Case Study Challenge

Machine Learning Data Scientist

Entrix Data Science Challenge

Welcome to the Entrix Data Science Challenge. This challenge consists of 3 questions designed to assess your skills in time-series forecasting, optimization, and reinforcement learning, within the context of energy markets and battery storage trading. These are skills we require from machine learning data scientists at Entrix, as well as the ability to quickly build an understanding of the problems our products solves. The challenge therefore intends to test these skills in the context of a simplified version of the problem our Entrix product also addresses.

Please read through the entire challenge before attempting any of the questions. Feel free to answer the questions in any order you like and to put more emphasis on the questions that you feel address some of your strengths, while moving through the other questions more quickly.

In all three questions of the Data Science Challenge, please add documentation that provides sufficient details to understand the underlying logic of written code where applicable. Where no documentation is provided, please ensure that the code can be understood easily. For documentation, a specific format is not required - in-code comments can be good enough.

Question 1 – Forecasting Day-Ahead Auction (DAA) Prices

Electricity prices fluctuate due to supply and demand variations, regulatory policies, and external market factors. The day-ahead auction (DAA) price represents the hourly electricity prices for the following day, set through market bidding mechanisms.

In this task, you are required to develop a forecasting model to predict hourly DAA prices based on the 60 minutes historical data. Your model should aim to capture underlying patterns and trends to provide reliable predictions. You may use any suitable time-series forecasting technique and your chosen approach should be justified with an explanation of why it is appropriate for the problem.

Once your model is built, evaluate its performance using relevant metrics such as Mean Absolute Error (MAE) or Root Mean Square Error (RMSE). Provide visualizations comparing your predicted prices with the actual market prices to illustrate the accuracy of your model.

Question 2 – Optimization Utilizing Forecast

Energy storage systems, such as batteries, provide opportunities to buy electricity when prices are low and sell when prices are high, thereby maximizing revenue. Using the forecasted DAA prices from Question 1, you will now design an optimized trading strategy for a battery energy storage system.

The battery is characterized by the following specifications:

- It has a maximum capacity of **1 MWh** and can charge or discharge at a maximum rate of **1 MW** per hour.
- It is not possible to charge the battery beyond its full capacity or discharge beyond its empty state.

Your task is to formulate an optimization problem that determines the best charge and discharge schedule to maximize profit. The optimization should respect the physical constraints of the battery while strategically leveraging price fluctuations.

Compare the profit made by your optimisation strategy using your forecasted DAA price from question 1 against profit made using actual DAA price from historical data (perfect foresight) and evaluate how well your forecast + optimisation trading strategy performed.

Question 3 – Reinforcement Learning-Based Optimization Approach

Instead of using a deterministic optimization approach, reinforcement learning (RL) can be applied to learn an optimal battery trading strategy through interaction with past price data.

In this question, you are **not required to implement a working RL model**. Instead, you should provide a structured and detailed explanation of how you would design an RL-based approach for optimizing battery trading.

In your response, clearly define the key components of the RL framework:

- **State space**: Describe what information the RL agent would observe at each decision step.
- **Action space**: Specify the possible actions the agent can take, such as charging, discharging, or holding the battery state.
- **Reward function**: Define how the agent receives feedback for its actions, ensuring that it aligns with the goal of profit maximization.
- **Training strategy**: Explain what RL algorithm you would use (e.g., Q-learning, Deep Q-Networks (DQN), Proximal Policy Optimization (PPO)) and why it is suitable for this problem.

Compare your proposed RL-based approach with the **optimization technique from Question 2**. Discuss the advantages and disadvantages of using reinforcement learning for this task, particularly in scenarios where **price forecasts may be uncertain**. Outline the challenges associated with training an RL agent in this setting and suggest strategies to improve learning efficiency.

If you want to any further into the topic, here are some recommendations:

It is not mandatory to read them and you can solve the case challenge without, but it might be helpful to understand the energy industry better.

- Current Status of Energy Transition
- How does the German Energy Market work? (available in German only)
- General Introduction to the Potential of Battery Storage
- Global Energy Storage Market Set to Hit One Terawatt-Hour by 2030
- Electricity Markets in Great Britain - A guide to understand it
- Fluence (Julian Jansen) on the energy system (Steffen as part of a Twaice Vision Speaker Series) - Watch 4:26 min - 22:45 min