Assignment 2 concentrates on day-ahead generation scheduling from the perspective of the system operator. In this assignment, you will investigate how to benefit from classification for this application. As in the previous assignment, the work to undertake involves a bit of mathematical modelling, implementation in your favorite modelling language (preferably Python), generation and discussion of results, as well as presentation of the work in a short report.

The aim of Assignment 2 is to evaluate

- your understanding of day-ahead scheduling from the system operator perspective,
- your ability to implement linear and non-linear classification,
- your critical analysis of the results generated.

The expected outcome of Assignment 2 includes:

- a report of maximum 10 pages (excluding appendices),
- code delivered as supplementary material.

The evaluation of Assignment 2 will count for 33.33% of the final grade. **Individual contributions to the assignment must be clearly stated in the report.** If not, equal contribution will be assumed. Good luck!

Description of the assignment

In this assignment, you will apply linear and non-linear classification for the day-ahead scheduling problem, specifically unit commitment. Envision yourself as the system operator, aiming to accelerate solving the unit commitment problem by learning from historical data. Unit commitment in energy systems involves determining which power units should be online and offline in each period. This problem is crucial for operational planning, ensuring optimal resource utilization and stability. In this assignment, you will see one way of using machine learning for this application.

Step 1) Understanding the problem: Please describe the unit commitment problem, model it as an optimization problem, and identify the variables, objective function and constraints involved. You can use the 6-bus test system from Figure 1 or any other test system that you prefer (you are not restricted to this use case) for your simulations. The 6-bus test system, shown in Figure 1, includes 3 conventional units (G1-G3), 2 wind farms (W1-W2), and 3 loads (L1-L3). The technical data related to generating units and transmission lines, capacity of wind farms, and the nominal value of loads have been attached.

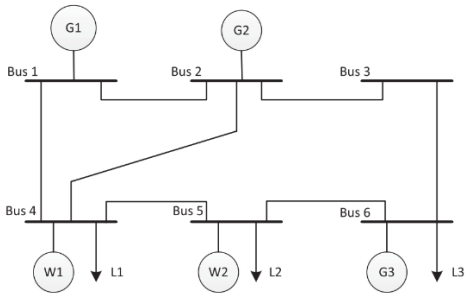

Figure 1. The IEEE 6-bus test system with three conventional units G1-G3, two wind farms W1-W2, and three loads L1-L3.

Step 2) Data preparation:

Step 2.1) Please create different samples of loads and production of the wind farms. Then, run simulations of the unit commitment for all samples and save the samples and their corresponding optimal values of binary variables for the next step.

Step 2.2) Using the results from the unit commitment simulations, define the features and labels and split the dataset into training, validation, and testing.

Step 3) Classification: Please train a set of binary classifiers (one for each generating unit at each time) using the prepared dataset to predict the optimal value of binary variables. Please try both linear and non-linear classification. Note that some generating units are always on or off. Therefore, you do not need to implement classification techniques for them.

Step 5) Evaluation: Please evaluate and compare the performance of linear and non-linear classifiers in predicting binary variables. Please elaborate the metrics you use for evaluating.

Step 6) Please discuss the limitations of classification models for the unit commitment problem. How can you deal with the limitations and still consider classification models as part of finding an optimal solution?

Note: You may use Gurobipy and scikit-learn API for this assignment.