Machine learning Canvas

Background

- Goals : Getting next days pv panels predictions for scheduling purposes (cost, economisations, maintenance, …).
- Pains : The ignorance concerning the amount of electricity produced. Link between the meteorological station and the solar panels production.
- Users: ULiège administrator of the solar panels and management of the electricity.

Value Proposition

- Product: friendly user calendar with prediction of electricity.
- Alleviates: The users will just have to look at the prediction and not have to infer themselves about the production.
- Advantages: Better agreement between teams inside Uliege and easier access/ use of the pv panel maintenance/usage

Objectives

- Production daily prediction module
- Energy management integration (in google calendar for example to provide a friendly interface.)
- Maintenance scheduling feature (implement a feature to tell which days are favorable for maintenance due to low prediction of electricity).

Solution

- Core Features: Solar panels productivity and meteorological data.
- Integration: Use google calendar to put results / Home assistant/ alexa, …, to leverage predictions in it.
- Alternatives: integration with other calendar platforms or create one specified to this.
- Constraints: Data privacy and security + Scalability
- Out-of-scope: Weather station maintenance + Electricity consumption predictions (focus is on predicting electricity production; predicting consumption is considered out-of-scope for this solution) + Physical Hardware Upgrades (The solution will focus on software and data-related aspects; any physical upgrades to solar panels or associated hardware fall outside the scope)

Feasibility

It is feasible due to the availability of the data (+-28k) which will let us use a simple machine learning model to predict.

- Data available and labeled (maybe a possibility to gather more data through time if uliege allows it, enabling the case of an online model.)
- 4 motivated team mates.
- Possibility to use simple ML models (large amount of available data).

Data

- Dataset of the University of Liège. (private)
- 2 files available

Metrics

Error between prediction and real production like MSE and/or MAE.

Evaluation

Offline:

Usage of common machine learning techniques to assess the best possible model, splitting the data between train and test set, k-fold cross validation for hyperparameter tuning and for ensuring the robustness of the different models.

Online:

Look at prediction latency and prediction precision over time

Modeling

Try different things that we learned in previous courses such as:

- Basic ML Models (KNN, Linear Regression, Decision Trees)
- Ensemble methods (Random Forests, GBM, Stacking)
- Small neural network
- Bayesian inference

The iterative approach for the modelization of our task is as follows:

- Data preparation
- Baseline Model
- Develop the different models mentioned above
- Offline and online evaluation
- Iterative refinement

Inference

Beyond point predictions, our project dives into the realm of probabilistic forecasting. Imagine your Google Calendar showcasing a range of possible solar energy production, not just a single value. This empowers you to:

- Make informed decisions: Plan energy usage and grid integration with confidence.
- Prepare for fluctuations: Ensure smooth energy flow by anticipating potential variations.
- Gain deeper understanding: Uncover factors influencing solar panel performance. Our continuously learning model refines these probabilistic predictions, providing an accurate picture of the future. We'll ensure timely updates while maintaining efficient calculations.

This project goes beyond prediction, unlocking the full potential of solar energy through a data-driven, probabilistic approach. Embrace the uncertainty, optimize energy management, and contribute to a sustainable future for our university.

Project

Imagine a future where your Google Calendar effortlessly displays the predicted daily output of our solar panels. This project makes that vision a reality. Every 15 minutes, our model receives fresh weather data and generates a precise prediction, seamlessly integrated into your calendar. Accuracy is paramount, but speed won't be neglected. We'll strike a perfect balance, ensuring predictions are reliable and timely. By constantly monitoring and fine-tuning the model, we guarantee its continued effectiveness, even as data volumes and user needs grow. This project is more than just prediction; it's a step towards a sustainable future, powered by the sun and guided by data-driven insights.