**PV Fault Detection, Diagnostics, and Estimation Algorithms**

Outline and short description of potential learning based algorithms for evaluating solar PV data.

**Anomaly Detection**

*Description:*

The identification of abnormal events, conditions, or observations that do not match with the expected.

*Methods:*

**User Defined Threshold**

Description: User sets the max and min values for certain parameters and when the limits are violated the data point is flagged.

**One-Class Support Vector Machine**

Description: An unsupervised algorithm that learns a decision function for novelty detection. The method can discover anomalies in the training data set.

Toolbox:

1. Sklearn

**Adaptive Resonance Theory**

Description: Unsupervised learning algorithm that divides sectors of the data set into multi-dimensional hyperboxes.

Toolbox:

1. <https://github.com/cbirkj/art-python>

**K-Means Clustering**

Description: Algorithm that separates data into groups based on distances between the centroid of each group.

Toolbox:

1. Python sklearn

**Classification**

*Description:*

The categorization or labeling of an event, condition, or observation based on shared characteristics.

*Methods:*

**Multi-Class Support Vector Machine**

Description: Supervised learning model that separates data using a maximum margin approach using an optimal hyperplane.

Toolbox:

1. Sklearn

**Gaussian Process**

Description: Inference where any finite subset of the data has a joint Gaussian distribution.

Toolbox:

1. Python Sklearn

**Laterally Primed Adaptive Resonance Theory Neural Network**

Description: The LAPART algorithm couples two Fuzzy ART algorithms to create a mechanism for classifying data based on learned associations.

Toolbox:

1. https://github.com/sandialabs/lapart-python

**K-Means Clustering**

Description: Algorithm that separates data into groups based on distances between the centroid of each group.

Toolbox:

1. Python sklearn

**Regression**

*Description:*

The estimation or prediction of continuous quantities. For example, an analyst may be interested in estimating the power output of a PV system based on the irradiance and ambient air temperature.

*Methods:*

**Least-Squares Linear, Polynomial, & Logistic Regression**

Description: Fits an equation to a set of data based on the minimization of the sum of squared residuals.

Toolbox:

1. Python numpy.linalg.lstsq
2. Python scipy.optimize.curve\_fit

**Gaussian Process Regression**

Description: Inference where any finite subset of the data has a joint Gaussian distribution.

Toolbox:

1. Python Sklearn
2. Python custom (birk)

**Convolutional Neural Network**

Description: Feed-forward, multilayer perceptron artificial neural network that mimics biological process such as the human brain. Commonly used for image processing.

Toolbox:

1. Python TensorFlow

**Coefficient-Based Models**

Description: PV models that estimate the voltage, current, and power of a PV system. Models include Sandia Array Performance Model (SAPM), Single Diode Model, and the PVWatts model.

Toolbox:

1. Python PVLIB