Clarification Note

Early Fault Detection in PV Inverters Using Unsupervised Learning with Autoencoders: A Case Study on the GPVS-Faults Dataset

Supervisors: Hai Canh VU; Nassim Boudaoud

Aya Benkirane, Shruti Debath, Maoye Guan, Ningyuan Zhang

Project Objectives

Develop and evaluate an unsupervised deep learning framework using autoencoders for early detection of PV inverter faults on the GPVS-Faults dataset. The model learns normal inverter behavior from healthy data to automatically identify deviations, offering an alternative to traditional rule-based or supervised methods that require labeled fault data and expert knowledge.

Project Deliverables

- A robust unsupervised framework for early detection of PV inverter faults using autoencoders.
- Benchmark results on the GPVS-Faults dataset demonstrating feasibility and sensitivity to early deviations.
- A reusable pipeline adaptable for real-world PV monitoring systems.
- A comprehensive project report.
- A presentation summarizing key findings.
- A clarification note to support understanding and future deployment.

Project Milestones

- **Data Preprocessing:** Import and clean .csv files from the GPVS-Faults dataset; extract relevant features such as DC/AC voltage, current, and inverter-side measurements.
- **Model Development:** Design and train an LSTM-based autoencoder using only healthy operation data to learn the baseline behavior of the inverter.
- **Anomaly Detection:** Identify deviations in test scenarios and set adaptive thresholds to flag early-stage faults.

- **Evaluation:** Assess detection performance using metrics such as anomaly detection accuracy, false alarm rate, and lead time to fault.

Project Team Presentation

Student's Name	Course	Skills
Aya Benkirane	ISC-AOS	ML, DL, Optimization
Shruti Debath	ISC-AOS	ML, DL, CNN, Optimization
Maoye Guan	ISC-AOS	ML, DL, Image Treatment, Optimization
Ningyuan Zhang	ISC-AOS	ML, DL, Semantic Segmentation

Project Team Organization

For convenience and to meet the requirements of the clarification note, each team member was assigned a primary task while all contributed across all areas.

- **Aya Benkirane — Evaluation & Integration:** common splits/metrics, thresholds & event-level alarms, result aggregation & slides.
- **Shruti Debath — Data & Windowing:** cleaning, windowing, healthy-only normalization, gray-zone & metadata, basic features (RMS/harmonics, etc.).
- **Maoye Guan — Traditional Unsupervised:** Threshold/EWMA/CUSUM, PCA, One-Class SVM/LOF/Isolation Forest.
- **Ningyuan Zhang — LSTM-AE:** model & training, reconstruction-error scores, thresholds/alarms with A, cross-mode tests.

Expectations of the Project Team

The project team expects to deepen their knowledge and practical skills in unsupervised deep learning, particularly using autoencoders and LSTM architectures for time-series analysis. Through this project, they aim to gain hands-on experience in data preprocessing, feature extraction, model development, and anomaly detection in real-world PV systems. Additionally, the team seeks to communicate results effectively through reports and presentations.

Project Time Schedule