling water flow
XMV(6)	Purge valve (stream 9)	~~XMV(12)~~	~~Agitator speed~~

Fault ID	Description	Type	Magnitude
IDV1	A/C Feed ratio, B Composition constant (stream 4)	Step	203%
IDV2	B Composition, A/C Ratio constant (stream 4)	Step	105%
IDV3	D Feed temperature (stream 2)	Step	5%
IDV4	Reactor cooling water inlet temperature	Step	9%
IDV5	Condenser cooling water inlet temperature	Step	15%
IDV6	A Feed loss (stream 1)	Step	342%
IDV7	C Header pressure loss – reduced availability (Stream 4)	Step	25%
IDV8	A, B, C Feed composition (stream 4)	Random variation	736%
IDV9	D Feed temperature (stream 2)	Random variation	8%
IDV10	C Feed temperature (stream 4)	Random variation	112%
IDV11	Reactor cooling water inlet temperature	Random variation	567%
IDV12	Condenser cooling water inlet temperature	Random variation	8%
IDV13	Reaction kinetics	Slow drift	16%
IDV14	Reactor cooling water valve	Sticking	1285%
IDV15	Condenser cooling water valve	Sticking	5%
IDV16	Unknown	Random variation	78%
IDV17	Unknown	Random variation	557%
IDV18	Unknown	Step	57%
IDV19	Unknown	Random variation	73%
IDV20	Unknown	Random variation	310%

Source: H. Chen, P. Tino, X. Yao, Cognitive fault diagnosis in Tennessee Eastman Process using learning in the model space, Computers and Chemical Engineering (67), 2014.

Study summary

Autoregressive formulation

$$f(x_{t-}) \longrightarrow \hat{x}_t$$

$$\varepsilon_t = x_t - \hat{x}_t$$

$$x_{t-T} \dots x_{t-2} \quad x_{t-1} \qquad x_t$$

MSE (Mean Square Error)

MD (Mahalanobis Distance)

Fault	LSTM		Bidirectional-LSTM		TCN		Attention LSTM		Transformer		Transformer	
	MD	MSE	MD	MSE	MD	MSE	MD	MSE	MD	MSE	MD	MSE
1	99.96%	99.74%	99.97%	99.73%	99.86%	99.61%	99.96%	99.72%	99.93%	99.64%	99.92%	99.66%
2	99.09%	97.99%	99.05%	97.98%	98.55%	97.93%	99.02%	97.96%	99.00%	98.45%	98.94%	98.40%
3	5.77%	5.12%	5.65%	5.15%	5.32%	5.12%	5.69%	5.08%	5.38%	5.00%	5.25%	5.03%
4	100.00%	99.95%	100.00%	99.96%	100.00%	100.00%	100.00%	99.96%	100.00%	99.96%	100.00%	99.99%
5	100.00%	20.63%	100.00%	20.66%	100.00%	26.46%	100.00%	46.25%	100.00%	28.86%	100.00%	25.74%
6	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
7	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
8	98.32%	93.42%	98.30%	93.55%	97.65%	94.68%	98.27%	94.40%	98.26%	96.43%	98.23%	95.80%
9	6.25%	5.21%	5.98%	5.22%	5.44%	5.19%	6.09%	5.18%	5.61%	5.19%	5.50%	5.17%
10	91.84%	21.06%	92.19%	18.11%	87.01%	35.57%	90.91%	22.56%	87.72%	17.48%	89.36%	15.51%
11	86.24%	80.43%	86.09%	80.26%	84.39%	80.51%	84.69%	80.44%	83.90%	77.51%	83.00%	76.83%
12	99.37%	95.87%	99.23%	95.71%	99.27%	96.63%	99.35%	96.69%	99.24%	98.20%	99.24%	97.89%
13	95.93%	93.21%	95.91%	93.12%	95.19%	93.48%	95.88%	93.36%	95.71%	94.01%	95.66%	93.66%
14	99.98%	99.96%	99.98%	99.97%	99.98%	99.97%	99.98%	99.96%	99.98%	99.97%	99.98%	99.97%
15	8.30%	5.37%	5.77%	5.35%	6.29%	5.36%	7.52%	5.35%	5.98%	5.39%	5.94%	5.35%
16	91.97%	16.99%	92.66%	14.31%	89.15%	23.10%	90.36%	18.16%	91.35%	13.43%	90.86%	12.46%
17	96.52%	95.46%	96.51%	94.92%	96.41%	96.14%	96.50%	95.92%	95.88%	91.53%	95.98%	90.62%
18	94.69%	93.70%	94.56%	93.73%	94.61%	93.90%	94.61%	93.70%	94.42%	93.76%	94.37%	93.78%
19	87.38%	24.10%	87.37%	24.09%	87.95%	23.39%	85.71%	24.22%	87.51%	25.13%	86.55%	24.40%
20	95.56%	71.11%	95.07%	69.11%	94.95%	47.92%	95.93%	60.51%	90.34%	48.05%	88.12%	46.87%
False Positive	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%
Average (TP)	**82.86%**	65.96%	**82.71%**	65.55%	**82.10%**	66.25%	**82.52%**	66.97%	**82.01%**	64.90%	**81.85%**	64.36%

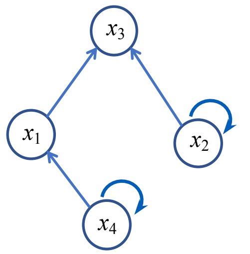
Other formulations

- Developed a new self-supervised formulation for model training with demonstrated performance improvement comparing with autoregressive formulation

- Developed a new approach to leverage a small sample of faulty cases along with normal operation data, demonstrated improved performance using TEP dataset with as little as 3 cases per fault for training

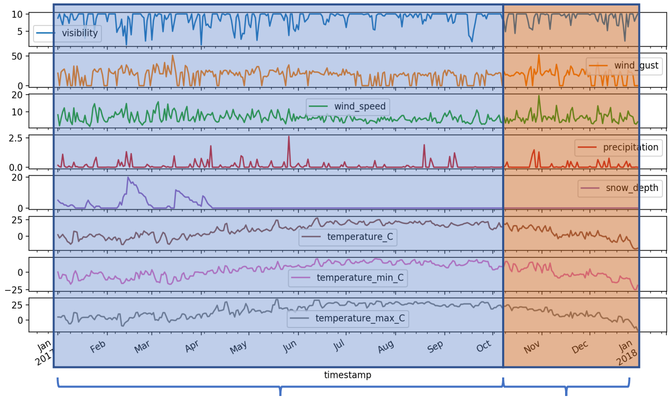
Anomaly detection and interpretation by learning variable dependency

Learn variable dependency graph ***N*** from normal operation data



		effect			
		x_1	x_2	x_3	x_4
cause	x_1	0	0	1	0
	x_2	0	1	1	0
	x_3	0	0	0	0
	x_4	1	0	0	1

Data

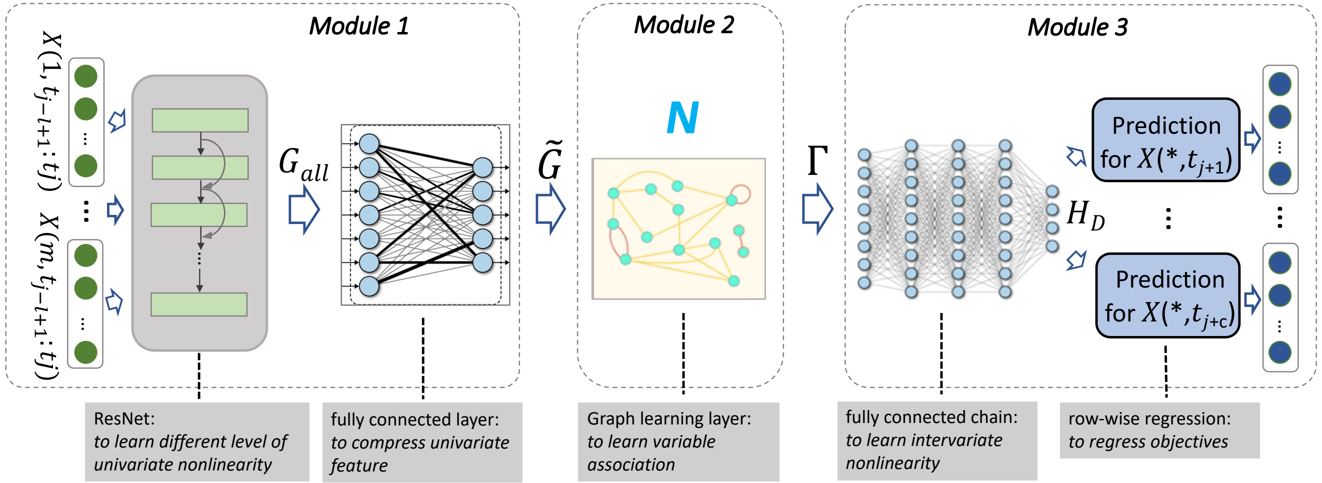


l

input window

w

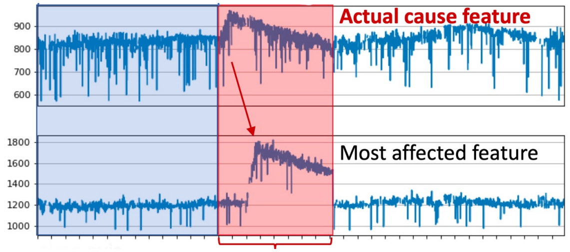
prediction window



Detection and Diagnosis of **A**nomaly with **V**ariable **A**ssociation **C**hange (***DAVAC***)

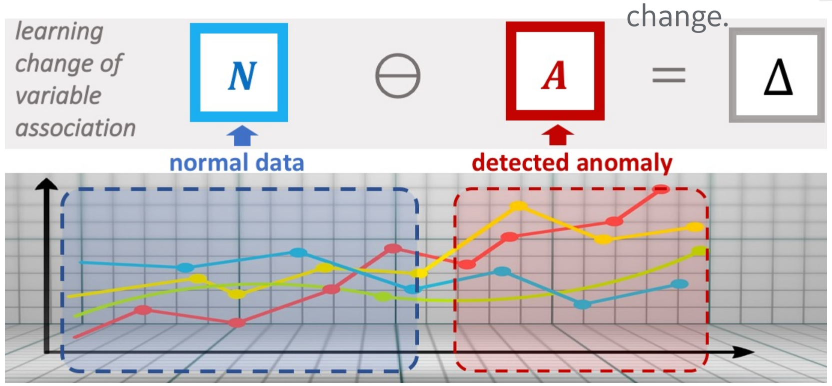
Anomaly interpretation

Anomaly induced association change



Univariate ranking methods usually only return **the most affected features**.

The **actual cause of anomaly** can be obtained from the association change.

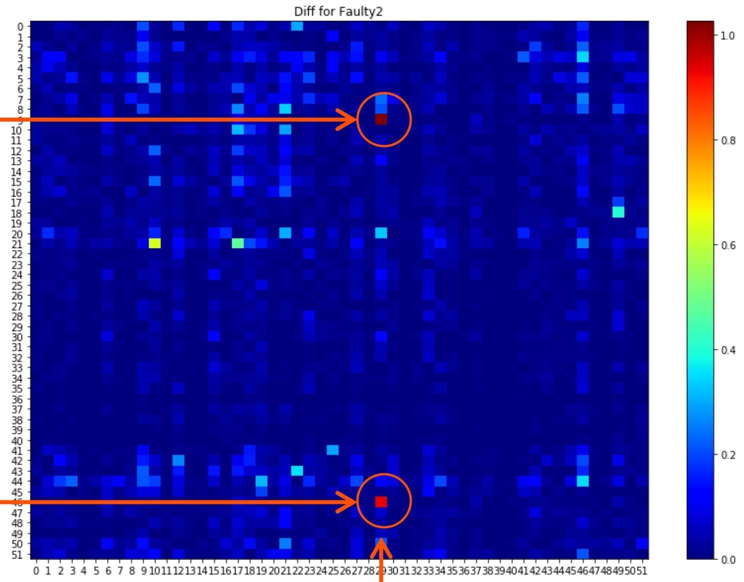


TEP data example

Fault 2: B Composition, A/C Ratio constant (step)

var_9: Purge_rate

var_46: Purge_valve



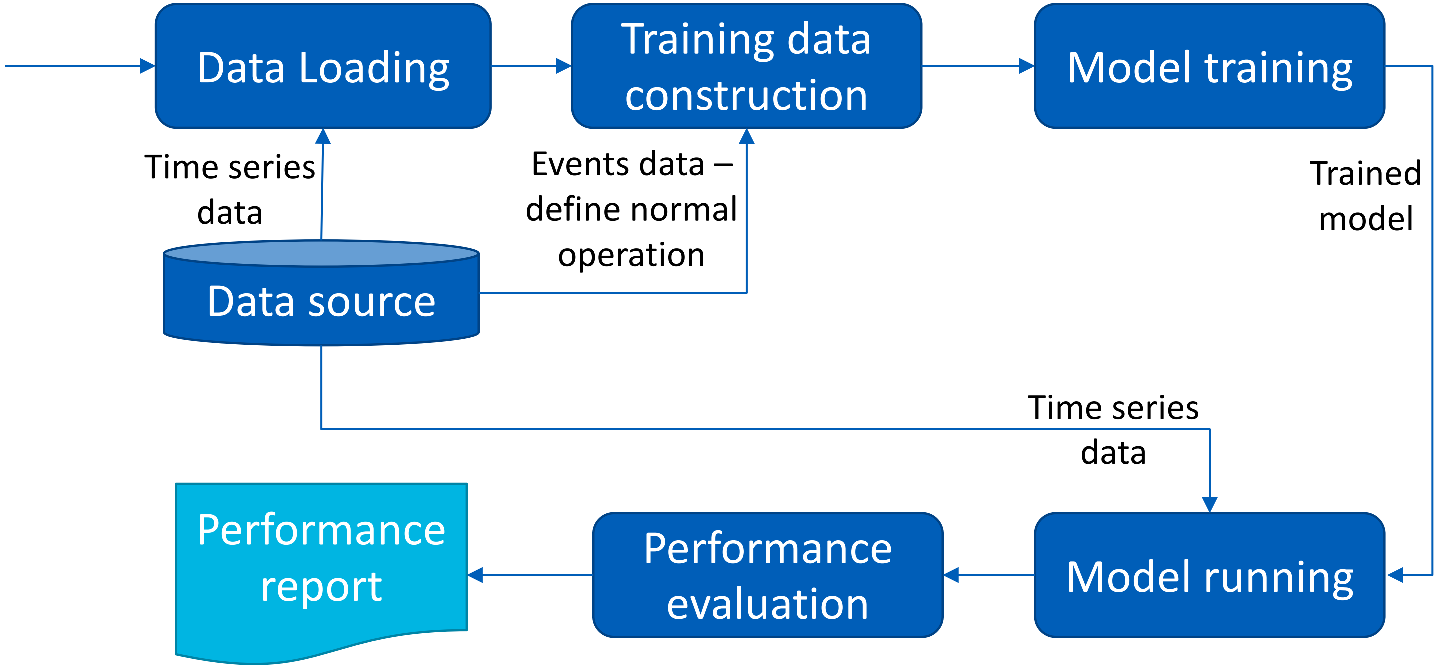
var_29: Composition_of_B_purge

Process steps for plant data

Data tags and configuration

InInput	InTarget	InControl	SourceName
1		1	5-F-FDR-DMD
			5-F-FDR-DSCH-PRS
			5-F-FDR-RATE
1	1		5-F-FDR-SPD
			5-F-MILL-AMP
			5-F-MILL-DP
1	1		5-F-MILL-INT-T
1		1	5-F-MILL-PA-1-F
			5-F-MILL-PA-F
1		1	5-F-MILL-PA-T
1	1		5-F-MILL-T
1		1	5F-FDR-SPD-BIAS

Input: all time series measurements as input to model
Target: tags to predict
Control: command and actuation measurements



Barry unit 5 plant data experiments

Experiment setup: 3 month for training; 13 months for evaluation

Use case / Data tags configuration / Performance evaluation

WET COAL (OMC) –FEEDER STOPPAGE:

InInput	InTarget	InControl	SourceName
1		1	5-F-FDR-DMD
			5-F-FDR-DSCH-PRS
			5-F-FDR-RATE
1	1		5-F-FDR-SPD
			5-F-MILL-AMP
			5-F-MILL-DP
1	1		5-F-MILL-INT-T
1		1	5-F-MILL-PA-1-F
			5-F-MILL-PA-F
1		1	5-F-MILL-PA-T
1	1		5-F-MILL-T
1		1	5F-FDR-SPD-BIAS

System	Total Events	Detected Events (TP)	False Alarms / 13 months (LSTM)	False Alarms / 13 months (CNN)
Mill A	4	4	48	50
Mill B	2	2	29	34
Mill C	9	9	39	46
Mill D	6	6	22	27
Mill E	1	1	10	8
Mill F	6	6	14	22

CONDENSER TUBE CLEANING SYSTEMS INCLUDING DEBRIS FILTER:

CIRCULATING WATER PUMP MOTORS:

InInput	InTarget	InControl	SourceName
1	1		5-DEBRIS-FLT-A-DP
1	1		5-A-CRC-WTR-PMP-MTR-AMP
1	1		5-A-CWP-MTR-LWG-BRG-T
1	1		5-A-CWP-MTR-UPG-BRG-T
1	1		5-A-CWP-MT-THR-BRG-T
1	1		5-A-CWP-MTR-STAT-1-T
1	1		5-A-CWP-MTR-STAT-2-T
1	1		5-A-CWP-MTR-STAT-3-T

System	Total Events	Detected Events (TP)	False Alarms / 13 months (LSTM)	False Alarms / 13 months (CNN)
CWP A	4	4	5	11
CWP B	4	4	3	4

BOILER FEED WATER PUMP:

InInput	InTarget	InControl	SourceName
			5-DEA-LVL
			5-A-BFP-BRG-OIL-P
1	1		5-A-BFP-DISC-P
1		1	5-A-BFP-DMD
1	1		5-A-BFP-F
			5-A-BFP-IB-BRG
			5-A-BFP-KPPH-F
			5-A-BFP-OB-BRG
			5-A-BFP-SEAL-WTR-T
1		1	5-A-BFP-SPD

System	Total Events	Detected Events (TP)	False Alarms / 13 months CNN, balanced	False Alarms / 13 months CNN, unbalanced
BFP A	7	7	11	17

Demonstrated AI learning system on initial plant use cases

Next steps

- Study AI system with simulation data from GE Steam Power
- Plant Barry Unit 5 data gathering for validation preparation
- Usability experiment with GE internal data
- Enhance AI framework maturity

Q & A