

Exploring E-REDES Open Data with Supervised Learning

Purpose

This lab introduces students to real electricity distribution data provided by **E-REDES**, Portugal's electricity distribution operator (<https://e-redes.opendatasoft.com/pages/homepage/>). The goal is to practice applying supervised learning algorithms to real-world datasets. Students will learn how to prepare the data, choose a target variable, and build predictive models.

Example Datasets (from the E-REDES portal)

- **Power outages** by region and time
- **Duration of interruptions**
- **Energy losses**
- **Voltage quality indicators**

Students may select one dataset or combine several.

Example Prediction Tasks

- Predict the **number of outages** next month in a given region.
- Estimate the **average duration** of outages.
- Classify regions into **high-risk vs. low-risk** for interruptions.
- Forecast **energy losses** based on network characteristics.

Example Algorithms

- **Linear Regression** – for predicting numbers such as duration or energy loss.
- **Decision Trees** – for classifying regions.
- **Random Forests / Gradient Boosting** – for more advanced predictions.
- **Logistic Regression** – for binary outcomes (e.g., high vs. low risk).

Tasks

1. Download one or more datasets from the E-REDES portal.
2. Explore the dataset to identify features (inputs) and a potential target (output).
3. Apply at least one supervised learning algorithm of your choice.
4. Compare predicted values with real values.

Outcome

At the end of the lab, students should be able to:

- Use real-world energy data in a machine learning workflow.
- Train a supervised learning model with Python or another tool.
- Understand how predictions can be applied in electricity distribution.