Machine Learning for Power Grid Control: A Project for Enhancing Resilience through Data Quality

Maximilian Gräser[1], Oscar H. Ramírez-Agudelo,[2] Michael Karl[3]
1 German Aerospace Centre (DLR) - Institute for AI Safety and Security, Germany, Maximilian.graeser@dlr.de
2 German Aerospace Centre (DLR) - Institute for AI Safety and Security, Germany, oscar.ramirezagudelo@dlr.de
3 German Aerospace Centre (DLR) - Institute for AI Safety and Security, Germany, michael.karl@dlr.de

Abstract

Machine learning methods offer promising capabilities for forecasting, decision support, and automated control, but their trustworthiness in critical infrastructures depends heavily on the quality of the underlying data. This project addresses these challenges by systematically analyzing and integrating data quality considerations into the entire ML lifecycle. The project's outcomes include a dedicated data quality framework that identifies leverage points where interventions can enhance the reliability of ML models. By explicitly embedding data quality into the design and application of AI methods, this project contributes to strengthening trustworthiness and resilience in critical energy infrastructures.

Keywords: Data Quality; Machine Learning; Resilience; Forecasting

Introduction

The transformation of the power system through the expansion of renewable energy has led to higher variability, decentralization, and uncertainty in grid operation. Traditional control mechanisms are increasingly insufficient to guarantee stability and resilience under such conditions. AI and machine learning (ML) are regarded as key enablers for future energy systems, offering advanced forecasting, situational awareness, and automated decision-making capabilities (Strecker et al., 2025). However, the deployment of such methods in critical infrastructures raises fundamental concerns about trustworthiness.

A central factor in this context is data quality (DQ). ML methods are only as reliable as the data they are trained and operated with. Low-quality, incomplete, or biased data can lead to poor performance, unreliable decisions, and, in the worst case, unsafe system states. Recognizing the critical role of data quality in ensuring reliable and safe ML applications, this project explicitly integrates DQ considerations throughout the ML lifecycle.

Project Objectives

The main objective of this project is to enhance the trustworthiness of ML applications in grid operation by explicitly embedding data quality into the entire ML development and deployment process. To achieve this, the project focuses on:

1. Designing a lifecycle-oriented data quality framework to systematically integrate DQ considerations into ML development and deployment.
2. Applying the framework to a representative use case: the prediction of solar irradiance as an important factor for solar energy integration into the grid.
3. Developing a data quality model and robustness analyses that demonstrate how the framework can be used to evaluate and strengthen ML models in practice.

Methodological Approach

The data quality framework developed in this project spans the entire ML lifecycle, from data acquisition to model deployment. It identifies leverage points where data quality issues most strongly affect model reliability and provides guidance on how to address them.

To validate the framework, it was applied to a solar irradiance forecasting task. This use case reflects a highly relevant challenge for renewable energy integration, where variability and uncertainty strongly impact grid operation. LSTM networks and other ML methods have been shown to provide effective and accurate forecasts for energy time series, such as power load (Tang et al., 2019), making them suitable for this use case. Within this context, several data quality measures for sensor data were defined to quantify characteristics such as accuracy, precision, and timeliness. In addition, a robustness analysis was carried out to assess how variations in data quality affect the performance of the ML model.

Results and Contributions

The project provides the following contributions to trustworthy ML in grid operation:

- A data quality framework applicable across the ML lifecycle, with clear leverage points for intervention.
- A validation of the framework through its application to solar irradiance forecasting, demonstrating its practical relevance and showing how ML model trustworthiness is affected by variations in data quality.
- A set of data quality measures tailored to sensor data.

These results contribute to improving the resilience of critical energy infrastructures by enabling more reliable forecasting and decision support based on trustworthy ML models.

References

Strecker, K., Behrends, H., Frey, U., Ramírez-Agudelo, O. H., Frison, L., & Paetzold, U. W. (2025). KI als Beschleunigerin der Energiewende. In ForschungsVerbund Erneuerbare Energien (FVEE) (Hrsg.), *Themen 2024 – Die Energiewende mit Forschung beschleunigen* (pp. 49–53). FVEE. https://doi.org/10.5442/t2024

Tang, D., Li, C., Ji, X., Chen, Z., & Di, F. (2019, February 22–24). Power load forecasting using a refined LSTM. In *Proceedings of the 2019 11th International Conference on Machine Learning and*

Computing (ICMLC '19), Zhuhai, China (pp. 104–108). Association for Computing Machinery. https://doi.org/10.1145/3318299.3318353