

Implementation of Machine Learning and Statistical Techniques in Solar Energy Generation Monitoring Systems

Lukas Voveris

Vilnius University

Faculty of Mathematics and Informatics

Naugarduko st. 24, LT-03225, Vilnius, Lithuania

lukas.voveris@mif.stud.vu.lt

Author(s) Name(s)

Author Affiliation(s)

Department/Institute

Full Address(es)

E-mail (s)

Abstract. This research explores the use of machine learning and statistical techniques to improve solar power monitoring. Its key contributions include automated data tagging and classification, interactive heat map based visualization, and performance ratio (PR) calculation. A unified tagging framework and an automated classification system were developed to standardize measurement data across multiple photovoltaic sites. The classification model, leveraging natural language processing techniques, achieved an accuracy exceeding 99.7 %. The study introduces an interactive heat map that tracks inverter-string performance in real time, allowing for fast fault detection. Additionally, a PR calculation system was developed using both real measurements and modelled irradiance data. The accuracy of various tilted irradiance models was evaluated ($MAE = 0.119$, $RMSE = 0.154$), and their suitability for PR estimation was assessed. The results indicate that automated methods can substantially improve monitoring efficiency, analytical precision, and the effectiveness of solar plant maintenance.

Keywords. solar energy, machine learning, heat map, classification, performance ratio, irradiance modelling, visualization, natural language processing

Acknowledgments

This research was done in collaboration with dr. Jolita Bernatavičienė at Vilnius University - Faculty of Mathematics and Informatics.

References

- Chioncel, C. P., Augustinov, L., Chioncel, P., Gillich, N., & Tiran, G. (2009). Performance ratio of a photovoltaic plant. *Bulletin of Engineering, University Politehnica Timișoara*, (Fascicule 2), 555-558.
- Hemetsberger, W., Schmela, M., & Dunlop, S. (2024). *Global Market Outlook for Solar Power 2024-2028. SolarPower Europe*.
- Lindig, S., Louwen, A., Moser, D., & Topič, M. (2020). *Outdoor PV-system monitoring-input-data quality, data imputation and filtering approaches*. *Energies*, 13(19), 5099.
- Mousavi Maleki, S. A., Hizam, H., & Gomes, C. (2017). *Estimation of hourly, daily and monthly global solar radiation on inclined surfaces: Models revisited*. *Energies*, 10(1), 64.
- Nicchiotti, G., Fromageat, L., & Etienne, L. (2016). *Machine-learning strategy for fault classification using only nominal data. Proceedings of the European Conference of the PHM Society*, 3(1).
- SolarEdge. (2025). *Monitoring Platform User's Guide for System Owners (Version 2.1)*. Retrieved May 10, 2025.
- Vidyanandan, K. (2017). *An overview of factors affecting the performance of solar PV systems*. *Energy Scan*, 27-28, 216.