# Optimizing Next-Day Energy Price Predictions

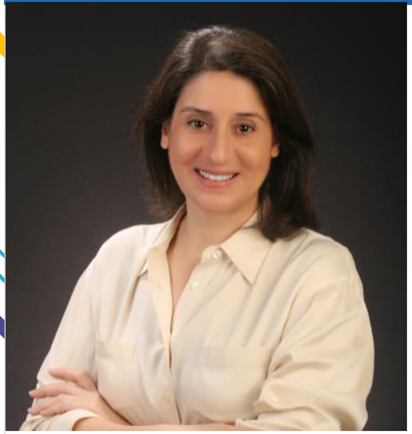
Integrating Forecaster Insights for Accurate Next-Day Energy Price Predictions

Katerina Psallida
Konstantinos Tsagkaropoulos

Prepared for Big Blue Data Academy, Data Science Course, July 2024

# B I O G R A P H I E S

## KATERINA PSALLIDA



**Work Experience**

- Digital Sales for Data & AI Products
- Marketing Professional
- SAP Netweaver / Basis Consultant

**Studies**

- MSc Computational Physics
- BSc Physics

## KOSTAS TSAGKAROPOULOS

**Work Experience**

- Internet/Field Sales
- Telecommunications
- Creative Marketing
- Internet Sales

**Studies**

- BA Marketing & Advertising

# A G E N D A

1. Present the Business Need and the Question Posed by PPC

2. Description of the Dataset Provided by PPC & **E**xploratory **D**ata **A**nalysis

3. Comparative Analysis of Forecasters

4. Forecasting the Best Next Price (Regression | Timeseries | RNN)

5. Conclusions

# A G E N D A

1. Present the Business Need and the Question Posed by PPC
2. Description of the Dataset provided by PPC & **E**xploratory **D**ata **A**nalysis
3. Comparative Analysis of Forecasters
4. Forecasting the Best Next Price (Regression  Timeseries | RNN Analysis)
5. Conclusions

# BUSINESS CHALLENGE

## SCOPE

- Utilize Historical Timeseries Data with Price Predictions from 3 Forecasters
- These are Combined with Actual Hourly Energy Prices
- Evaluate each Forecaster's Historical Performance compared to the Actual Energy Prices
- Enhance the Predictive Accuracy of the 3 Forecasters, by using their predictions to predict a more accurate Next Day Price

# A G E N D A

1 Present the Business Need and the Question Posed

2 Description of the Dataset provided by PPC & **E**xploratory **D**ata **A**nalysis

3 Comparative Analysis on Forecasters

4 Forecasting the Best Next Price (Regression & Timeseries analysis)

5 Conclusions

# THE DATASET

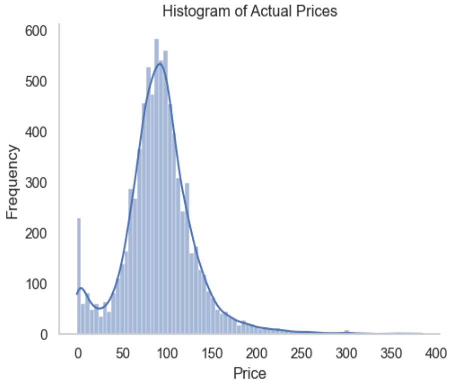
**8k** Rows

**5** Columns

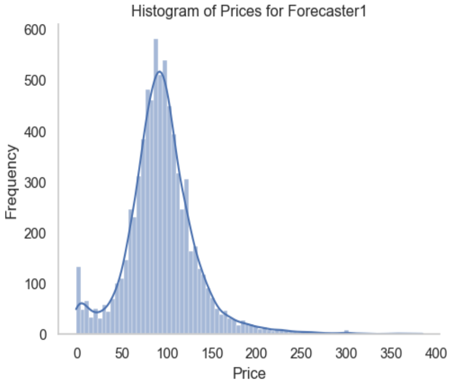
**Hourly TimeSeries Data**

28/07/2023 - 05/07/2024

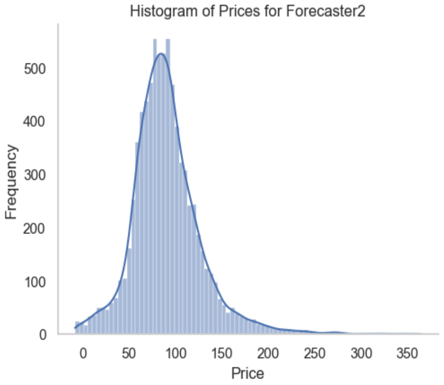
Actual Price



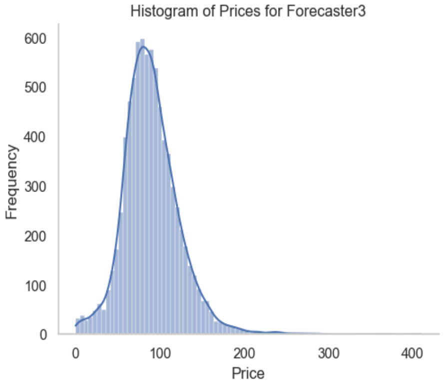

Price Forecaster1




Price Forecaster2




Price Forecaster3




**Preprocessing**

Interpolation

Merging

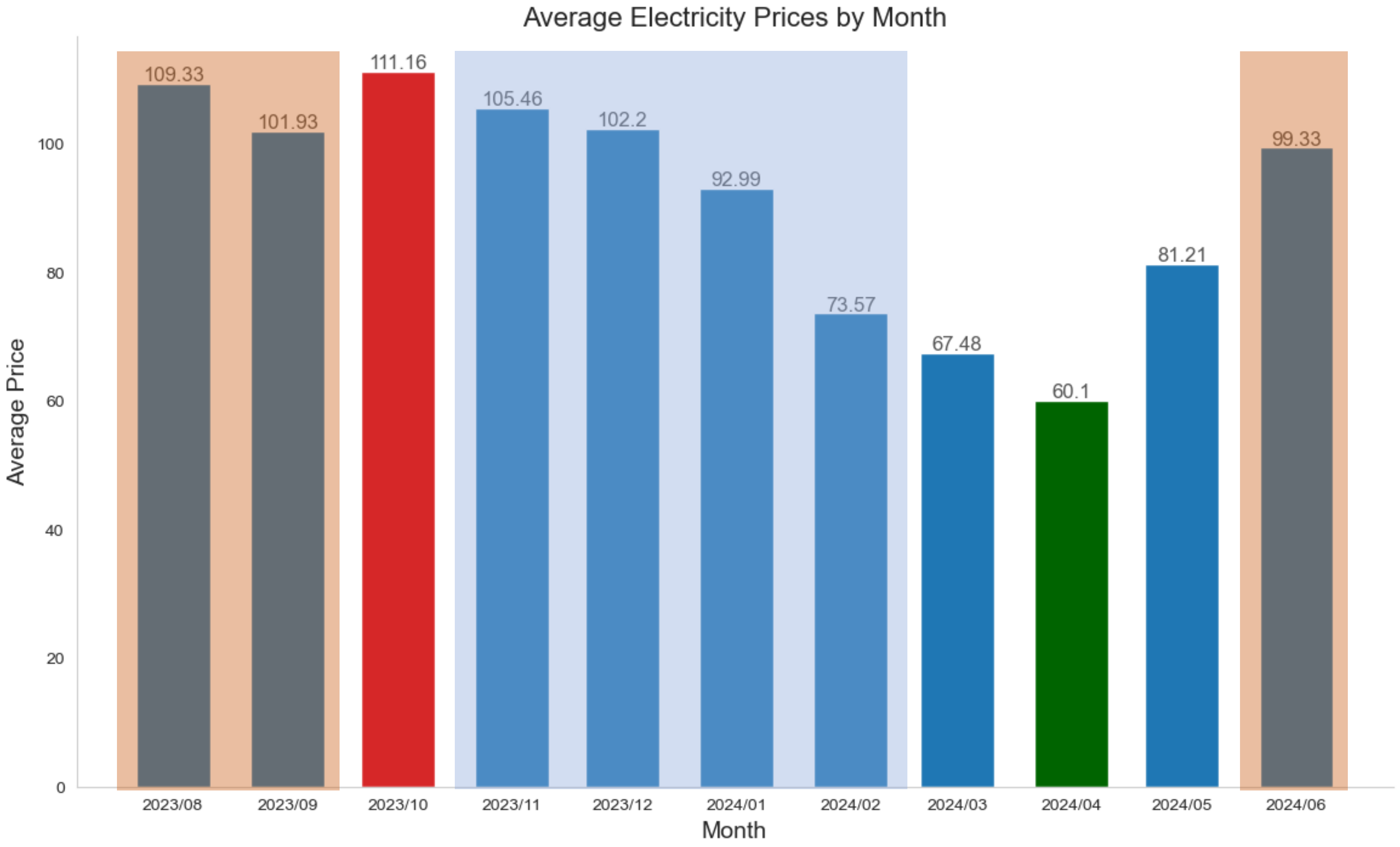
**Feature Engineering**

Temporal Feature Engineering

Cyclical Feature Encoding

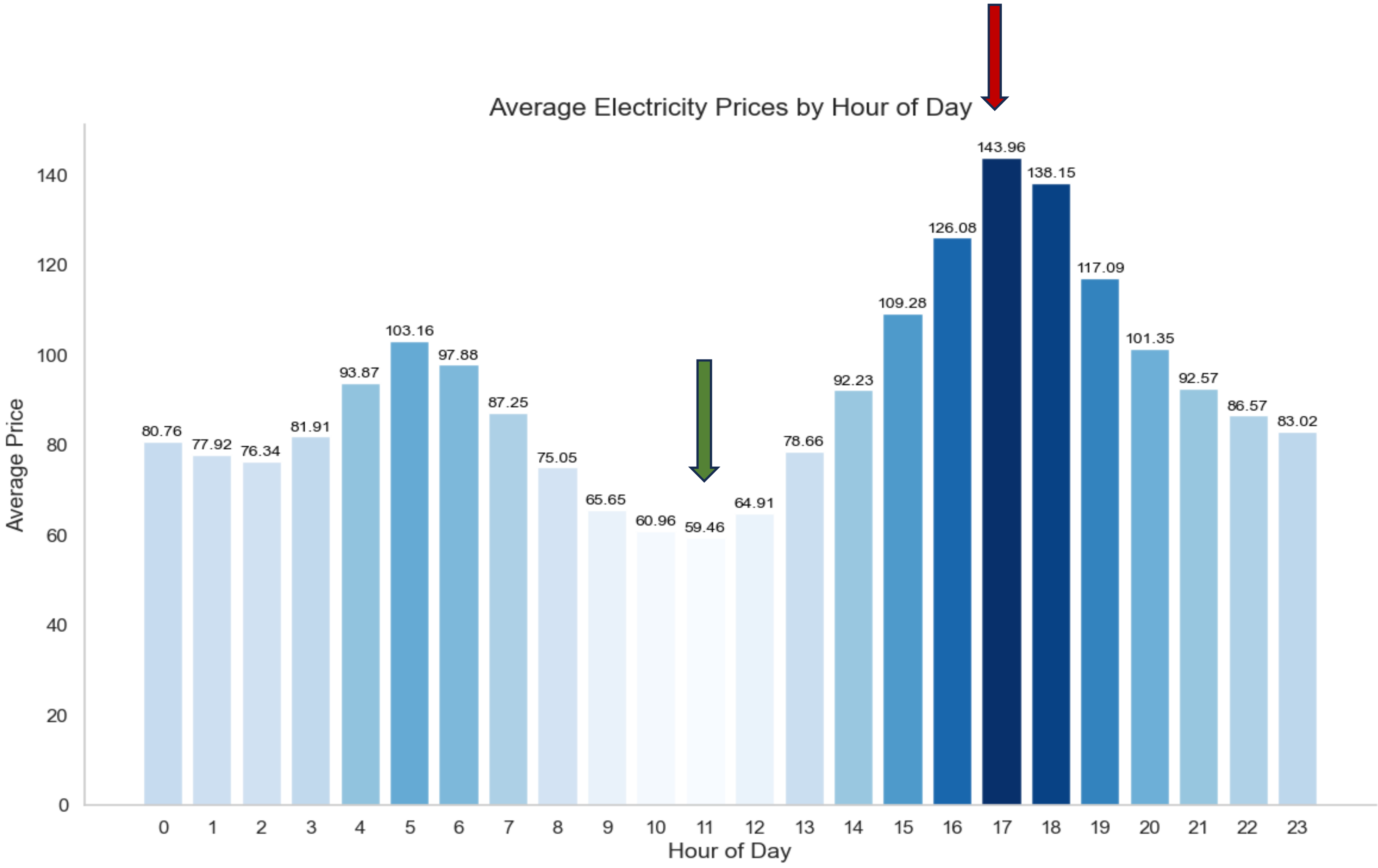
External Feature Implementation

# MONTHLY ANALYSIS



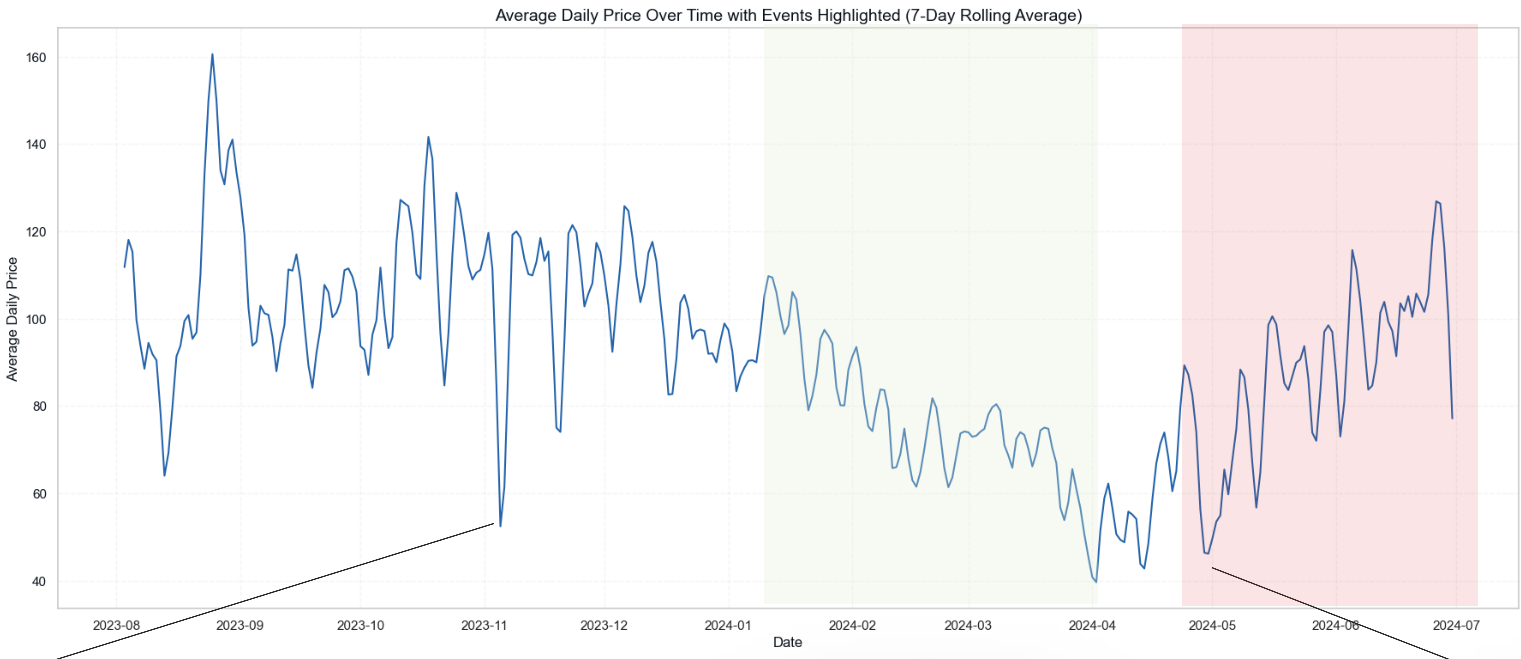

- Decline in prices during the winter months.
- Prices start increasing again in the summer months indicating a seasonal pattern in electricity pricing.

# HOURLY ANALYSIS




- Afternoon is expensive
- Cheapest prices during the peak of Photo Voltaic production

# E N E R G Y   R E L A T E D   E V E N T S   &   P R I C E




Geopolitical tensions and continued sanctions in Russian energy in **November 2023** caused a significant spike in energy prices.

**Jan 2024**: Downward Trend in Electricity prices in Greece. The repercussions of the crisis appear to be diminishing.

**May 2024**: Electricity prices in Greece increased by 41%. Decline in renewable energy production. Unfavorable weather conditions( African dust, cloud cover)

# A G E N D A

1 Present the Business Need and the Question Posed

2 Description of the Dataset provided by PPC &
**E**xploratory **D**ata **A**nalysis

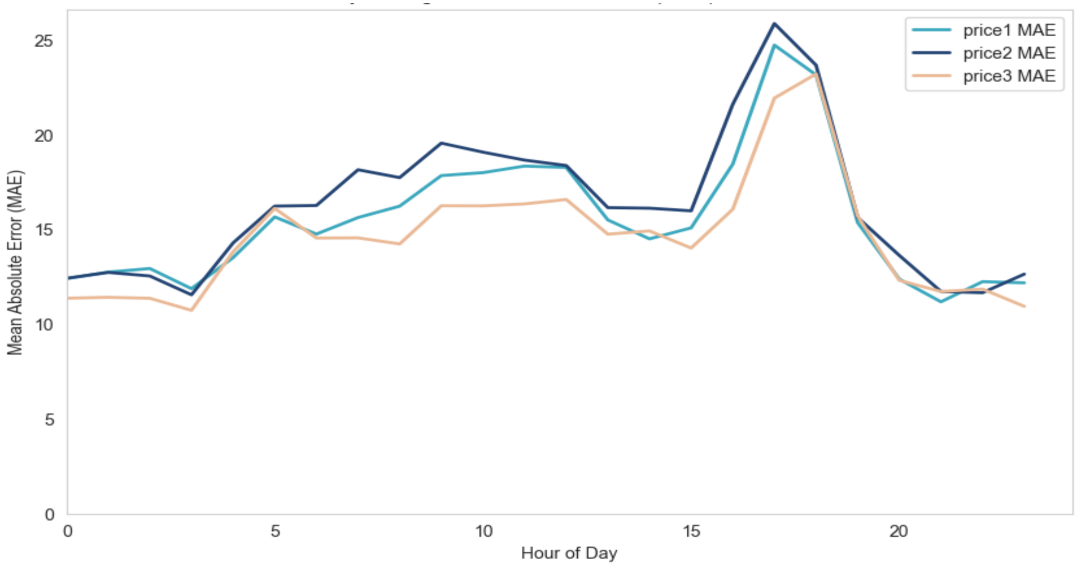
3 Comparative Analysis of Forecasters

4 Forecasting the Best Next Price
(Regression & Timeseries analysis)

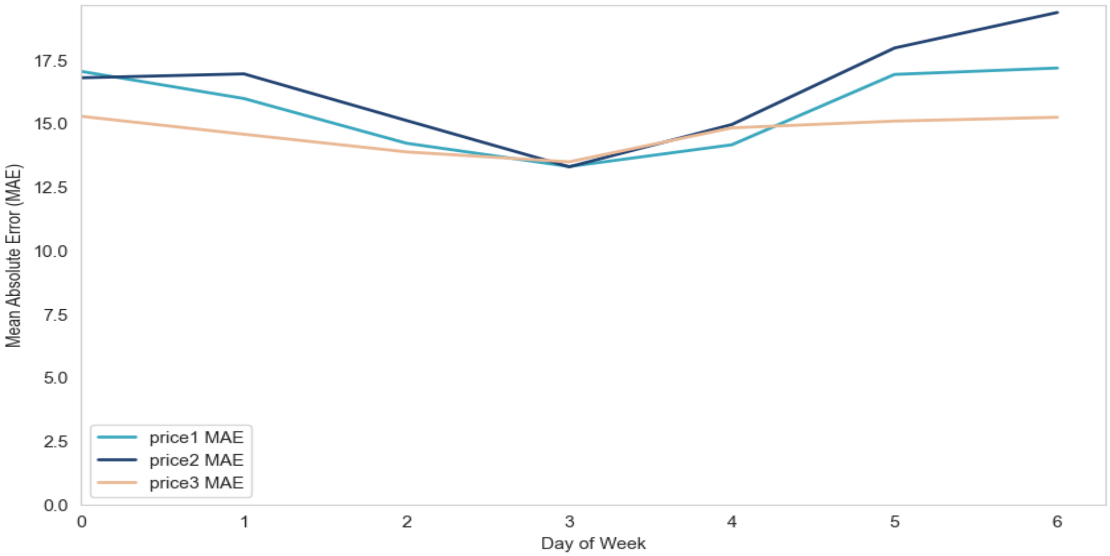
5 Conclusions

# FORECASTER MEAN DEVIATIONS (MAE) OVER TIME INTERVALS

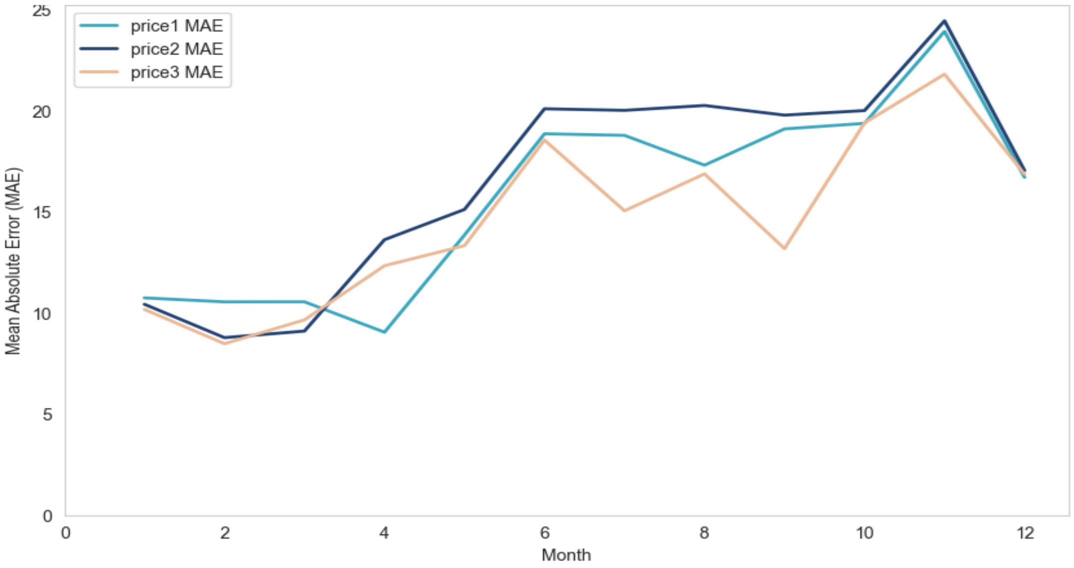
Hour of Day




Day of Week




Month




- **Similar** trends across all the time intervals
- **Monthly Fluctuations:** Forecaster 3 diverts from the rest exhibiting lower deviations from the actual price (months 6 to 10)
- **Forecaster 3** is consistently the best performer

# ERROR COMPARISONS : FORECASTERS

| | Forecaster1 | Forecaster2 | Forecaster3 |
|---|---|---|---|
| MAE | 15.53 | 16.33 | **14.61** |

# A G E N D A

1 Present the Business Need and the Question Posed

2 Description of the Dataset provided by PPC & **E**xploratory **D**ata **A**nalysis

3 Comparative Analysis of Forecasters

4 Forecasting the Best Next Prices (Regression | Timeseries | RNN Analysis)

5 Conclusions

# MACHINE LEARNING

| Regressors | Timeseries | Recurrent Neural Networks |
| --- | --- | --- |
| ○ Linear Regression<br>○ XG Boost<br>○ Adaptive Boost<br>○ Light GBM | ○ Arima<br>○ Prophet | ○ LSTM (Long-Short Term Memory)<br>○ GRU (Gated Recurrent Unit) |

# Mean Absolute Error

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

# MACHINE LEARNING

## Regressors

- Linear Regression
- XG Boost
- Adaptive Boost
- Light GBM

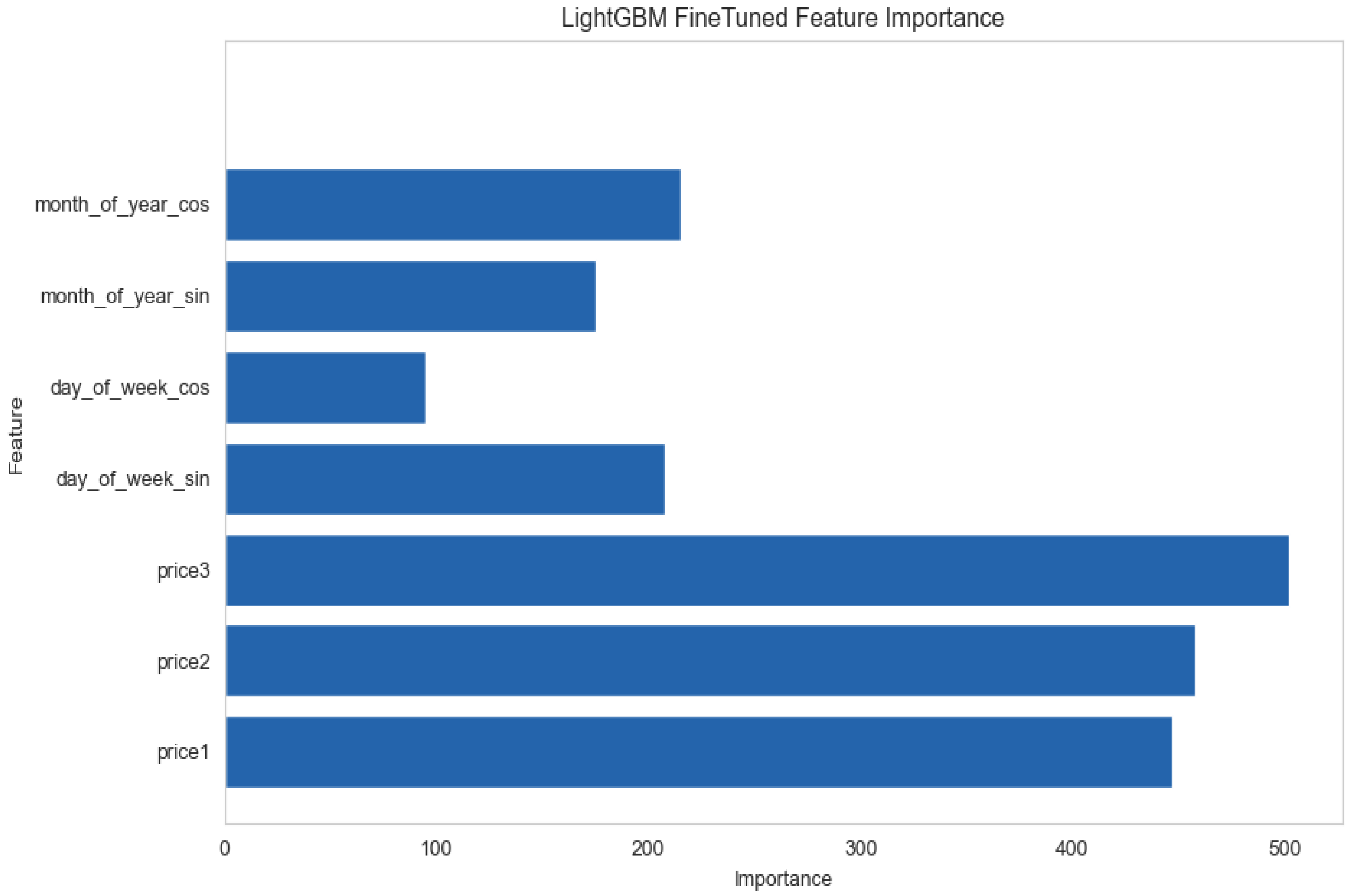
## Timeseries

- Arima
- Prophet

## Recurrent Neural Networks

- LSTM
- GRU

# REGRESSORS: FINE TUNED LIGHT GBM




**EXTRA FEATURE IMPLEMENTATION**

- Temporal Feature Engineering
- Cyclical Feature Encoding
- Light GBM shows dependency on DAY and MONTH

| | Light GBM |
|---|---|
| MAE | 11.87 |

# MACHINE LEARNING

| Regressors | Timeseries | Recurrent Neural Networks |
| --- | --- | --- |
| Linear Regression | Arima | LSTM |
| XG Boost | Prophet | GRU |
| Adaptive Boost | | |
| Light GBM | | |

**ARIMA vs Prophet**

- SARIMA through Auto-ARIMA
- Prophet Grid Search Hyper-Tune
- Very Similar Results
- Not Good Enough

| | Forecaster1 | Forecaster2 | Forecaster3 | ARIMA | Prophet |
|---|---|---|---|---|---|
| MAE | 15.53 | 16.33 | 14.61 | 14.60 | 14.46 |

# MACHINE LEARNING

| Regressors | Timeseries | Recurrent Neural Networks |
| --- | --- | --- |
| ○ Linear Regression<br>○ XG Boost<br>○ Adaptive Boost<br>○ Light GBM | ○ Arima<br>○ Prophet | ○ LSTM<br>○ GRU |

# LONG SHORT – TERM MEMORY (LSTM) | GATED RECURRENT UNIT (GRU)

Batch Size, Epochs
-0.7
MAE~12

Lagged Observations
1 – 24 – 48 Hours
-3.1
MAE~13

Network Architecture
The Simplest the Better
-3
MAE~15

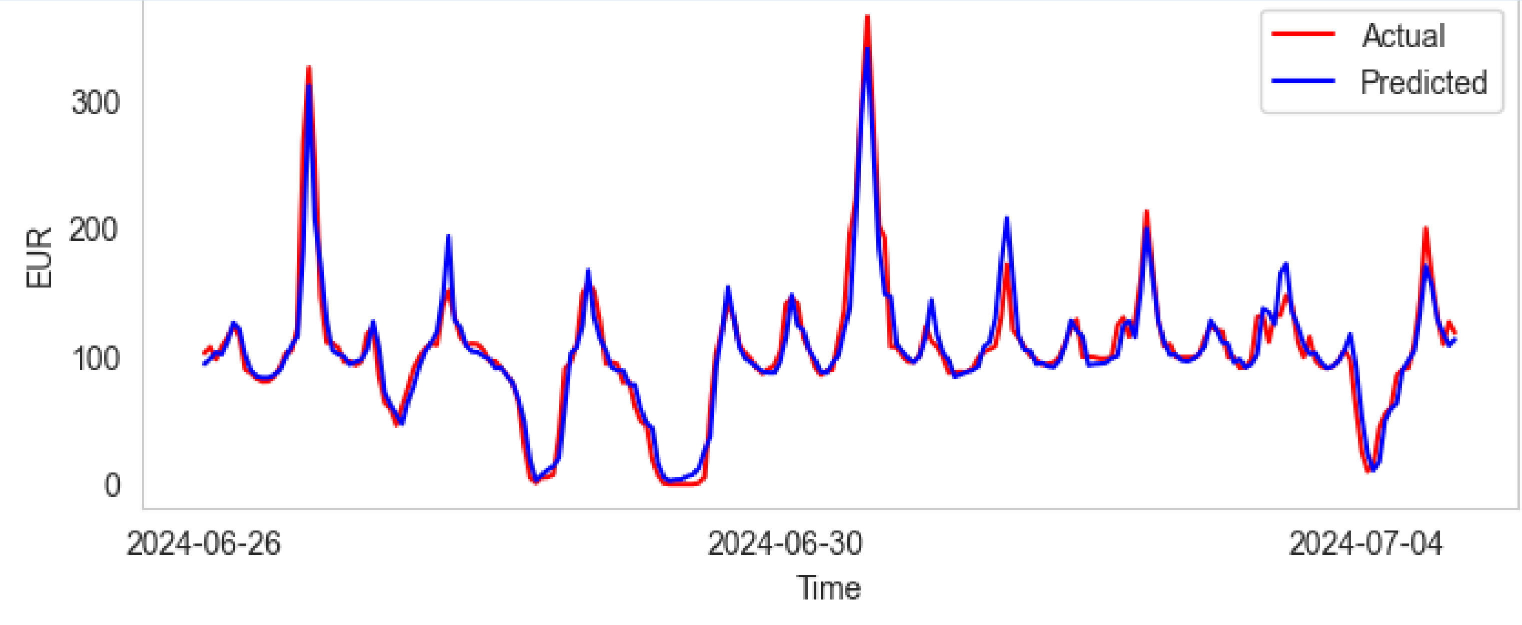
Multivariate Analysis
(Added Temporal Features)
-0.3
MAE~11.9

Feature Engineering
( Cyclical Feature Encoding)
-2
MAE~10.2

Hyper-Parameter Tuning
-0.28
MAE~9.8 *

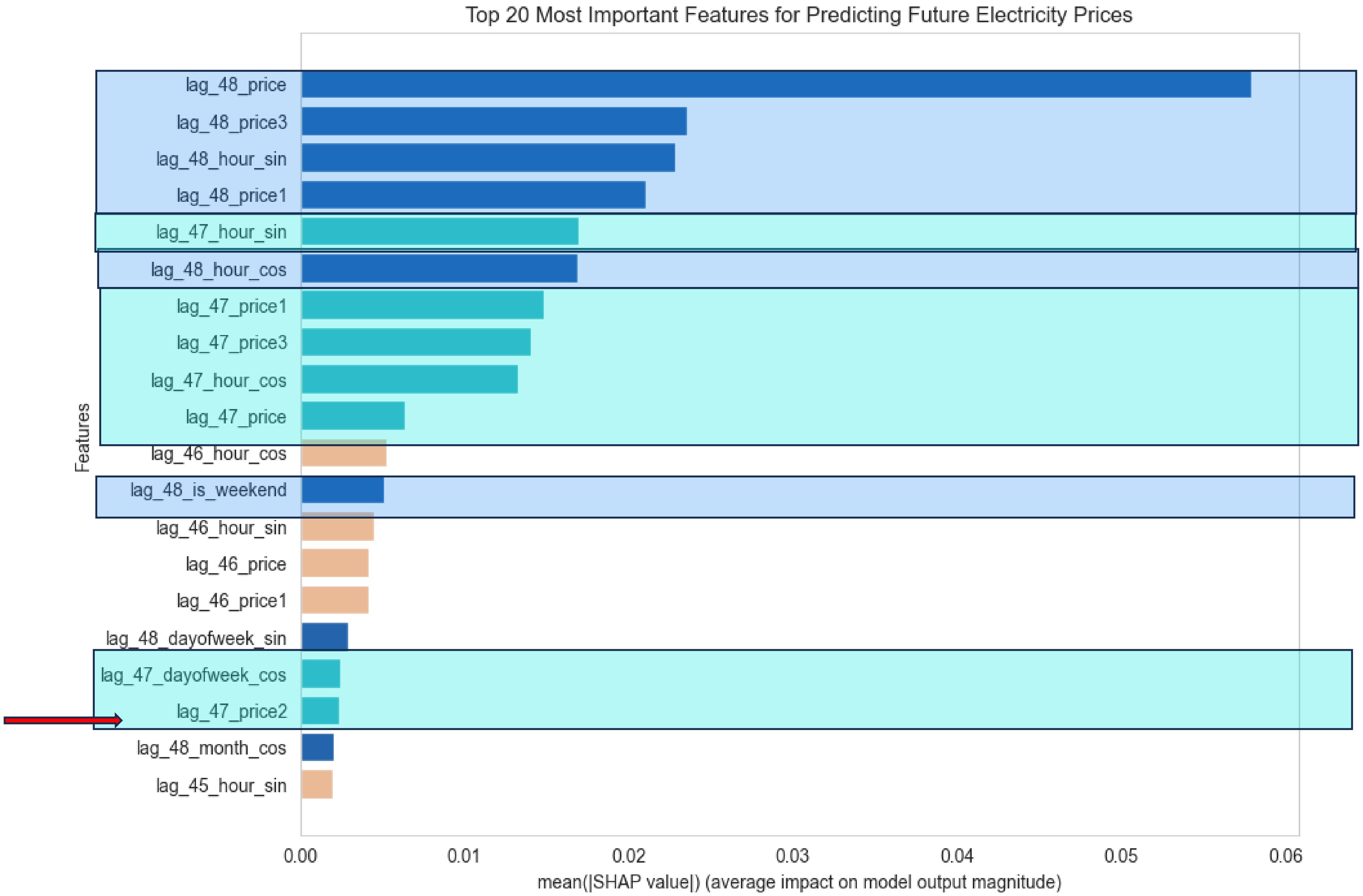
***BEST LSTM Model :[ SingleStacked** : 32 neurons**, n_lags** = 48 (**cyclical feature enconding** on hour of day, day of week, month of year) (**Additioanal Features**: Public Holidays, is_weekend) , (**batch_size** = 32, **epochs** = 40, Activation: **'elu'**, optimizer = 'adam'), ]

# RNN: GRU BEST MODEL PREDICTIONS vs. ACTUALS




| | GRU |
|---|---|
| **MAE** | 9.78 |

# GRU BEST MODEL - FEATURE DEPENDENCIES



# FORECASTERS | REGRESSORS | TIMESERIES | RNN

## C O M P A R I S O N

| | Forecaster 1 | Forecaster 2 | Forecaster 3 | Light GBM | LSTM | GRU |
|---|---|---|---|---|---|---|
| MAE | 15.53 | 16.33 | 14.61 | **11.87** | **9.84** | **9.78** |

neptune.ai

# A G E N D A

1. Present the Business Need and the Question Posed
2. Description of the Dataset provided by PPC & **E**xploratory **D**ata **A**nalysis
3. Comparative Analysis of Forecasters
4. Forecasting the Best Next Price (Regression | Timeseries | RNN Analysis)
5. Conclusions

# CONCLUSIONS

- We can provide an algorithm designed to deliver more precise energy price predictions by leveraging forecasts from the three different forecasters.

- Further research could involve incorporating additional features that include lagged observations, such as weather and seasonal data, to enhance the robustness of our analysis.

# T H A N K  Y O U

## KATERINA PSALLIDA

katherinepsallida@gmail.com

## KOSTAS TSAGKAROPOULOS

kotsagaro@gmail.com

# A P P E N D I X

## Articles

- Geopolitical Tensions Affecting Supply (November 2023) https://shorturl.at/4i07t
- Gas Supply drop due to Infrastructure Issues (February 2024) https://shorturl.at/pMGW7
- Temperature Data Acquisition: https://www.visualcrossing.com/weather/weather-data-services
- April Temperature Decrease: https://www.sunsave.energy/blog/april-2024-energy-price-cap
- Online Application to Log AI Experiment Results : https://app.neptune.ai/
- Should you use lagged features in LSTM : https://medium.com/@dclengacher/lstms-with-lagged-data-cc03a3a8cfcf

## Best Model Parameters after Hyper Parameter Tuning

- Best **Light GBM** Best GRU parameters: [**colsample_bytree**: 0.8, **learning_rate**: 0.1, **max_depth**: 11, **min_child_samples**: 10, **n_estimators**: 70, **num_leaves**: 31, **subsample**: 0.1]
- Best **ARIMA**: [**p**: 0, **q**: 1, **d**: 0 (cyclical feature enconding on day of week, month of year)]
- Best **Prophet** parameters: [**changepoint_prior_scale**: 0.01, **seasonality_mode**: additive, **seasonality_prior_scale**: 10, **changepoint_range**: 0.9, **weekly_seasonality**: True(cyclical feature enconding on day of week, month of year) ]
- Best **LSTM** parameters: [(**MinMaxScaler**(0,1)) (**SingleStacked, OneDense Output**, **n_lags** = 48, **batch_size** = 32, **epochs** = 40, Activation: **'elu'**, optimizer = 'adam', **units**: 32 (neurons)) (**cyclical feature enconding** on hour of day, day of week, month of year) (**Additioanal Features**: Public Holidays, is_weekend)]
- Best **GRU** parameters: [(**MinMaxScaler**(0,1)) (**Two Consecutive Stacks** + **OneDense**(25, linear activation), **OneDense**(1, linear activation) **n_lags** = 48, **batch_size** = 32, **epochs** = 40, **Activation1**: 'elu', **Activation2**:'relu', **optimizer** = 'adam', **units1**: 64 (neurons),**units2**: 128 (neurons)) (**cyclical feature enconding** on hour of day, day of week, month of year) (**Additioanal Features**: Public Holidays, is_weekend)]