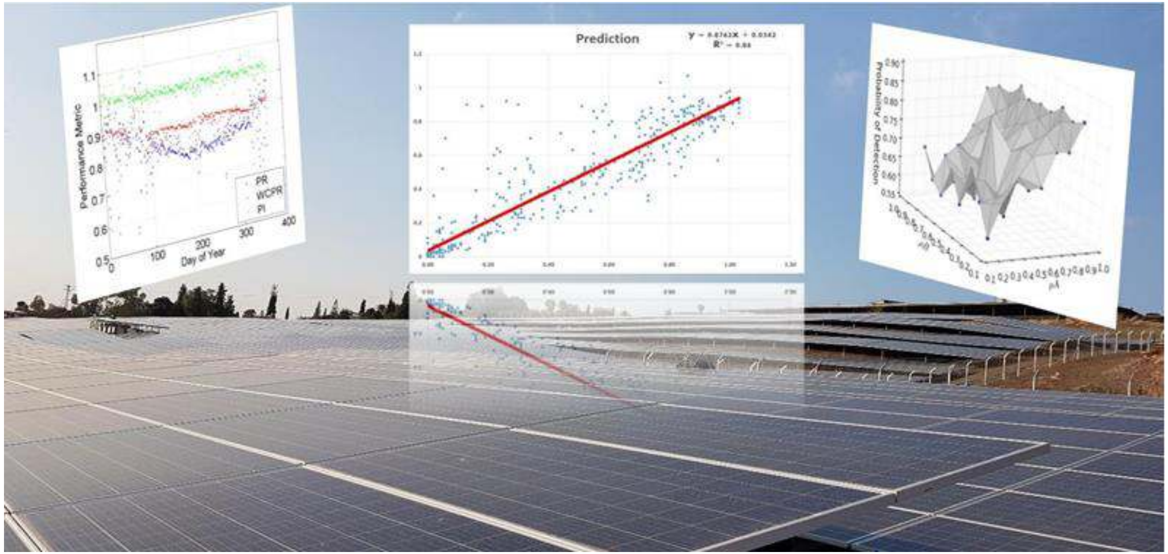


IEA PVPS

Data-driven mitigation measures in Advanced PV plant monitoring

Paolo Graniero, Helmholtz-Zentrum Berlin für Materialien und Energie, Berlin

Intersolar 2021, Munich, October 06th 2021

Agenda

- Report: The Use of Advanced Algorithms in PV Failure Monitoring
- Machine Learning 101
- Overview of ML application characteristics in PV fault detection
- Case Study: Unsupervised PV fault detection algorithms
  - Algorithms
  - Comparison / results
- Conclusions

PVPS

The Use of Advanced Algorithms in PV Failure Monitoring

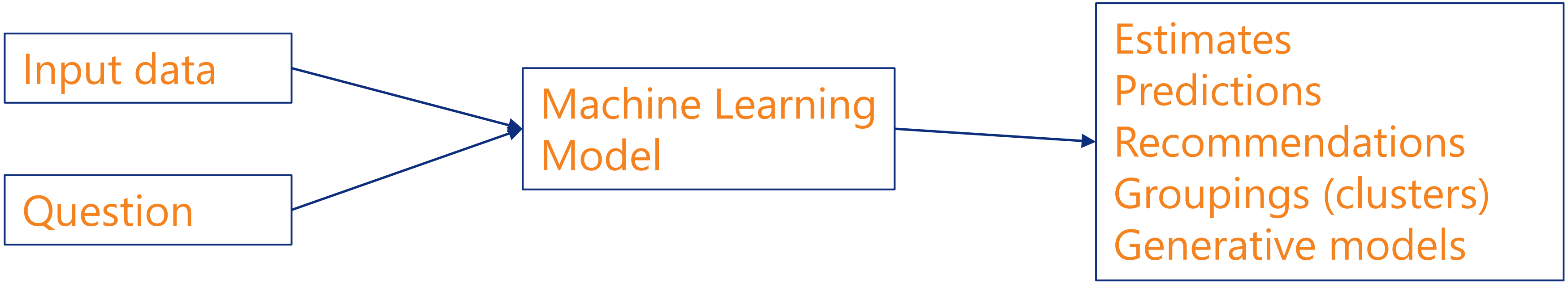
- Statistical and machine learning (ML) approaches
- Data sources

- 22 Fault detection methods included
- **Comparison of 8 unsupervised algorithms**
- Assessment of information content of different data sources

Machine Learning 101

Making machines (algorithms/computers) learn patterns looking only at the data, without explicit rules given

Examples of ML models: artificial neural networks, linear regression, decision trees, support vector machines

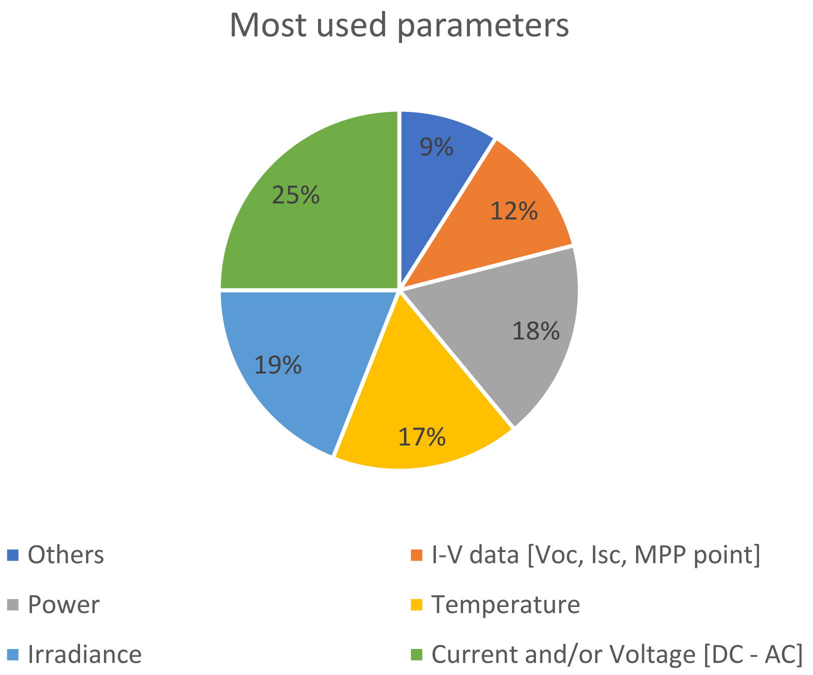


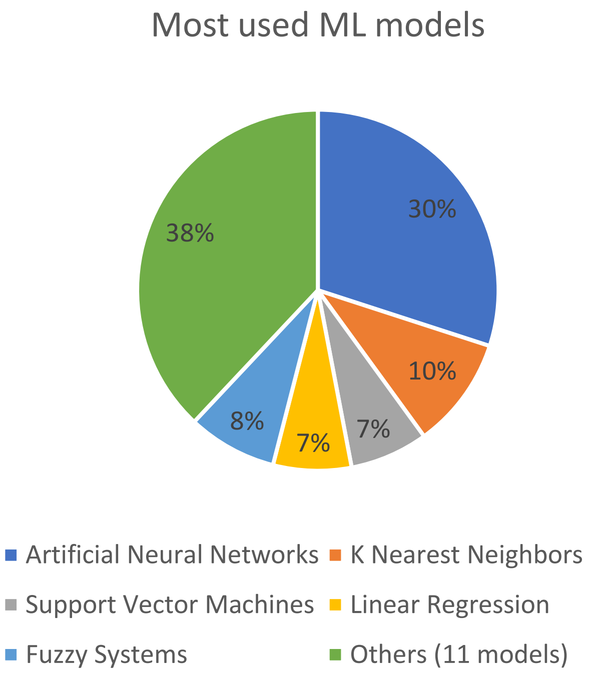
Machine Learning 101

- **Supervised learning**: the data **contains a target variable**, also called label, which we try to estimate/predict (e.g. identify a data points as faulty, having other data points labeled as faulty and normal)

- **Unsupervised learning**: there is **no target variable**, the interest is in finding interesting patterns in the data (e.g. finding patterns in the data, identifying different groups of data points that can be then studied and labeled as faulty/normal)

Overview: ML applications in PV faut detection





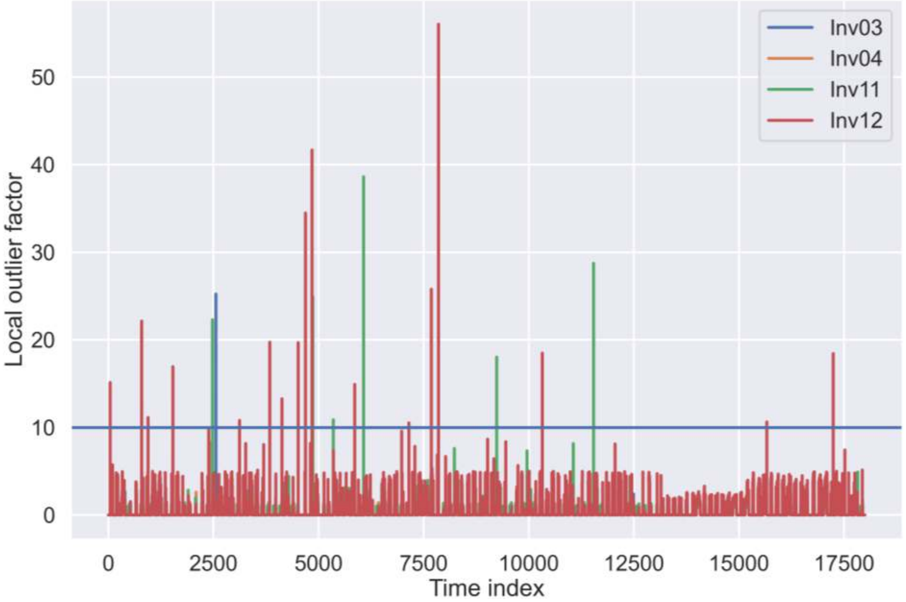
PVPS

Unsupervised Fault Detection Algorithms

Input data does not contain labeled faults

We find patterns in the data, and map them to threshold-dependent classes

No Label!



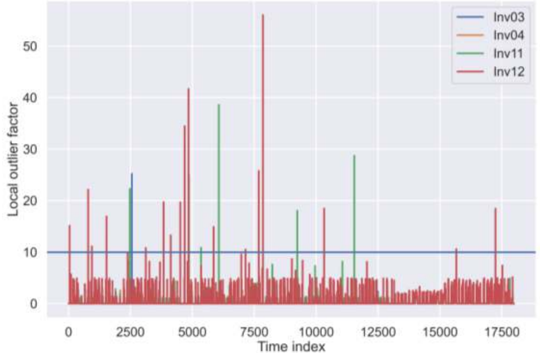
PVPS

High-level distinction

First Group

Estimation of the normal behavior of the system

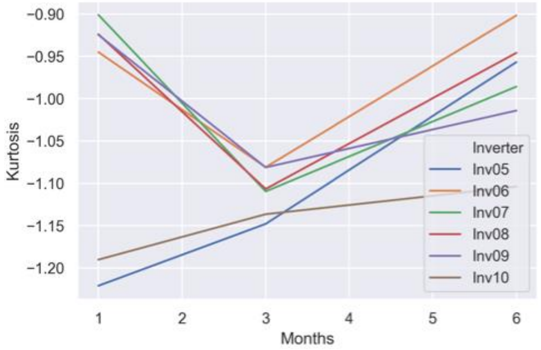
A fault is identified as a significant deviation from the estimated normal behavior



Second Group

Study of the statistical properties of data

A fault is identified by a user from the observation of the time evolution of statistical/structural indices



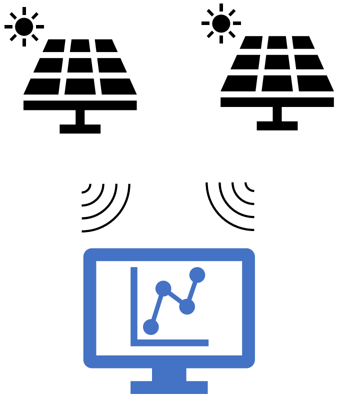
PVPS

Algorithms: Group 1

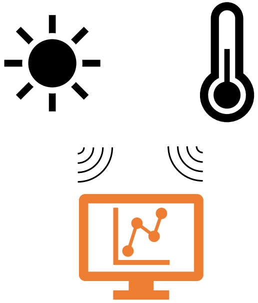
Estimating normal behavior and identifying significal deviations

How is the normal behavior estimated?

According to other systems' behavior



From external factors (irradiance, temperature)



Algorithms: Group 1

- SolarClique: Detecting Anomalies in Residential Solar Arrays
- Local outlier factor-based fault detection and evaluation of photovoltaic system

Normal behavior estimated according to other systems' behavior

- Real-time fault detection in massive multi-array PV plants based on machine learning techniques
- Online Fault Detection in PV Systems
- Intelligent Real-Time Photovoltaic Panel Monitoring System Using Artificial Neural Networks

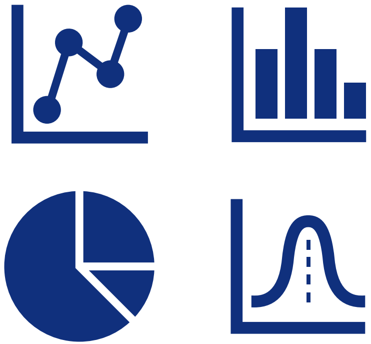
Normal behavior estimated from external factors (Irradiance, temperature)

Algorithms: Group 2

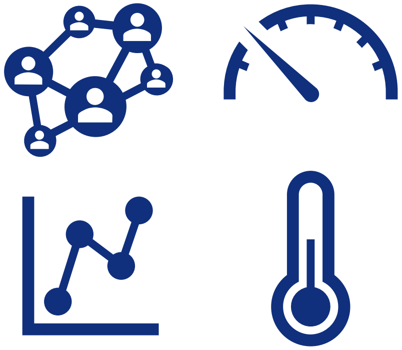
Study of the statistical properties of data

Which one?

Time evolution of statistical indices



Time evolution of network measures



PVPS

Algorithms: Group 2

- A Statistical Tool to Detect and Locate Abnormal Operating Conditions in Photovoltaic Systems
- Statistics to Detect Low-Intensity Anomalies in PV Systems

Time evolution of statistical indices

- Complex Network Analysis of Photovoltaic Plant Operations and Failure Modes

Time evolution of statistical interdependencies of selected monitored sensors (network measures)

Results – Group 1

Algorithm	LOF	SC	OFD	RTFD	NN
LOF	**8**				
SC	1	**1742**			
OFD	0	11	**235**		
RTFD	1	23	39	**494**	
NN	5	1010	190	156	**2739**

~25k data points

High variability in the number of identified faults

Generally low agreement

SC - SolarClique: Detecting Anomalies in Residential Solar Arrays

LOF - Local outlier factor-based fault detection and evaluation of photovoltaic system

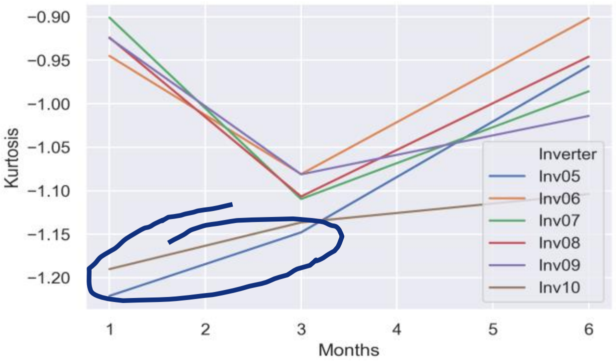
RTFD - Real-time fault detection in massive multi-array PV plants based on machine learning techniques

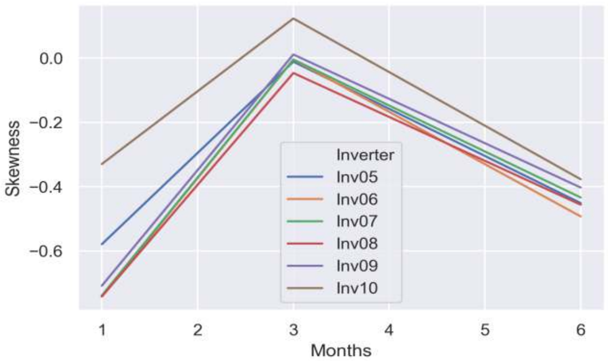
OFD - Online Fault Detection in PV Systems

NN - Intelligent Real-Time Photovoltaic Panel Monitoring System Using Artificial Neural Networks

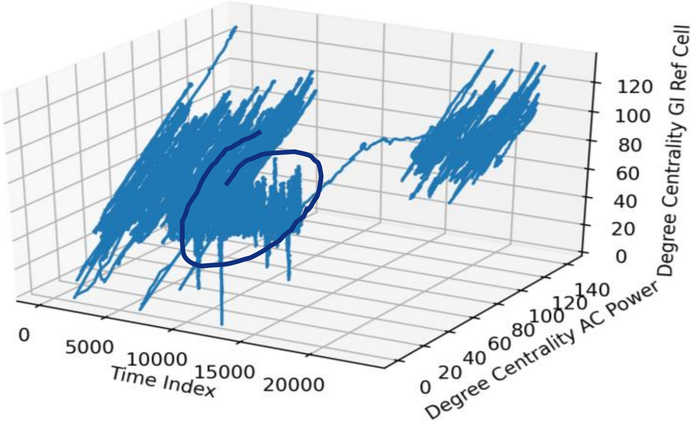
Results – Group 2

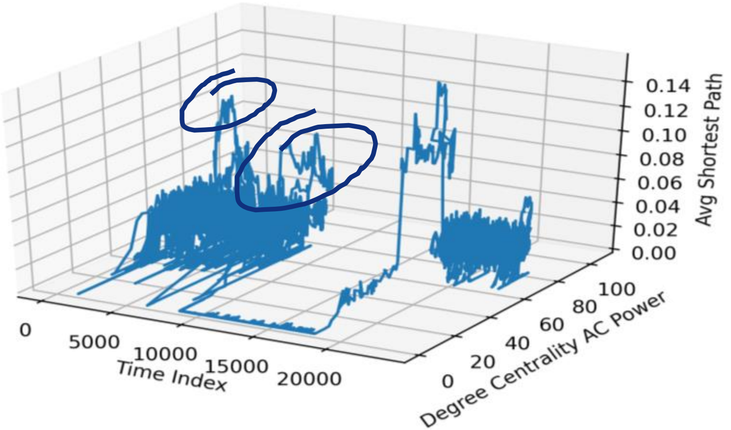
Statistical indices





Network measures





PVPS

Conclusions

- The application of ML and statistical approaches is still new, with high variability in type of application (input data and aim) and potential results
- Need for benchmarking datasets for unsupervised approaches
- Sharing of data and algorithms might help in:
  - The benchmarking process
  - Assessment of stability and reproducibility of ML models

Paolo Graniero, IEA PVPS Task 13 – 2.2
paolo.graniero@fu-berlin.de