Project Reinforcement Learning (2024)

8 January 2024 - 2nd of February 2024

One of the main application of reinforcement learning is in operations research. Among the many domains of operations research, one of increasing importance is in the smart grids domain. The proportion of new vehicles being electric vehicles (hybrid or pure electric) has surpassed 50% in many European countries. If a household is equipped with a smart meter that makes the consumer pay a different price of electricity depending on the time and day, batteries from electric vehicle could be charged at the time that is cheapest or even be discharged at times when prices of electricity is very high. One key challenge is the daily operation of these batteries. In this project you are in the role of the data science team and need to define a control strategy for new clients making use of such a system.

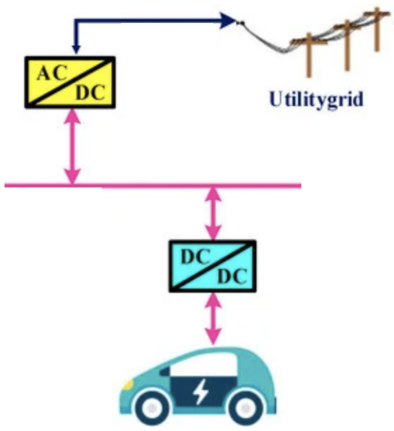

Figure 1: Illustration of a household including PV panels and electric vehicule

You consider a particular client that has the following characteristic. The battery of the car has a battery of 50kWh of usable capacity (you can imagine that the car owner never allows the battery to fall below 50kWh of maximum charge). The efficiency of the battery is 90% for the production of electricity and 90% for the storage (by releasing 1 kWh of energy from the battery, you only generate 0.9 kWh of electricity and by using 1 kWh of electricity, you only store 0.9 kWh of energy into the battery). The maximum power available is 25kW to ensure a better battery life span.

You need to suggest a control strategy for the operation (with unseen time series, i.e. next electricity prices) of the battery given 3 years of past electricity prices on the grid. These electricity prices (euros/MWh) are given on Canvas in the file *'train.xls'* that you can use to train/validate your model. You can assume that the price of electricity on the market is independent of the quantity of electricity that you trade with the market. Electricity is sold at the spot market price but is bought at twice the price to pay for the transmission costs and various taxes.

One important element is that the operation algorithm of your battery can only access the data up to the current time step. The future price is stochastic and unknown at the time of taking a decision because it depends on the offer and demand at the later time steps. This is also why RL is useful (otherwise an optimization formalization of the problem could solve this).

The car (hence the battery) is attached to the grid all the time during the night but one day out of two (that you can estimate with random chance), the car becomes unavailable from 8am to 6pm and comes back with 20kWh less. The car must always have at least 20kWh of usable capacity in the battery every morning at 8am, which is enforced by charging the battery from 7am to 8am to at least 20kWh if it was charged lower than 20kWh before.

Presentation 1 (19th of January)

You need to make a presentation where

- you have implemented the environment (suggestion: you can do it as a gymnasium environment https://gymnasium.farama.org/),
- you have implemented at least one baseline algorithm (not necessarily RL),
- you can provide visualization of the operation of the battery as well as estimated performance (with the validation time series),
- you explain the key next steps to improve the model.

You need to make a presentation (5 minutes presentation + 5 minutes feedback/questions).

Report (1st of February) and presentation 2 (2nd of February)

The final report and presentation need to show

- an RL solution that you have developed for the problem at hand,
- how you have validated the RL algorithm and what the estimated performance is,
- the code with a pre-trained model that can be run on new time series that will be provided. The additional data will be provided in the same format but possibly less days (e.g. with 1 year of data). The code should also have a README file that explains how to run the code for training the model and how one can use the pre-trained model.

You need to hand in a report (max 8 pages, appendix allowed) along with source code. You also need to make a presentation (5-8 minutes presentation + 5-8 minutes questions).

Additional research questions

As part of your final report and final presentation, you can also add variants to this problem and/or perform ablation analysis to highlight the key parts of your algorithm.

Additional avenues that you can investigate:

- Run the same algorithm with less/more features and see how the performance is impacted.
- Explain different validation strategies and what are the pro and cons of using a given technique. You can also try to estimate how confident you are about the performance of the operation (e.g. by estimating the variance of the performance via for instance cross-validation).

Other

- Report and presentation (8 points)
  - Formalization (clarity)
  - Description of the results
  - Additional more open research questions that is investigated
- Code and its performance (2 points)
  - Details on how the required API that your code will need to have will be provided on Canvas such that your code can be tested with new unseen time series.