OAK RIDGE National Laboratory

Lawrence Livermore National Laboratory

Grid Signature Library (GSL)

Aaron Wilson*, Jamie Lian* and Jhi-Young Joo+

* Oak Ridge National Laboratory

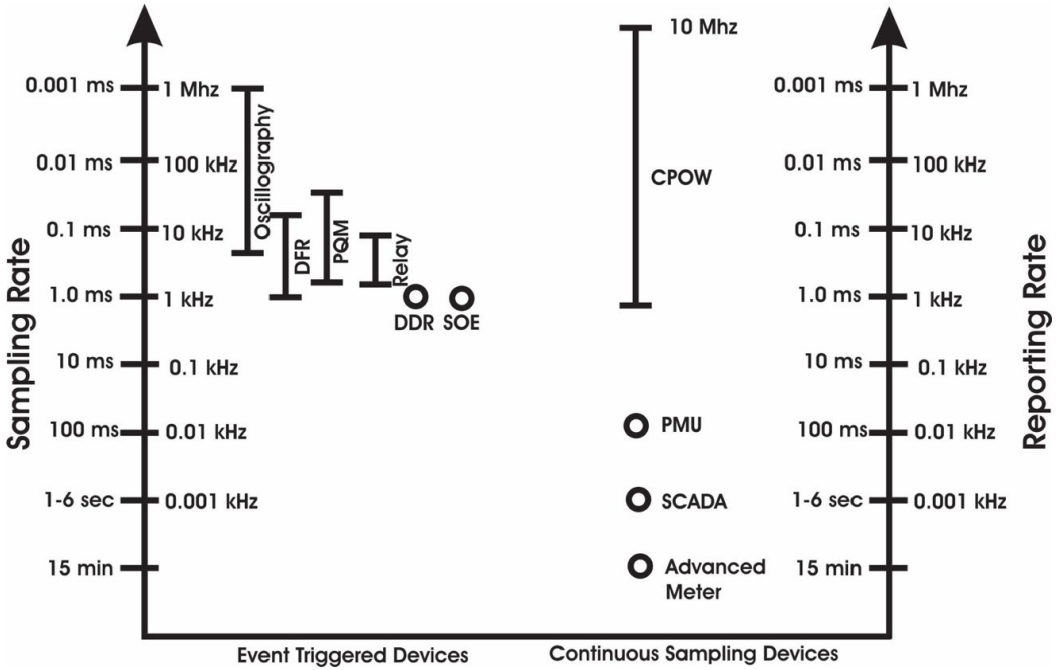
+ Lawrence Livermore National Laboratory

Oct. 19th, 2022

U.S. DEPARTMENT OF ENERGY

Towards more observable grid

- Event triggered measurements
  - Relays
  - Digital fault recorders
  - Power quality meters
- Continuous measurements
  - SCADA
  - AMI (advanced metering infrastructure)
  - PMU (phasor measurement unit)
  - Point-on-wave (POW) measurements
- Event records
  - Outage and maintenance records
  - Device activation records



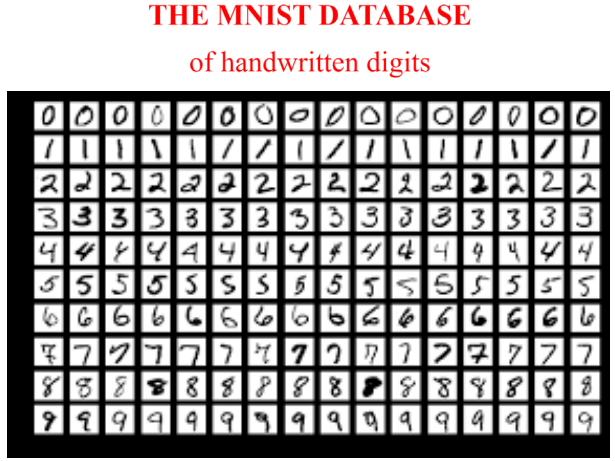
Grid Monitoring devices by resolution and data continuity*

AI/ML-based Grid Health Monitoring

*A. Silverstein and J. Follum, "High-resolution, time-synchronized grid monitoring devices," PNNL, Tech. Rep. PNNL-29770, Mar. 2020.

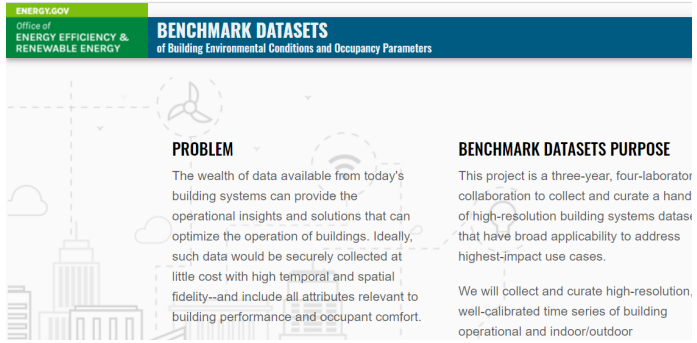
But still one step away

- Data labeling is critical to AI/ML
  - MNIST
  - ImageNet
  - BTO Building Benchmark Datasets
- Challenges exist for grid events
  - Data is decentralized and inaccessible
    - Limits actionable data available for analytics
  - Data is multimodal and unstandardized
    - Prevents integration of different data sources
  - Data is unprocessed and unvalidated
    - Lacks critical metadata and proper labeling



http://yann.lecun.com/exdb/mnist/

https://syncedreview.com/

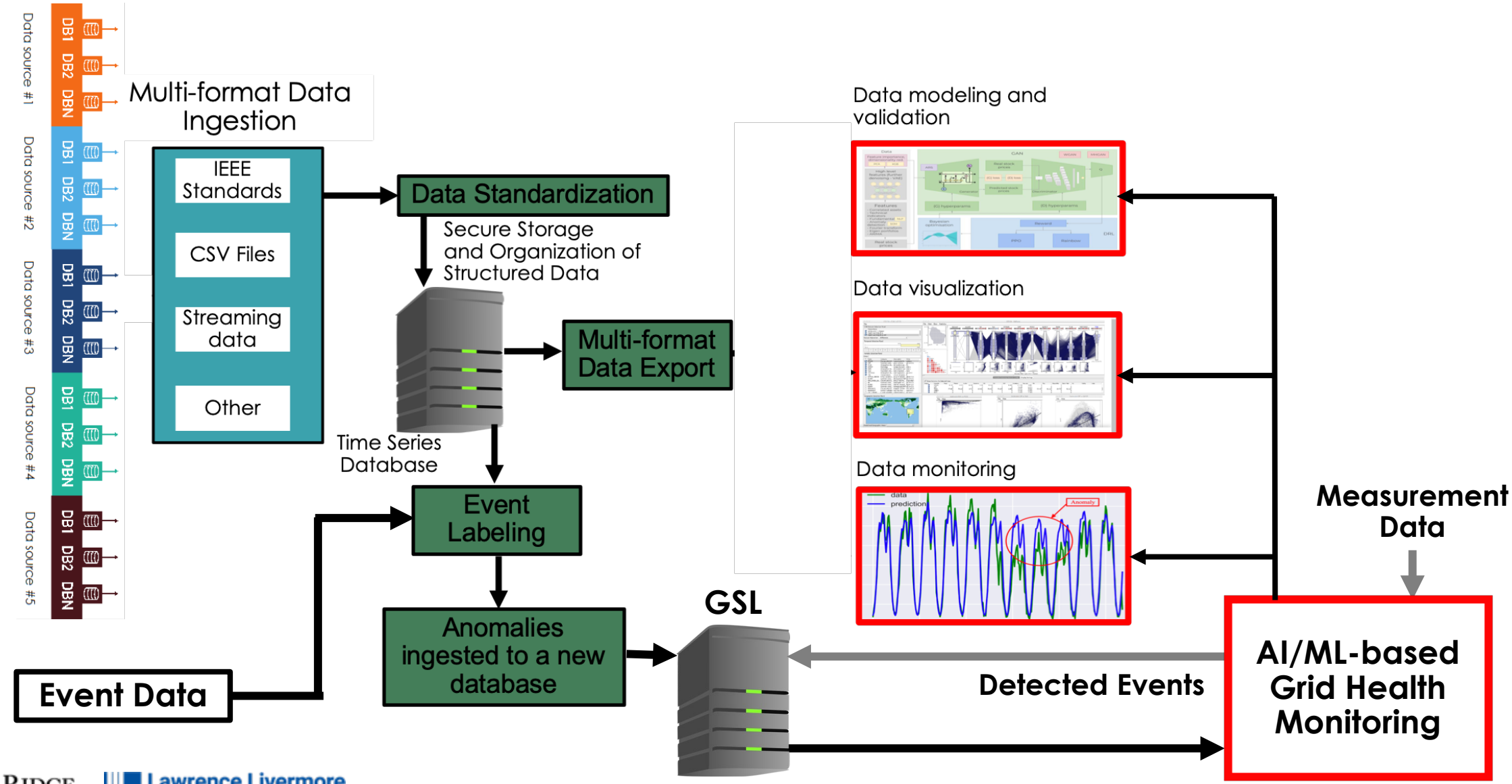


https://bbd.labworks.org/

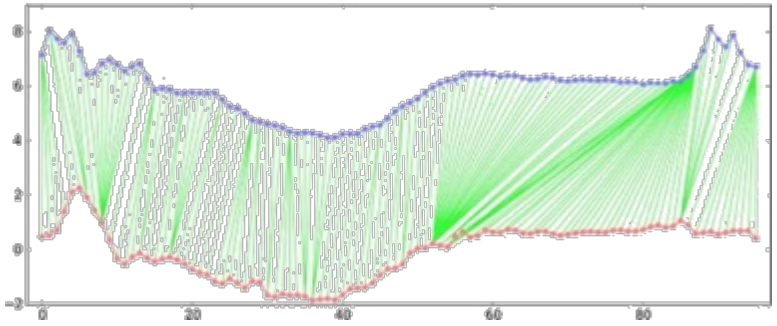
Project Overview

- ORNL and LLNL, funded by DOE Office of Electricity, partnered to develop an open-source Grid Signature Library (GSL)
  - Measurement data: raw data with signatures yet to be extracted
  - Signature data: labeled events with data provided in specific formats
- Goal
  - Facilitate, tag and fuse data feeds from multiple sources
  - Implement a modular architecture for expandable design
  - Anonymize event sources to enable open data sharing
  - Provide go-to resources for event detection and algorithm validation

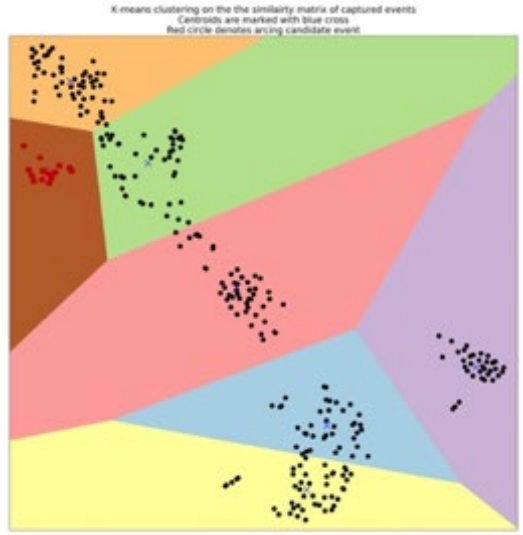
Library Framework



Signature Data Extraction – Example

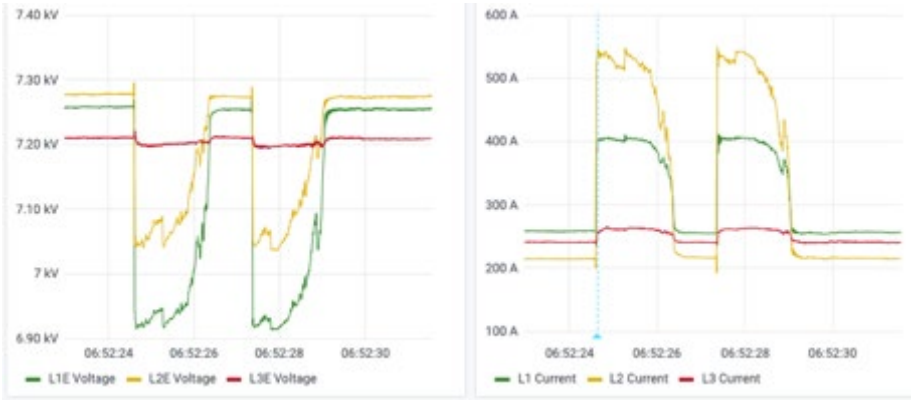


Dynamic time warping (DTW)

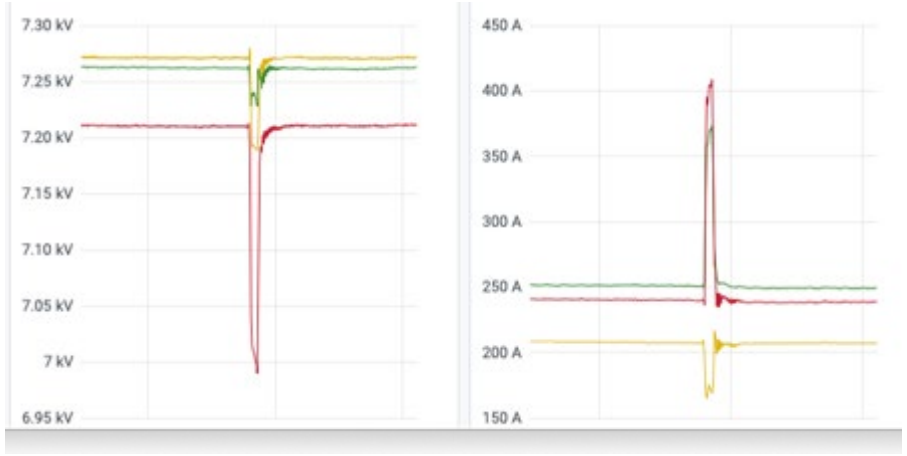


k-means clustering

Sample Signature Data



Blown fuses



Line recloser open (wire down)

Recloser close

Line reclosers open/close (vegetation)

Labeling from analysis of outage records after clustering events

Progress to date

- Collected, reviewed, processed and labeled over 2,600 grid event data from public/private providers, for example
  - DOE/EPRI National Database Repository of Power System Events
  - University of Tennessee Knoxville (UTK) FNET/GridEye Data
  - DOE FOA 1861 Data
    - Eight project teams
    - PMU dataset
      - Covers the three US interconnections
      - Two-year duration
      - Includes event logs

Signature Summary by Data Provider:

No.	Data Provider	Num Signatures	Num Event Tags
1	Provider 1	295	939
2	Provider 2	242	319
3	Provider 3	96	113
4	Provider 4	143	256
5	Provider 5	105	377
6	Provider 6	14	25
7	Provider 7	16	55
8	Provider 8	13	0
9	Provider 9	1031	2286
10	Provider 10	663	1097
	Total	2618	5467

Database Metadata & Schema

- PostgreSQL + TimescaleDB

Signature ID: 4

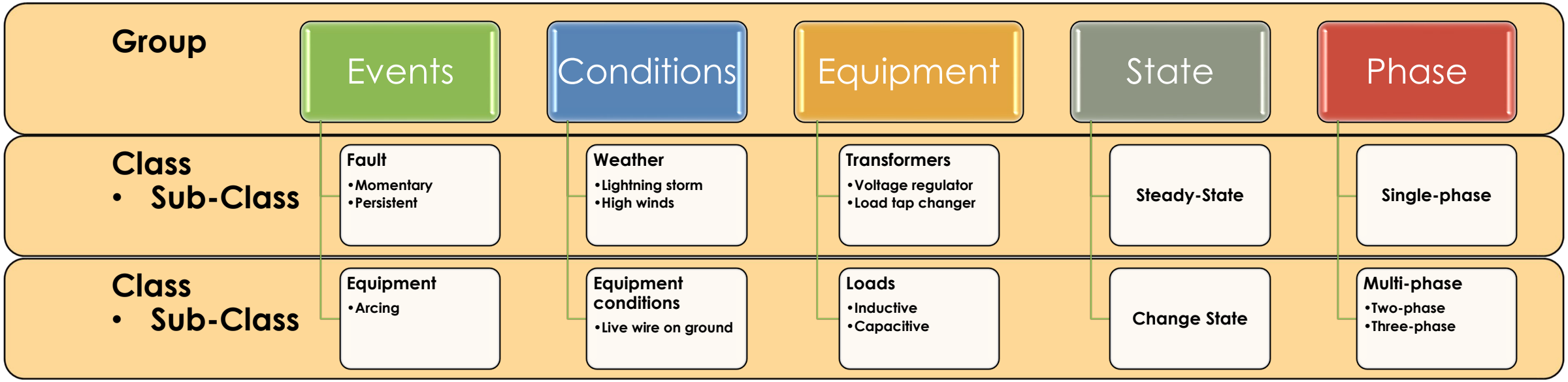
Event Label:	recloser	**Data Source:**	Provider 1
Device Name:	Site0010		
Sample Rate:	7680.0 Hz	**Data Type:**	PoW
Wave Start Timestamp:	2006-08-01T00:00:00	**Wave End Timestamp:**	2006-08-01T00:00:00.199870
Wave Duration:	0.200 sec	**Links:**	Graphs Download data (CSV)
Event Start Timestamp:	2006-08-01T00:00:00	**Event End Timestamp:**	2006-08-01T00:00:00.199870
Measurement Types:	Current(A), Voltage (V)		
Description:	A lightning strike caused a recloser on F_0000026 to operate twice and caused primary wire to fall to the ground.		
Event Tags:	Conditions::Equipment Conditions::Live wire on ground Equipment::Interrupting Device::Recloser Events::External::Lightning Strike		

Signature ID: 975

Event Label:	Trip	**Data Source:**	Provider 9
		Sensor Type:	PMU
Sample Rate:	30.0 Hz		
Wave Start Timestamp:	2016-10-01T00:00:00	**Wave End Timestamp:**	2016-10-01T00:03:59.967000
Wave Duration:	239.967 sec	**Links:**	Graphs Download data (CSV)
		Fault Type:	Trip
Measurement Types:	Frequency(Hz), Frequency - ROCOF (Hz/sec), Positive Sequence Current Angle (deg), Positive Sequence Current Magnitude (pu), Positive Sequence Voltage Angle (deg), Positive Sequence Voltage Magnitude (pu)		
Description:	Number of measuring PMUs: 1		
Event Tags:	Equipment::Overhead Equipment::Conductor State::Change State::Trip		

Hierarchical Event Tagging

- Useful for grouping similar types of disturbances
  - Avoids long list of unique disturbance types/conditions
- Flexible & expandable
  - creating entirely new entry when adding new disturbances



User Interface

Dashboard

- Data visualization & downloading
- Visual analytics of signature data and AI/ML approaches

Application

- Signature matching tool based on AI/ML
- Library API with Shell & Python scripts

Forum

- User's survey
- External data uploading

Signature Matching Tool

- Objective
  - Identify and classify unknown/unlabeled events based on the repository of existing labeled events
- Pre-processing
  - Standardization of measurement data
  - Use variables such as voltage, current and frequency for feature extraction
- Feature extraction
  - Statistical moments: mean, variance, skewness, kurtosis
- Event classification
  - Tested approaches so far
    - Gaussian Naïve Bayes, decision tree, random forest
  - Use unlabeled events in the Library as testing dataset

Basic Classifiers being Tested

Gaussian Naïve Bayes

- Determine class label that maximizes objective function

$$\hat{y} = \arg\max_y P(y) \prod_{i=1}^{n} P(x_i \mid y)$$

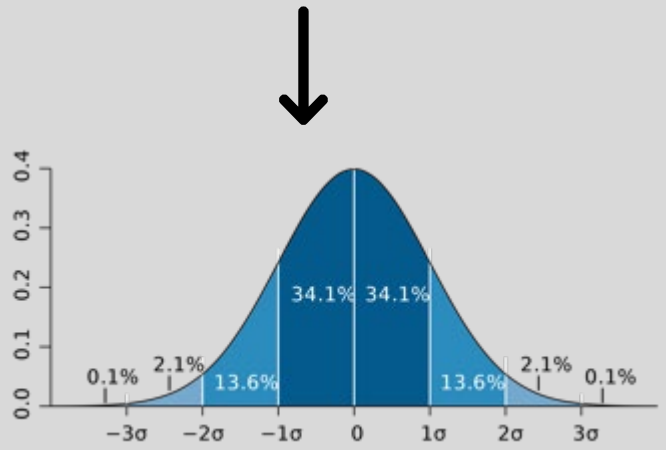


Figure 1. Gaussian Distribution (By M. W. Toews - Own work, based (in concept) on figure by Jeremy Kemp, on 2005-02-09, CC BY 2.5, https://commons.wikimedia.org/w/index.php?curid=1903871)

Decision Tree

- Find "rules" that separate data in correct class

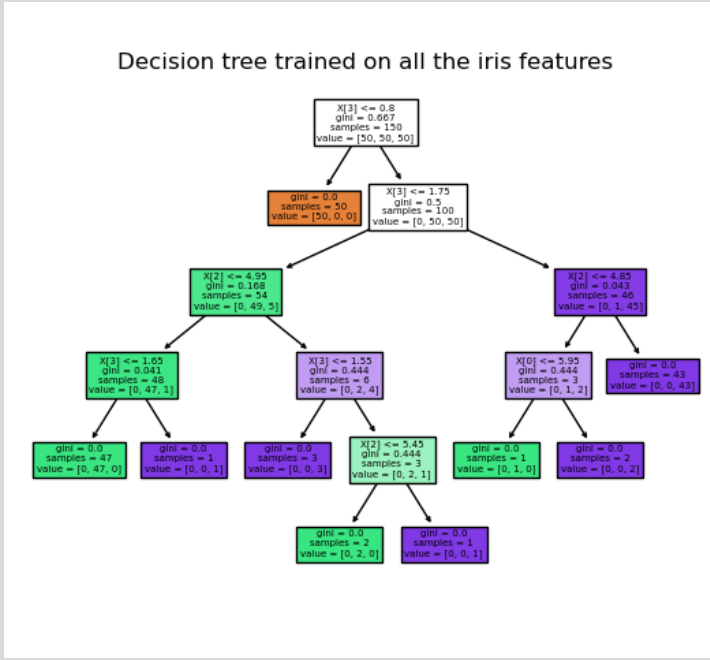


Figure 2. Example decision tree ("1.10. Decision Trees." Scikit, https://scikit-learn.org/stable/modules/tree.html.)

Random Forest

- Create many "short" trees and vote

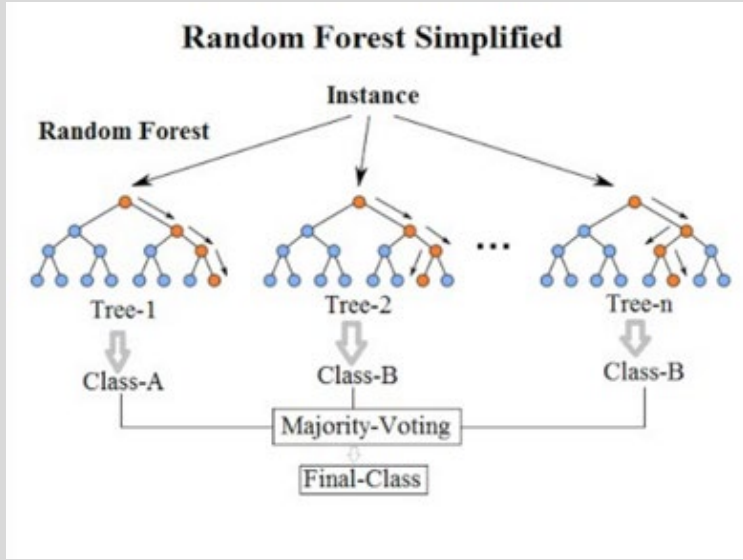


Figure 3. Random Forest Algorithm (By Venkata Jagannath - https://community.tibco.com/wiki/random-forest-template-tibco-spotfirer-wiki-page, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=68995764)

Future Works

- Continue to collect real-world power grid event data
  - Perform intensive outreach and engagement with utility professionals
  - Establish and lead community efforts to speed up data collection
- Collaborate with developers of synthetic event dataset
  - Electric Grid Datasets – Texas A&M University
  - pmuBAGE – University of California, Riverside
- Advance AI/ML technologies for grid health monitoring
  - Enable performance testing, benchmarking and comparison
- Facilitate industrial adoption of AI/ML-based approaches
  - Provide an intuitive and visual understanding of AI/ML