

CAPSTONE PROJECT

IMBALANCE ENERGY PRICE PREDICTION

neuefische
Data Science Bootcamp 2021
Cologne

2021-08-10

Team

Christian **Dischke**
Jaouad El Yousfi
Boris **Geller**
Matthias **Schmidt**



An aerial, high-angle photograph of a large, modern architectural complex, possibly a university or corporate campus. The building features multiple levels, courtyards, and a complex arrangement of wings. The entire image is overlaid with a semi-transparent red filter. In the bottom right corner, a small white box contains the number '7'.

01

Introduction

INTRODUCTION :: IMBALANCE

Physical and Technological Constraints

1. Conservation of Energy
2. Grid Instability
3. Low Ability to Store Electrical Energy

→ *The need to match supply and demand every single moment!*

→ *Imbalance is result for false predictions of power supply and demand.*

→ *How to handle energy imbalances!?*



IMBALANCING ENERGY :: REGULATIVE & QUANTITATIVE



Imbalancing Energy

- Compensate for unforeseen deviation in power demand or power generation
- Allocation is executed by Transmission Operators

Imbalancing Price

- Updated every 15 min
- Calculated by the Transmission Operators
- Published in the following month within 10 business days

→ *Highly delayed price information*

Price Range €/MWh

Year	Min	q.05	Median	q.95	Max
2018	- 1.873	- 37	46	108	2.013
2019	- 2.320	- 39	39	97	2.865
2020	- 6.504	- 61	29	118	15.859

→ *High price range*

Ø Rate of Change

$$|\text{Median}|: \frac{36\%}{15 \text{ min}}$$

$$|\text{Mean}|: \frac{275\%}{15 \text{ min}}$$

→ *Extremely dynamic price development*

IMBALANCING ENERGY :: REGULATIVE & QUANTITATIVE



Imbalancing Energy

- Compensate for unforeseen deviation in power demand or power generation
- Allocation is executed by Transmission Operators

Imbalancing Price

- Updated every 15 min
- Calculated by the Transmission Operators
- Published in the following month within 10 business days

→ *Highly delayed price information*

Price Range €/MWh

Year	Min	q.05	Median	q.95	Max
2018	- 1.873	- 37	46	108	2.013
2019	- 2.320	- 39	39	97	2.865
2020	- 6.504	- 61	29	118	15.859

→ *High price range*

Ø Rate of Change

$$|\text{Median}| : \frac{36\%}{15 \text{ min}}$$

$$|\text{Mean}| : \frac{275\%}{15 \text{ min}}$$

→ *Extremely dynamic price development*

Business Implication

What, if we know the Imbalance Price ahead?

→ *Responding anti-cyclically to the supply situation*

OUR OBJECTIVE :: PRICE PREDICTION

U



→ *Can we predict the
imbalance price
several hours ahead?*

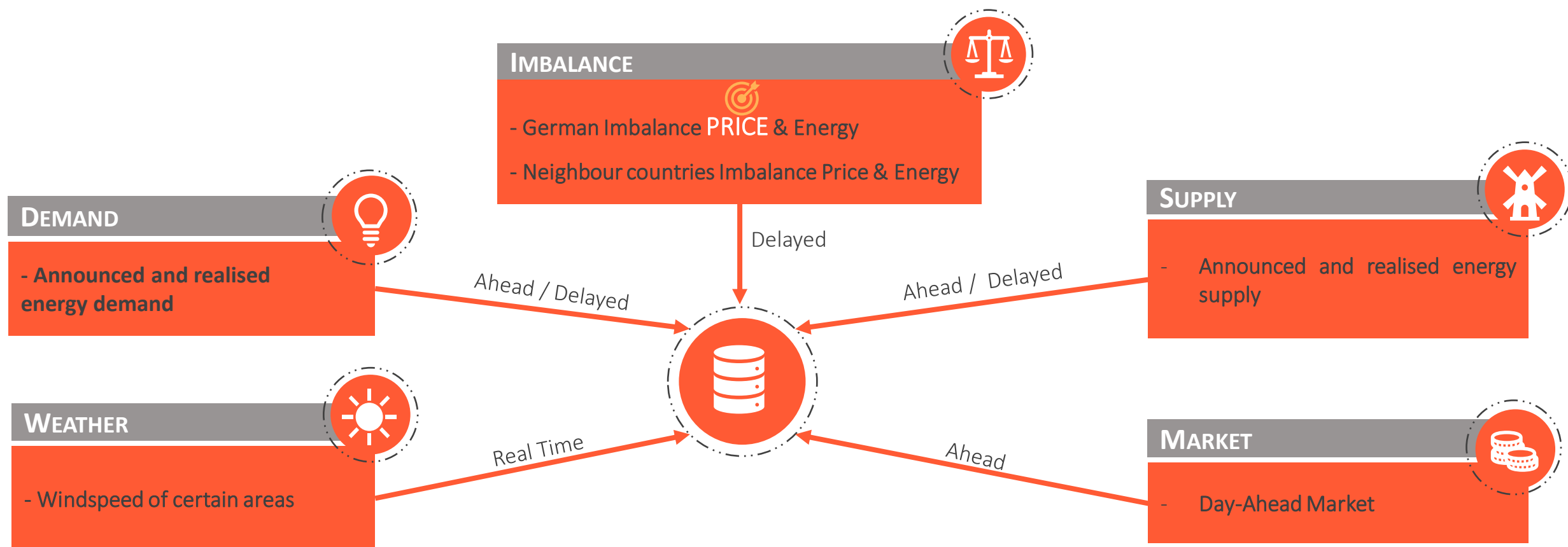




02

Our Data

OUR DATA SET :: SOURCES AND DESIGN



03

Insights :: Imbalance Price

*"You Are The Average Of The Five People
You Spend The Most Time With"*

Jim Rohn

INSIGHTS :: IMBALANCE PRICE

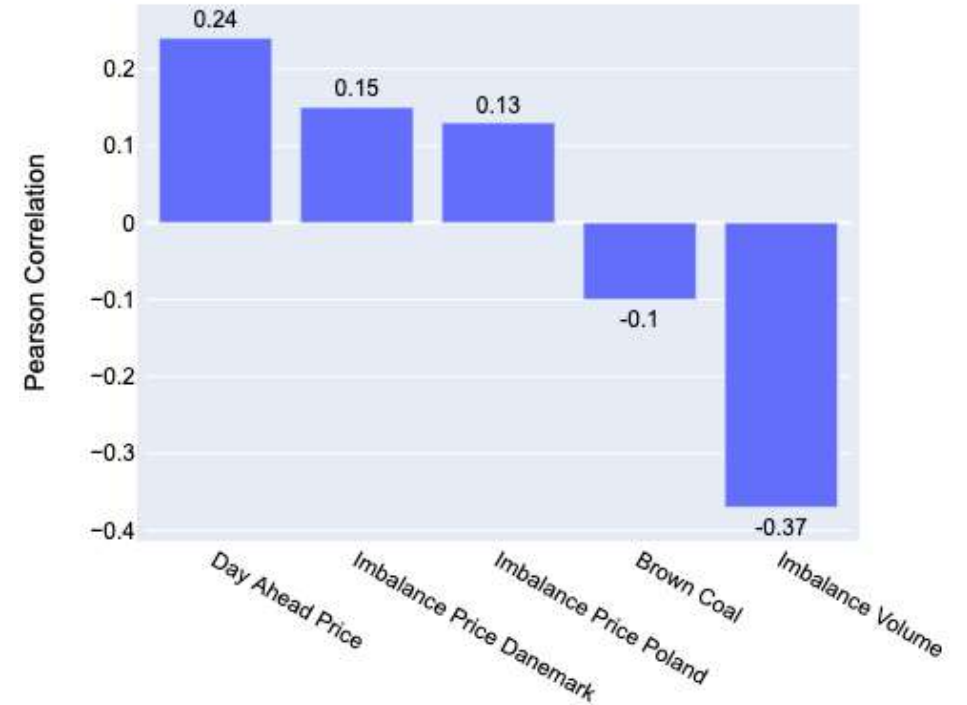
Time Series Analysis

We didn't detect any significant repeating patterns or cycles in the Imbalance Price. 😞

→ *The highly delayed Imbalance Price can't be predicted from.*

Data Environment

Imbalance Price :: Top 5 Correlation Feature



04

Regression



TIME SERIES :: RESULTS

Hours Ahead	Best Model	RMSE €/MWh
1	Encoder-Decoder	59,27
3	Bi-Directional LSTM	71,68
6	GRU	68,23
12	GRU	76,38
24	GRU	79,14

DESIGNED MODELS
<ul style="list-style-type: none">- GRU- CNN- LSMT- Bi-Directional LSTM- Encoder-Decoder

GRU: Gated Recurrent Unit



TIME SERIES :: RESULTS

Hours Ahead	Best Model	RSME €/MWh
1	Encoder-Decoder	59,27
3	Bi-Directional LSTM	71,68
6	GRU	68,23
12	GRU	76,38
24	GRU	79,14

Reference

Year	q.05 €/MWh	Median - RMSE _{1h} €/MWh ≈	Median €/MWh	RMSE _{1h} + Median €/MWh ≈	q.95 €/MWh
2018	- 37	-14	46	106	108
2019	- 39	-20	39	99	97
2020	- 61	-30	29	99	118

DESIGNED MODELS
<ul style="list-style-type: none"> - GRU - CNN - LSMT - Bi-Directional LSTM - Encoder-Decoder

Observation

The RMSE is slightly less, than the difference between median an the 90% boundaries of the last three years

Conclusion:

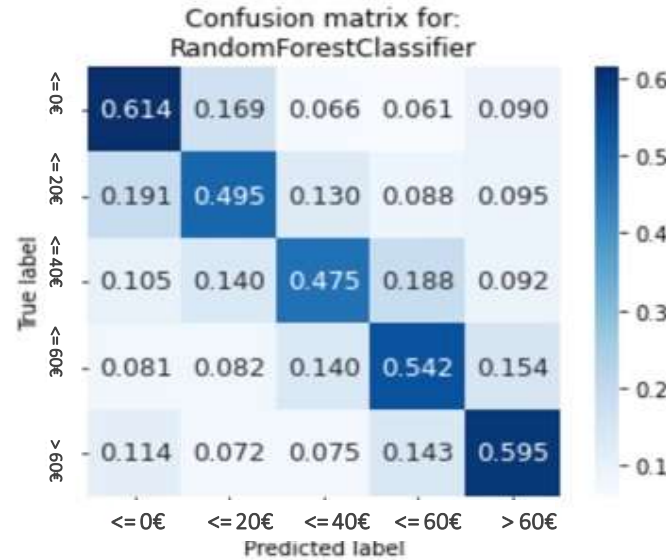
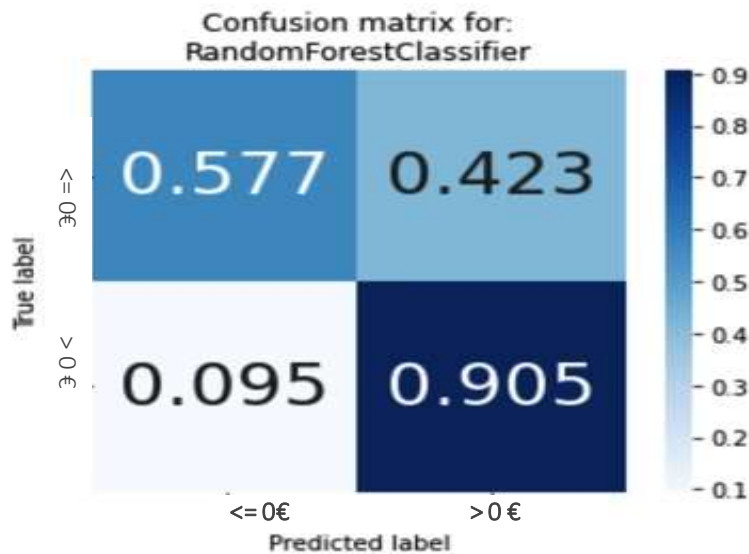
Low precision at present stage.

05

Classification



TIME INDEPENDENT APPROACH



DESIGNED MODELS

- PCA + Random Forest Classifier
- Random Forest
- PCA + Gaussian Naive Bayes

DESIGNED CLASSES

- Positive and Negative Prices
- < 0 | ≤ 20 | ≤ 40 | ≤ 60 | ≥ 60

Observations

Random Forest is capable of predicting both the lowest and highest rebap prices with close to 60% accuracy.

Conclusion: The model works very well for the classes *positive and negative prices* and acceptable for *5-Bucket classification*.

→ What is the **overall** performance of our best model?

Classification Setup	Accuracy
Positive and Negative Prices	83.5 %
5 Price Segements	54%

06

Summary



SUMMARY

Insights ¹

Imbalance Price Characteristics

- Delayed publication,
- High price range and volatility,
- Price can't be predicted from its history.
- Low correlation with given data

Predictions ²

Regression Approach

- Low precision at present stage.

Classification Approach

- Classification of **positive and negative prices** with a total **accuracy of 83.5 %**.
- Price classification with **5 price segments** leads to an **accuracy of 54%**.

The background of the slide features a close-up of a thick, braided rope with a metal carabiner attached. The rope and carabiner are positioned diagonally across the frame. The entire image is overlaid with a semi-transparent orange filter. In the upper left corner, the number '07' is displayed in a large, white, sans-serif font.

07

OUTLOOK

OUTLOOK :: WHAT COULD LEAD TO BETTER RESULTS?

Additional Data Mining

- Expand with additional Weather Data

Feature Engineering

- Methods Technical Analysis of the Financial Markets

Modeling

- Error Analysis

