# THE MACHINE LEARNING CANVAS

Designed for: MLOPS Triple P Designed by: A. Birtles, G. Delporte, R. Lambermont, A. Louis Date: 05/01/2024

## PREDICTION TASK

**Type of task?** Regression (predicting a continuous value - the amount of electricity generated).

**Entity on which predictions are made?** Photovoltaic (solar) panels.

**Possible outcomes?** Predicted watt-hours of electricity produced by the panels.

**Wait time before observation?** Immediate for predictions; actual observation wait depends on the data collection intervals (e.g., every 15 minutes).

## DECISIONS

**How are predictions turned into proposed value for the end-user?** Predictions can be integrated into a Google Calendar, allowing the administrator to visualize and manage energy production schedules, maintenance, and usage optimization. This aids in planning and operational efficiency without manual inference.

## VALUE PROPOSITION

**Product**: Friendly user calendar with the 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 and/or use of the photovoltaic panel maintenance and usage

## DATA COLLECTION

**Strategy for initial train set & continuous update:** Utilize historical data from the University's dataset.

Trying to get new data with the help of the University.

## DATA SOURCES

**Where can we get (raw) information on entities and observed outcomes?** Data will be sourced from the University's private dataset. Additional data may be obtained from integrated meteorological sensors and APIs that provide real-time weather data, from the weather station of the University.

## IMPACT SIMULATION

**Can models be deployed?** Yes, models can be deployed after testing.

**Cost/gain values for (in)correct decisions?** Cost includes potential inaccuracies in energy management and maintenance scheduling; gains from accurate predictions include cost savings, optimal resource utilization, and improved maintenance scheduling.

**Fairness constraint?** Ensure the model's predictions do not inadvertently favor certain times/days.

## MAKING PREDICTIONS

**Time available for this + featurization + post-processing?** Immediate processing required; limited to a few seconds to ensure timely updates.

**Compute target?** Must be lightweight enough to run on existing university servers or cloud infrastructure without significant cost increases.

## BUILDING MODELS

**Techniques:** Start with basic models (Linear Regression, Decision Trees), then explore ensemble methods) and a small neural network for better performance as well as a neural network that outputs some distribution parameters..

## FEATURES

**Input representations available at prediction time, extracted from raw data sources:** Features might include time of day, historical energy production data, current and forecasted weather conditions. WE want to have some sort of dashboard that will display the results for decision making.

## MONITORING

**Metrics to quantify value creation and measure the ML system's impact in production (on end-users and business)?** Key performance indicators like Mean Squared Error (MSE)

ONLINE COURSE

# Master the Machine Learning Canvas

Learn a step-by-step process to get to a complete and detailed Machine Learning Canvas. This will help you...

- Validate the feasibility of your ML use case ideas.
- Boost collaboration within your team.
- Anticipate issues that would otherwise come up during implementation or in production.

More details at [ownml.co/plan](ownml.co/plan)