UNIVERSITÀ DI TRENTO

A forecasting system for electricity production and customer demand

Department of Information Engineering and Computer Science

Master’s Degree in Artificial Intelligence Systems

Supervisors

Elisa Ricci

Daniele Miorandi

Student

Samuel Bortolin

Academic year 2021/2022

Problem Statement

- Global energy crisis caused an **increased volatility** in the electricity market
- Crucial for electricity suppliers to optimize the purchase of the necessary electricity
- Avoid to resort solely to the instantaneous electricity market
- **Plan energy procurement strategies avoiding cost fluctuations and imbalance fees**
- Offer customers the necessary electricity and supply it at a competitive price

Problem Statement

- Vital to forecast **customers' electricity demand**
- Crucial to forecast **production from own electricity generation facilities**
- Understand the consumption habits of individual customers by establishing a reference **consumption baseline**
- Personalized energy solutions (dynamic pricing scheme, energy storage systems, …)

MIWenergía Datasets

- Aggregated consumption data over all their customers, from **2 to 4 thousand** (19 months)
- Production data from **8 photovoltaic plants** (from 2 to 10 months)
- Individual consumption data from **3 customers** (from 9 to 15 months)
- Weather data from Murcia airport, including solar energy data, as additional features

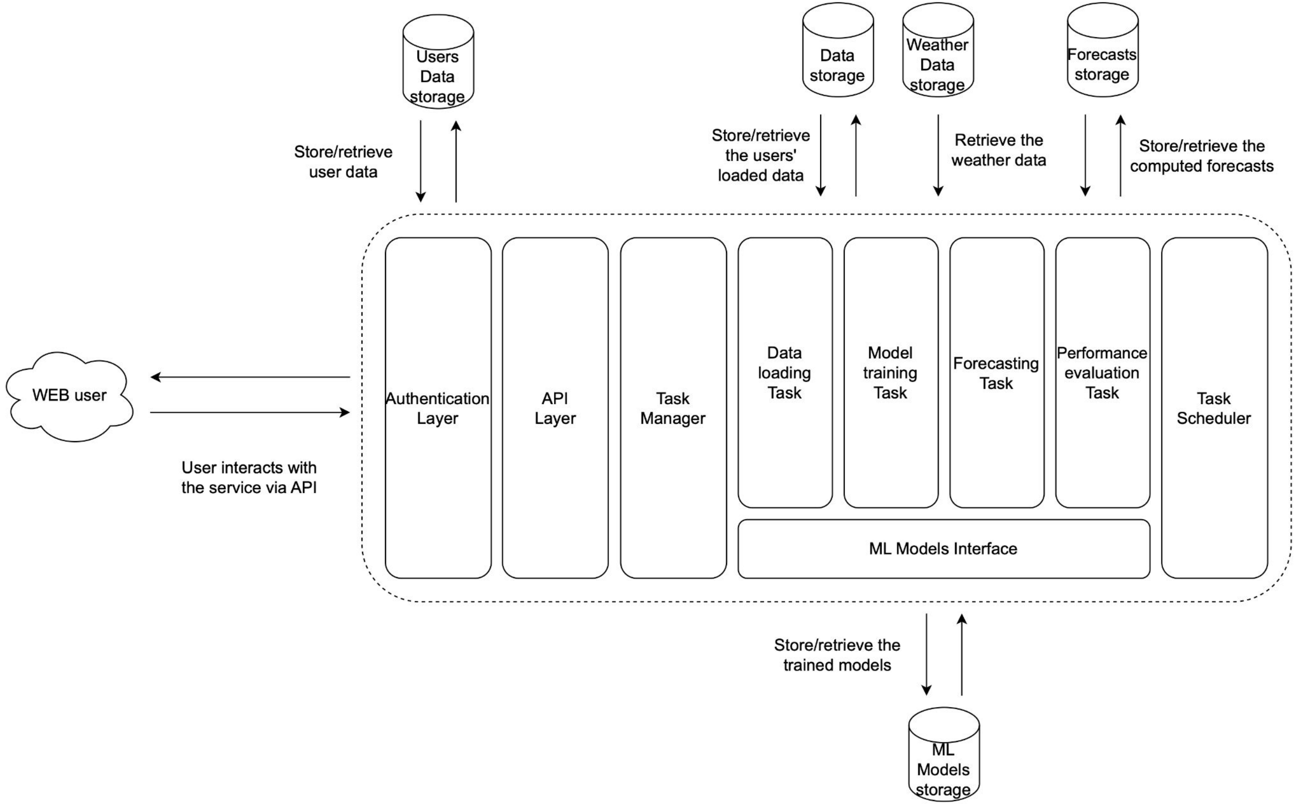
Research Questions

- Can the customers' electricity demand be forecasted based on past aggregated **consumption data over the customers** and historical weather data?
- Can the photovoltaic plants' production be forecasted based on past aggregated **production data over the photovoltaic plants** and historical weather data, including solar energy data?
- Can a consumption baseline of individual customers be established based on past **consumption data of individual customers** and historical weather data?

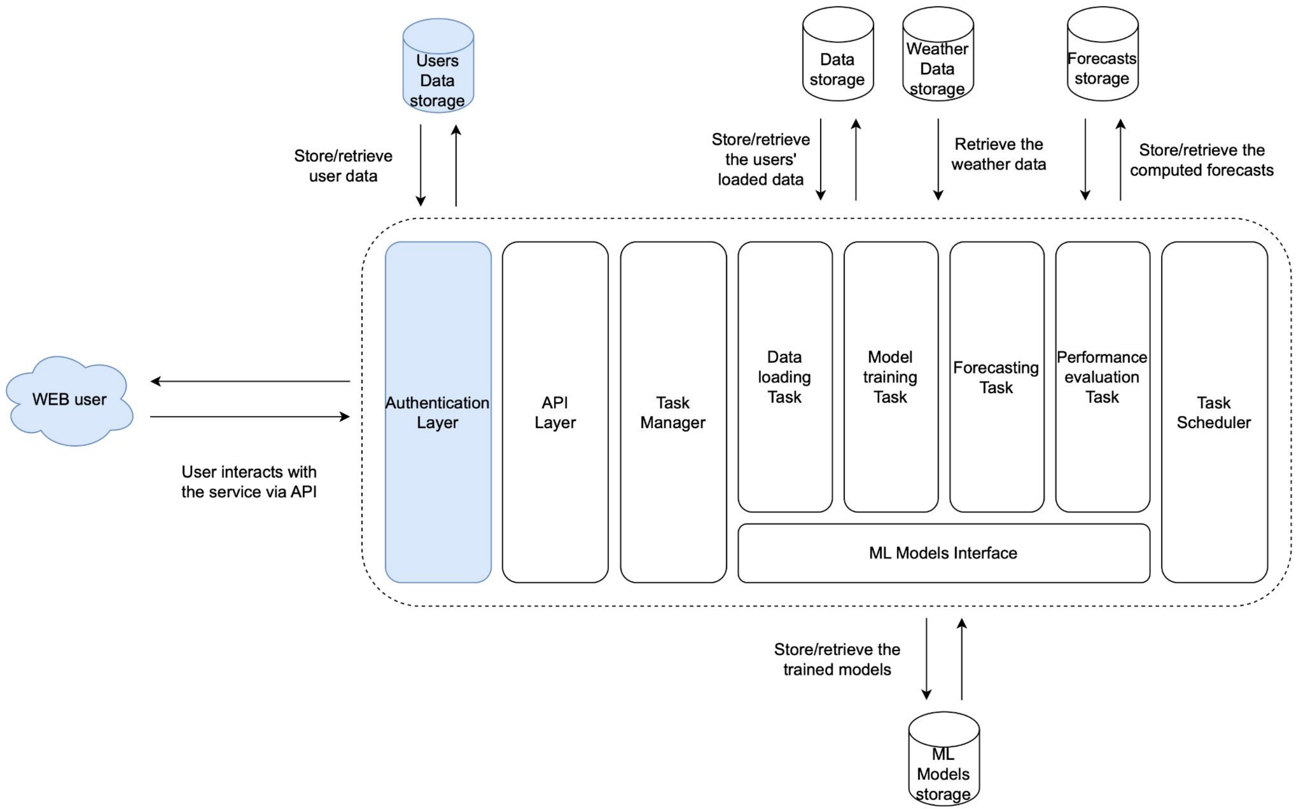
Achievements

- Designed a **novel forecasting system** for addressing the research questions
- Implemented a **prototype** of the system with a focus on key components
- Comprehensively evaluated the performance of a set of **developed models** for :
  - Electricity demand forecasting
  - Electricity production forecasting
  - Consumption baseline forecasting

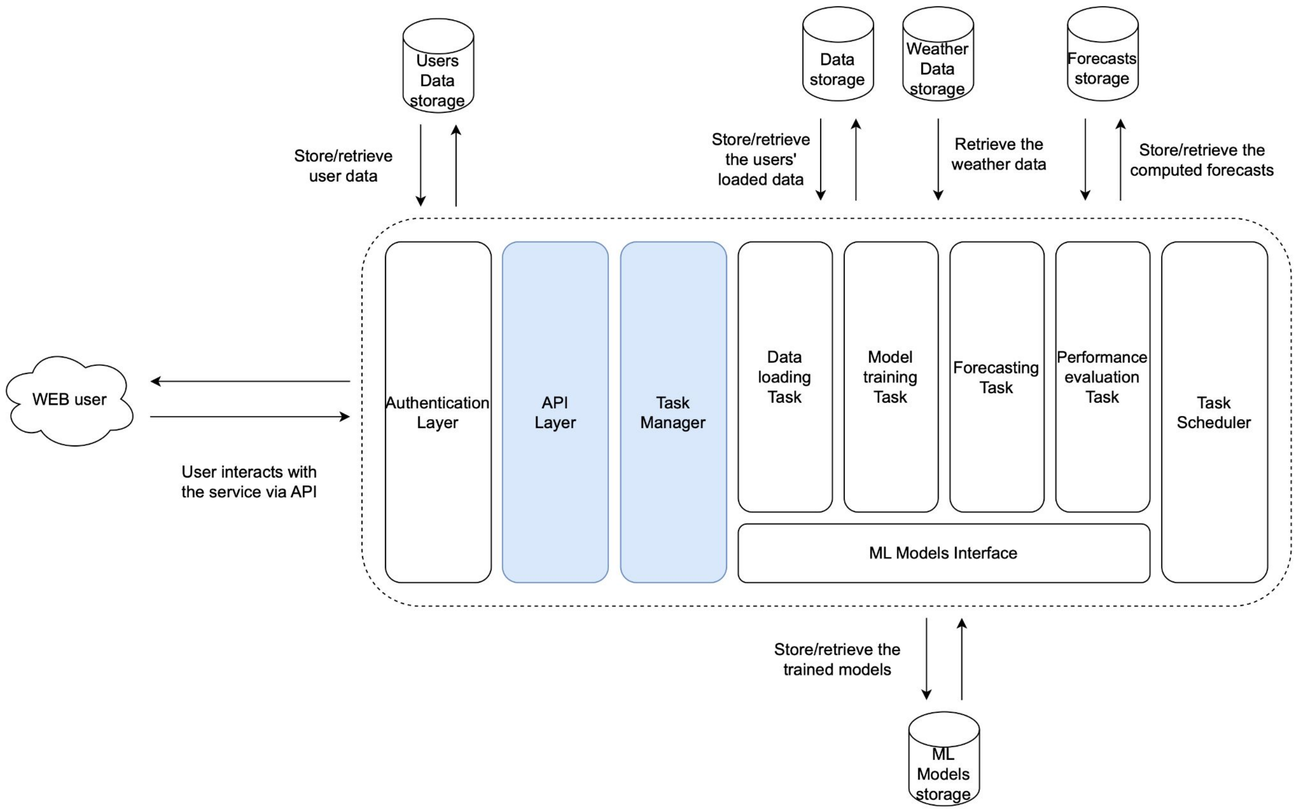
Forecasting System



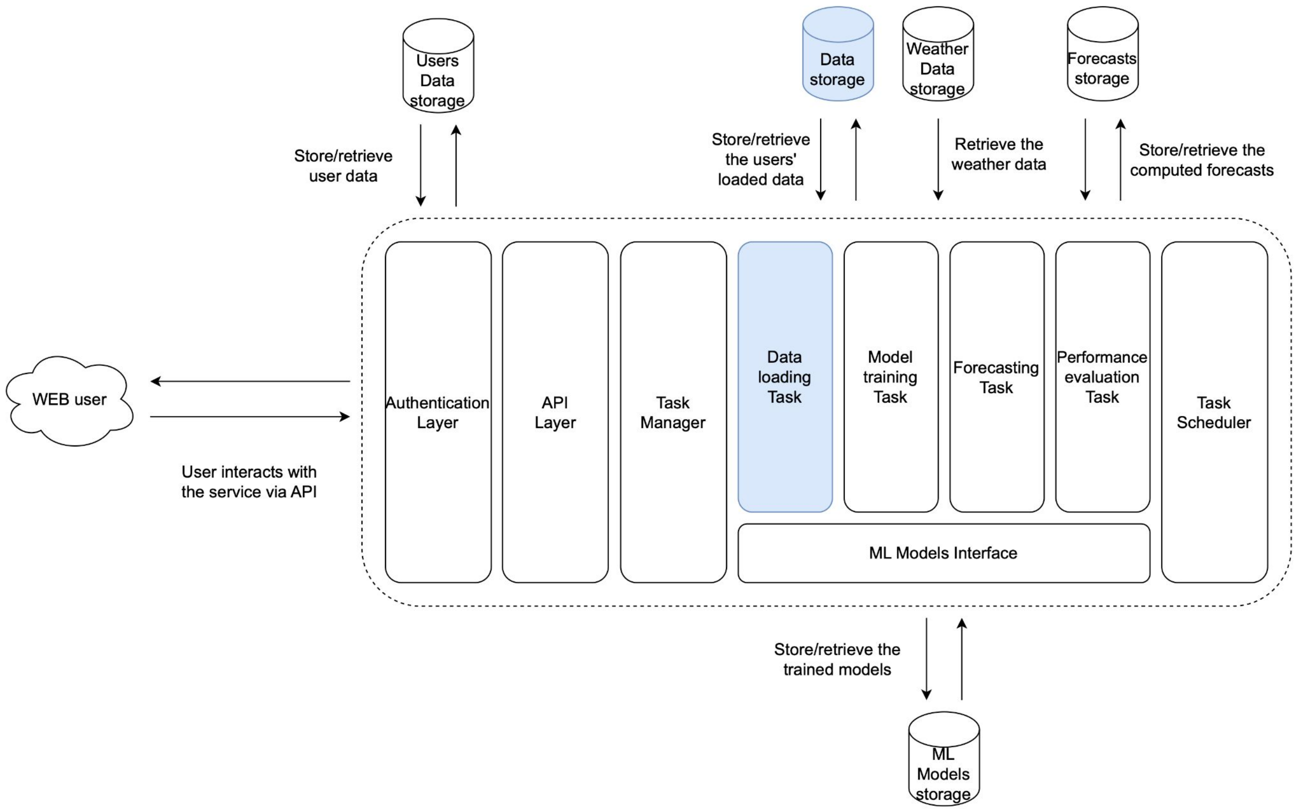
Forecasting System



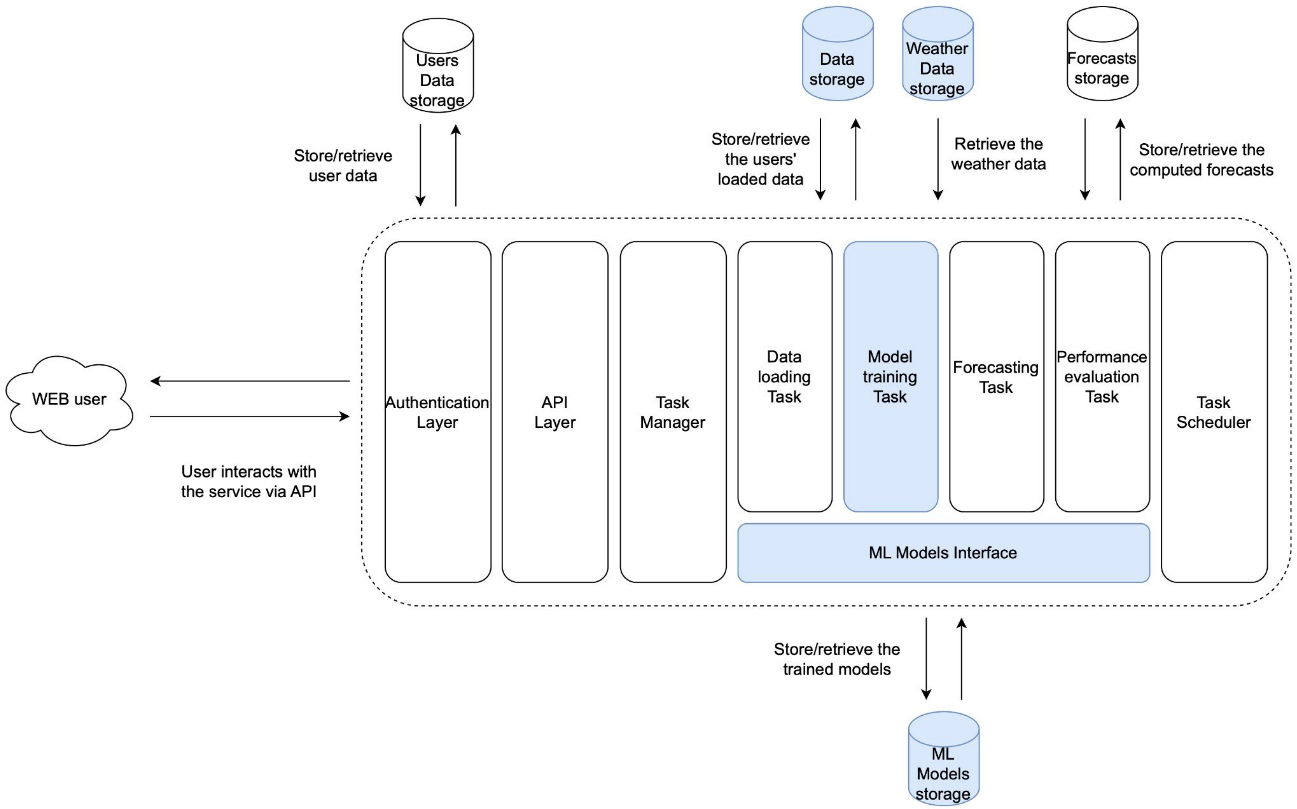
Forecasting System



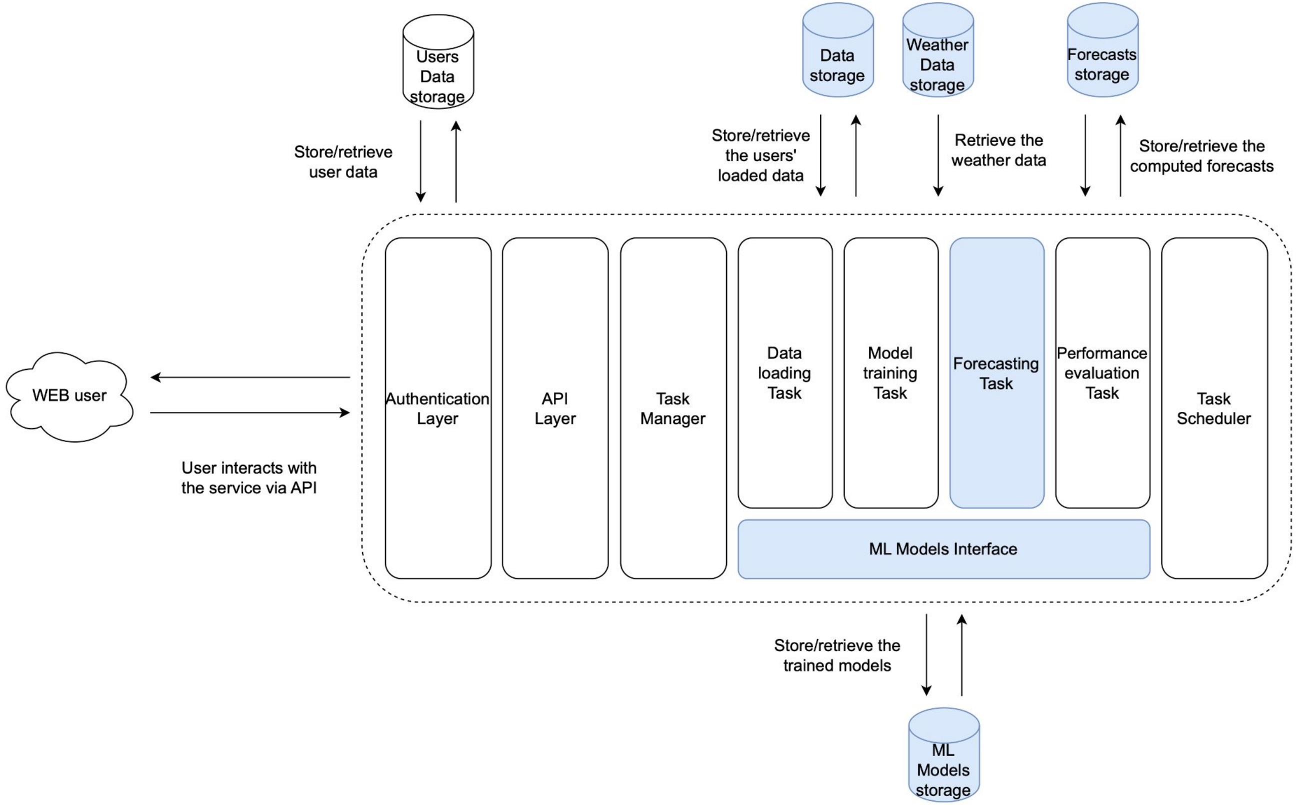
Forecasting System



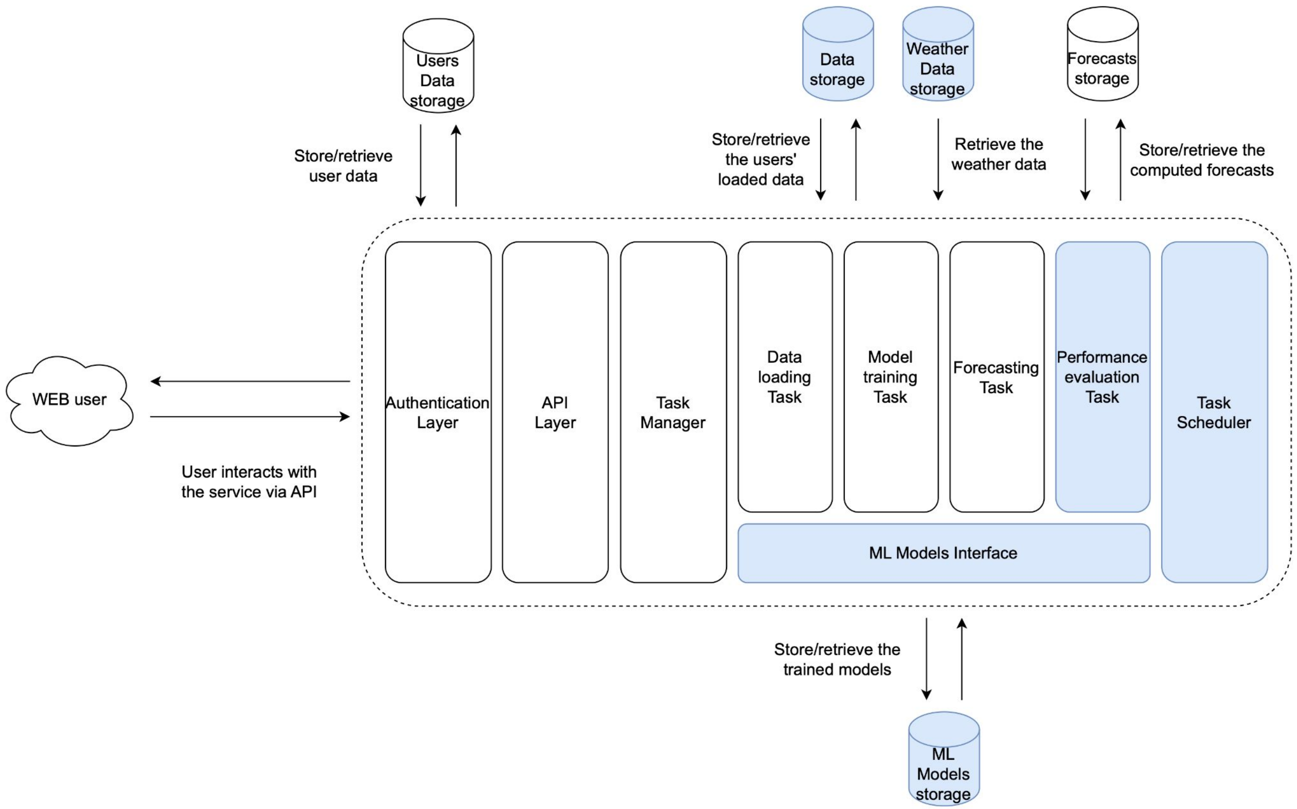
Forecasting System



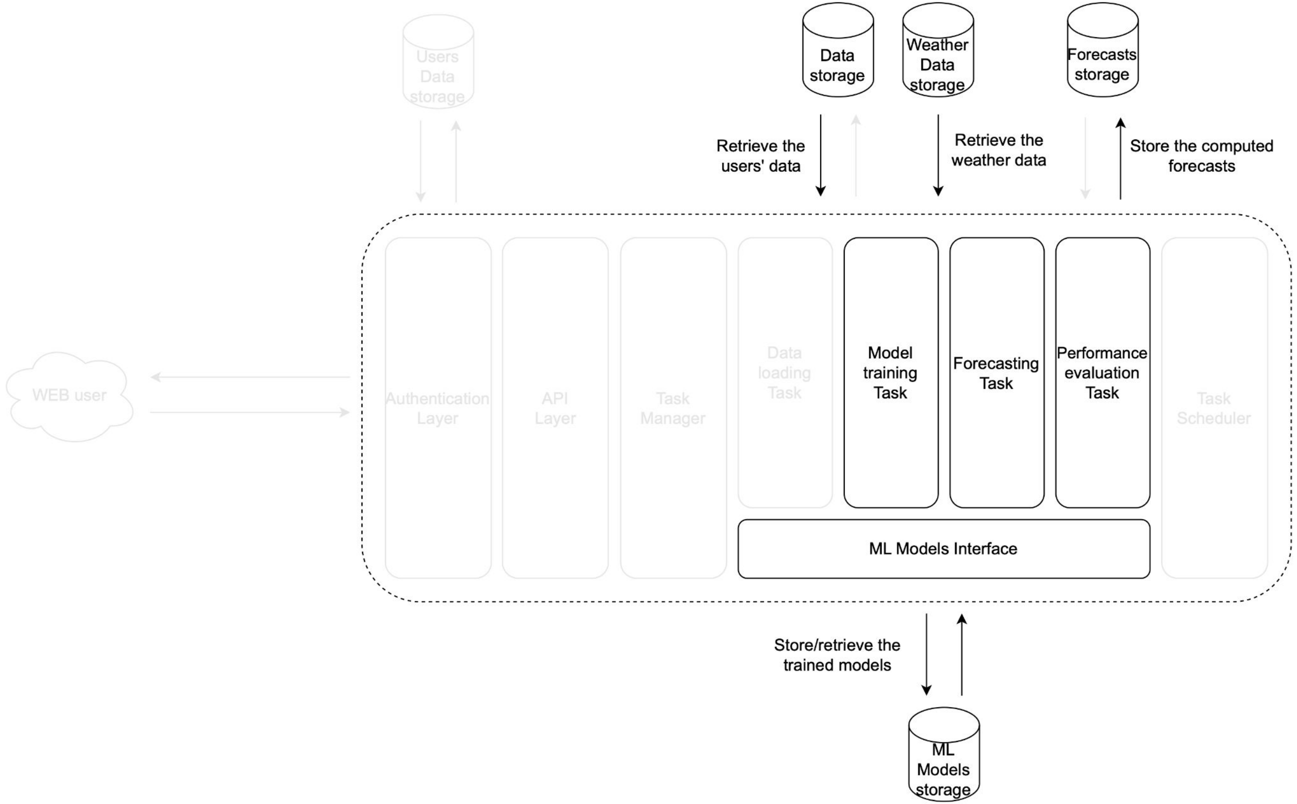
Forecasting System



Forecasting System



Implemented Prototype



Models

- Baselines: combination of previous days or weeks
- Statistical: ARIMA/SARIMA
- ML: SVR, Hist gradient boosting regressor, XGBoost regressor, and Prophet
- DL: LSTM, GRU, CNN, and TFT
- AutoML
- Combinations of Baselines, Prophet, and DL models

Performance Evaluation

MAPE and MAE as error metrics

Blocked k-fold cross-validation

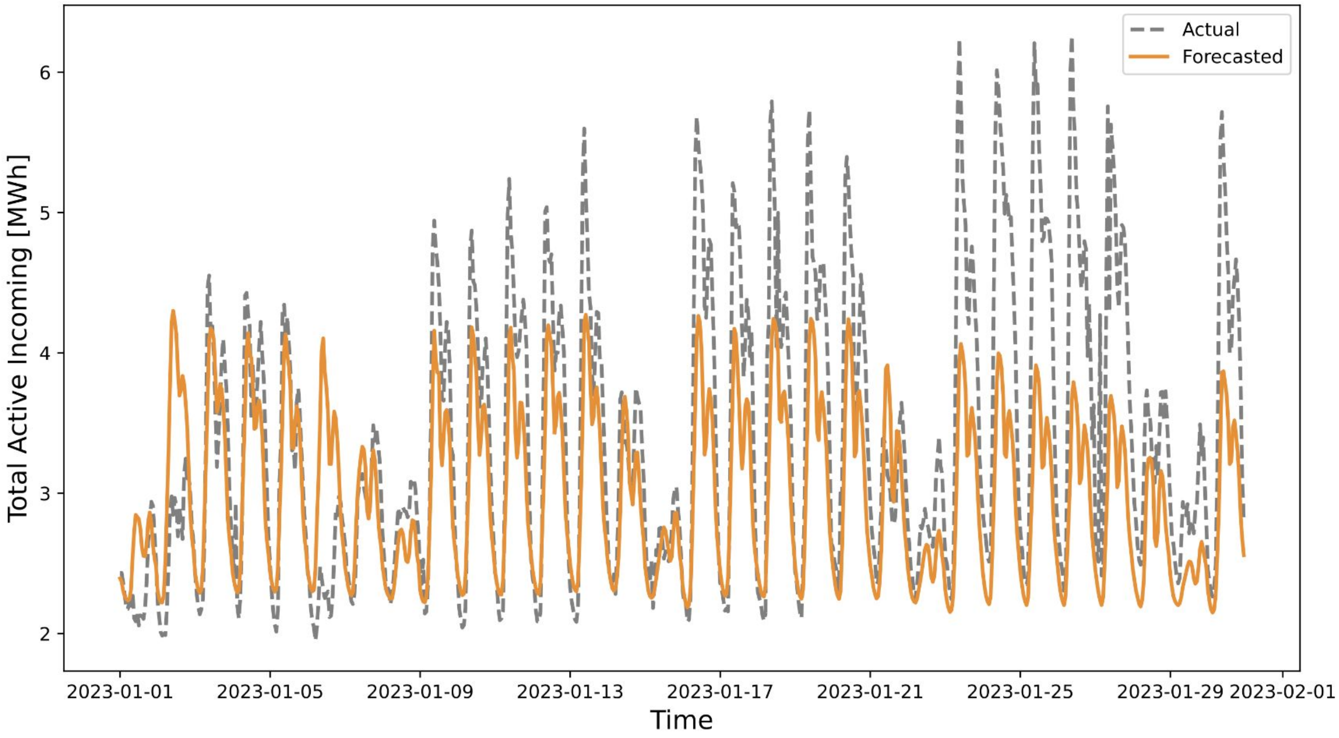
Test on the **last split**

Hourly Results

k = 12
Split size = 30 days

Model	Blocked k-fold cross-validation MAPE	Test on the last split MAPE
TFT	15.0 ± 4.3	14.8
Best combination: GRU + One Week Baseline	/	14.8
GRU	30.8 ± 14.5	16.8
AutoML	/	17.0
CNN	34.6 ± 15.8	17.5
One Week Baseline	12.9 ± 4.3	17.6
SARIMA	15.4 ± 4.6	18.5
One Day Baseline	22.9 ± 7.1	20.5
LSTM	34.1 ± 11.9	23.2
Hist gradient boosting regressor	14.5 ± 5.5	29.3
SVR	61.8 ± 22.6	29.5
Prophet	28.7 ± 7.1	31.6
XGBoost regressor	16.9 ± 8.9	37.8

TFT Forecasts

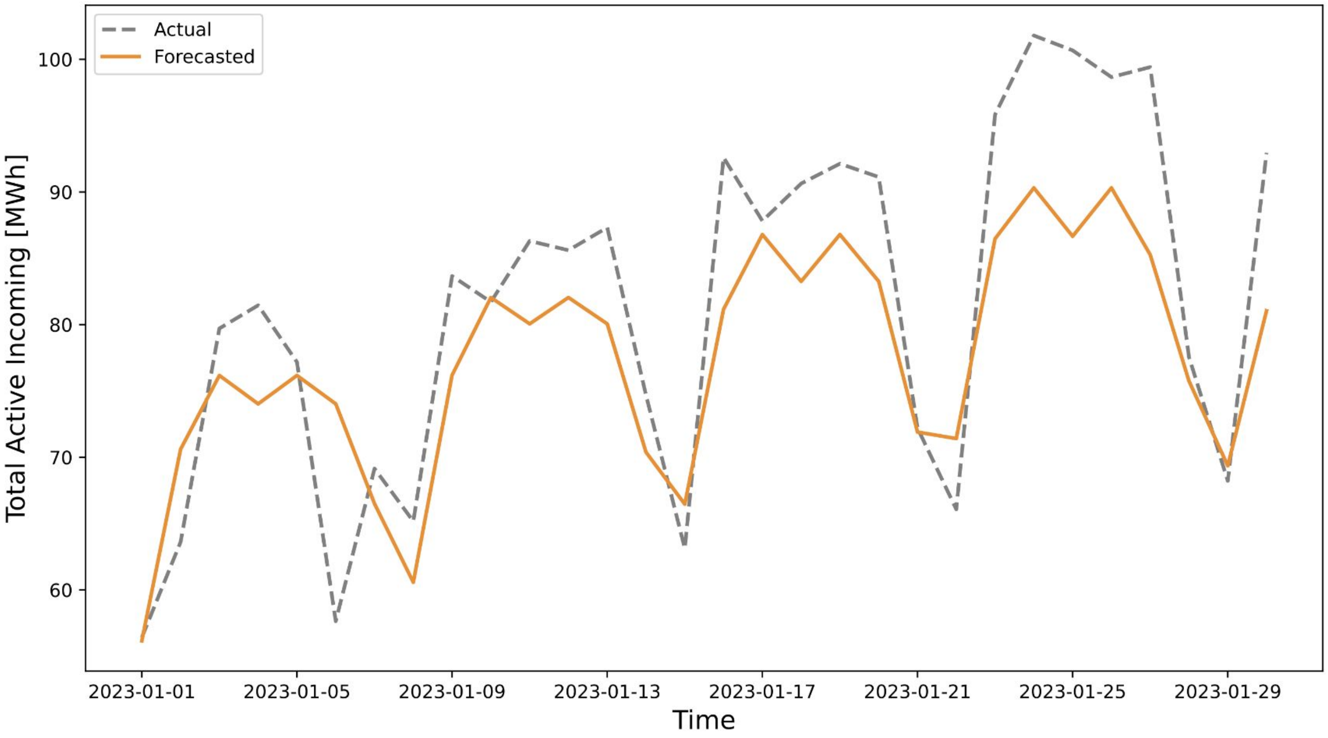


Daily Results

k = 12
Split size = 30 days

Model	Blocked k-fold cross-validation MAPE	Test on the last split MAPE
CNN	50.6 ± 36.8	7.5
XGBoost regressor	13.0 ± 4.4	7.7
Hist gradient boosting regressor	13.0 ± 5.5	9.8
Best combination: CNN + One Week Baseline	/	12.3
LSTM	29.6 ± 13.8	14.5
TFT	13.7 ± 5.1	14.9
One Week Baseline	11.5 ± 4.8	18.4
SARIMA	11.9 ± 7.0	19.6
GRU	28.8 ± 29.3	20.5
AutoML	/	21.0
One Day Baseline	22.8 ± 8.2	21.5
Prophet	19.1 ± 10.8	26.3
SVR	53.6 ± 22.4	39.8

CNN Forecasts

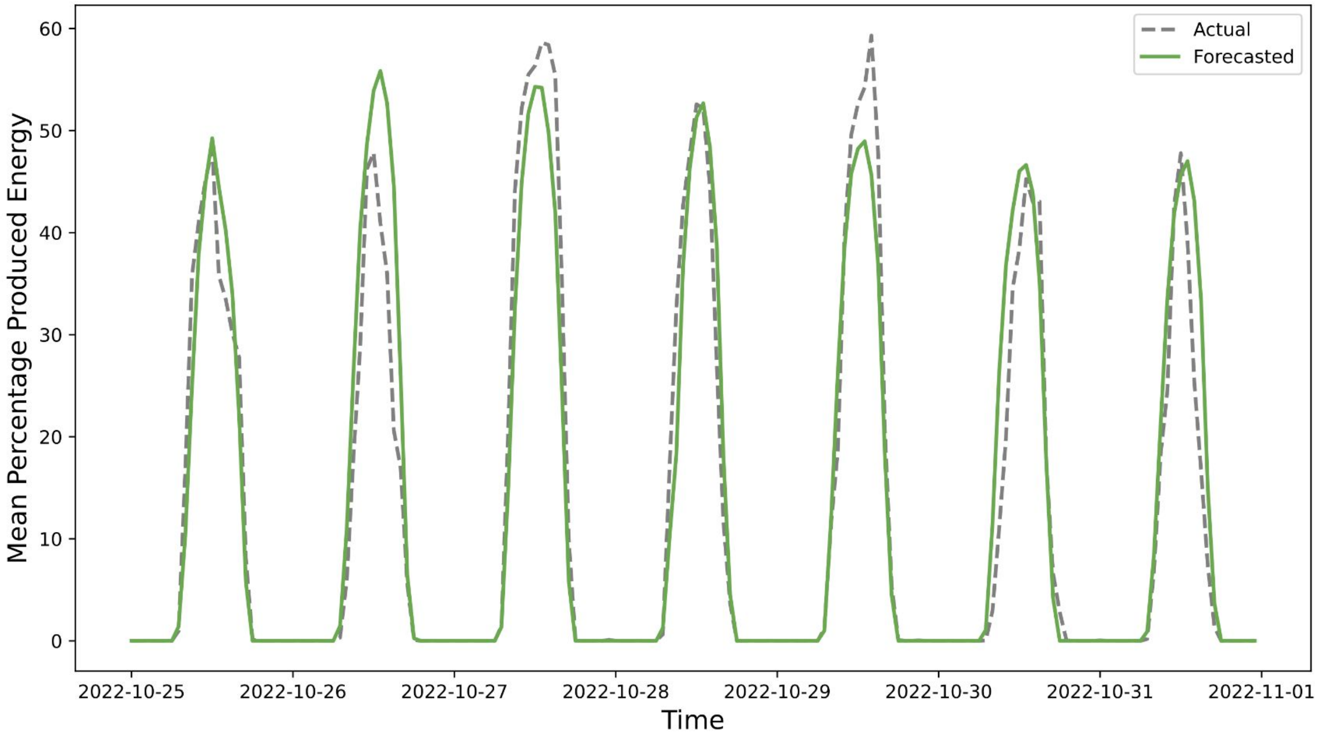


Hourly Results

k = 12
Split size = 7 days

Model	Blocked k-fold cross-validation MAE normalized on PV plants' nominal power	Test on the last split MAE normalized on PV plants' nominal power
GRU	5.7 ± 1.8	2.8
Hist gradient boosting regressor	4.2 ± 1.0	3.0
SVR	5.2 ± 1.1	3.1
CNN	6.0 ± 1.5	3.3
Best combination: LSTM + One Day Baseline	/	3.4
XGBoost regressor	4.6 ± 1.0	3.4
LSTM	5.2 ± 1.5	3.5
TFT	4.7 ± 1.4	3.6
ARIMA	5.1 ± 1.4	3.6
Prophet	7.9 ± 1.6	4.4
One Day Baseline	6.8 ± 3.1	4.6
AutoML	/	5.9

GRU Forecasts



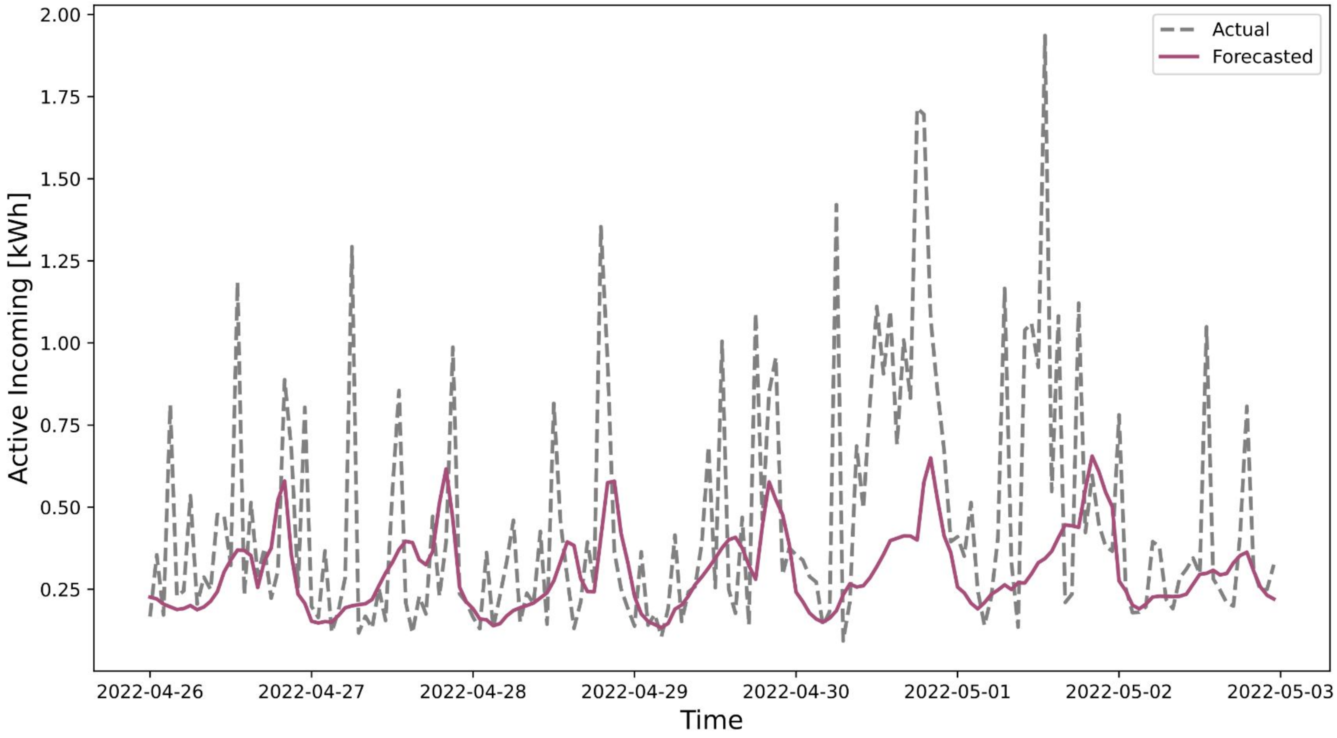
Hourly Results

k = 12
Split size = 7 days

Model	Blocked k-fold cross-validation MAPE — MAE [kWh]	Test on the last split MAPE — MAE [kWh]
TFT	47.3 ± 5.6 — 0.29 ± 0.03	44.2 — 0.23
Hist gradient boosting regressor	54.0 ± 7.0 — 0.26 ± 0.02	50.4 — 0.23
SVR	60.3 ± 6.8 — 0.28 ± 0.03	60.4 — 0.24
SARIMA	74.2 ± 10.0 — 0.27 ± 0.02	78.0 — 0.24
XGBoost regressor	73.7 ± 9.5 — 0.28 ± 0.02	68.1 — 0.25
Prophet	83.8 ± 10.2 — 0.28 ± 0.02	81.9 — 0.25
GRU	58.9 ± 19.6 — 0.31 ± 0.04	42.6 — 0.26
LSTM	111.9 ± 212.1 — 0.46 ± 0.52	44.2 — 0.27
4 Week Baseline	74.1 ± 12.5 — 0.28 ± 0.03	79.1 — 0.27
12 Week Baseline	74.3 ± 13.7 — 0.27 ± 0.02	89.2 — 0.28
One Week Baseline	82.7 ± 12.8 — 0.33 ± 0.03	70.5 — 0.28
One Day Baseline	75.8 ± 14.0 — 0.32 ± 0.04	71.3 — 0.29
CNN	148.2 ± 30.1 — 0.51 ± 0.11	70.2 — 0.32
AutoML	/	61.6 — 0.34

MAE

TFT Forecasts

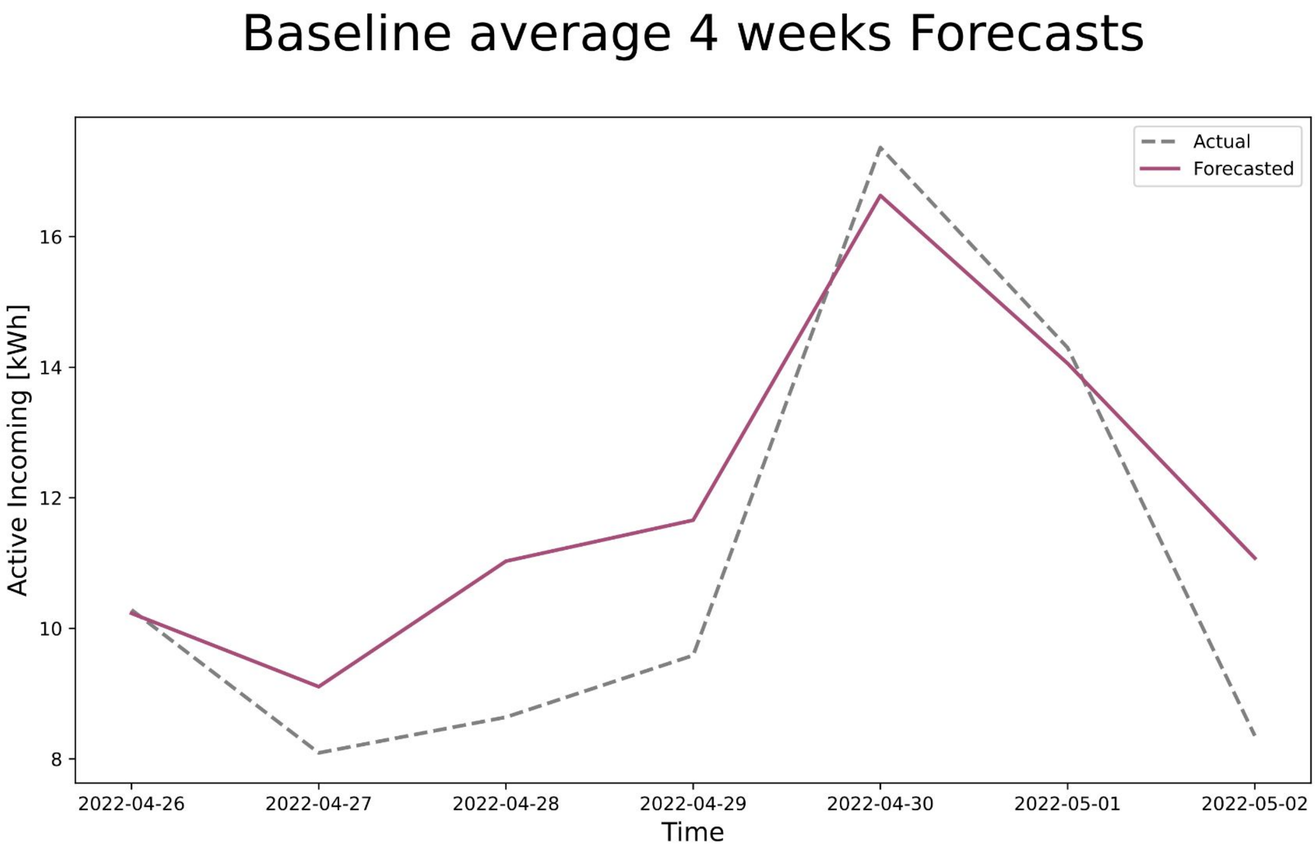


Daily Results

k = 12
Split size = 7 days

Model	Blocked k-fold cross-validation MAPE — MAE [kWh]	Test on the last split MAPE — MAE [kWh]
4 Week Baseline	16.2 ± 4.9 — 1.9 ± 0.6	14.4 — 1.3
One Week Baseline	19.6 ± 7.5 — 2.3 ± 0.9	14.5 — 1.6
CNN	28.5 ± 17.5 — 3.3 ± 1.9	13.5 — 1.8
SARIMA	16.4 ± 4.8 — 1.9 ± 0.5	17.1 — 1.8
TFT	16.7 ± 6.3 — 2.0 ± 0.6	14.6 — 1.8
Hist gradient boosting regressor	15.4 ± 3.7 — 1.9 ± 0.4	16.8 — 1.9
12 Week Baseline	15.6 ± 4.4 — 1.9 ± 0.5	20.0 — 2.0
XGBoost regressor	17.6 ± 4.9 — 2.1 ± 0.6	21.4 — 2.1
AutoML	/	19.9 — 2.2
GRU	18.7 ± 3.4 — 2.5 ± 0.5	17.9 — 2.4
LSTM	19.0 ± 3.5 — 2.5 ± 0.6	18.4 — 2.4
One Day Baseline	20.4 ± 3.9 — 2.7 ± 0.6	19.0 — 2.4
Prophet	17.4 ± 5.2 — 1.9 ± 0.4	25.5 — 2.5
SVR	18.1 ± 3.2 — 2.4 ± 0.6	23.4 — 2.7

MAE


Baseline average 4 weeks Forecasts
Actual
Forecasted
Active Incoming [kWh]
16
14
12
10
8
2022-04-26
2022-04-27
2022-04-28
2022-04-29
2022-04-30
2022-05-01
2022-05-02
Time

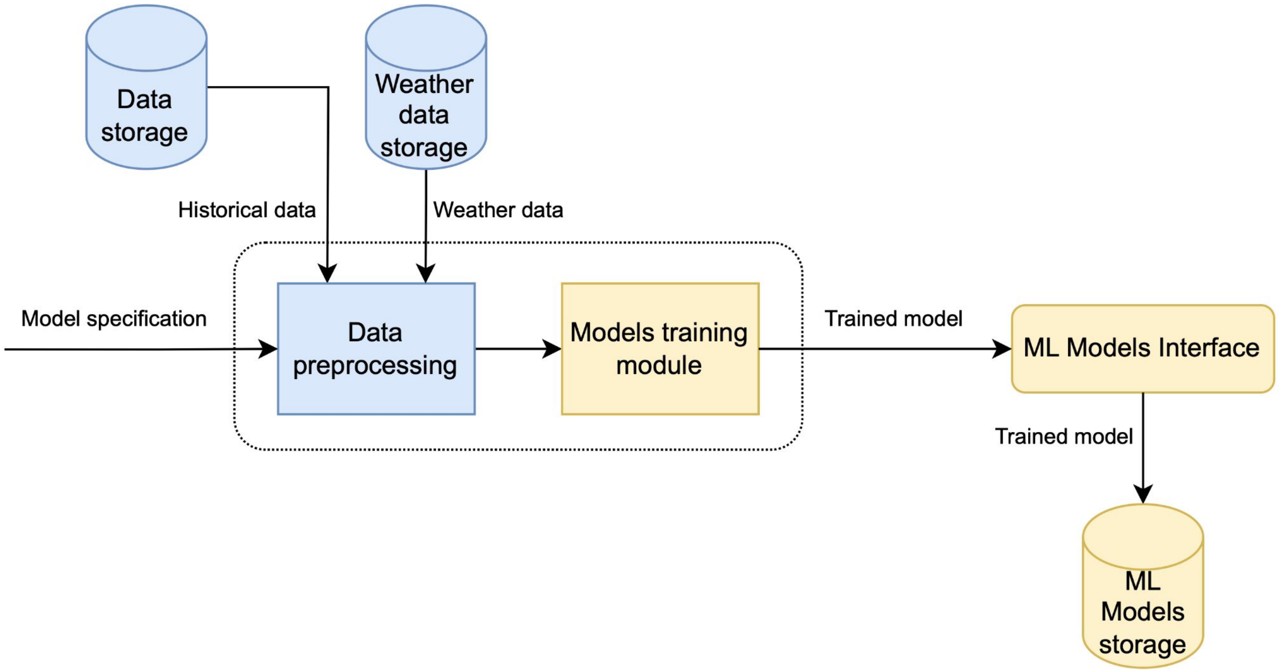
Conclusions

- Global energy crisis caused an **increased volatility** in the electricity market
- Designed a **novel forecasting system** and implemented a **prototype** of the system
- **Electricity production and customer demand can be accurately forecasted**
- **Consumption baseline** cannot be accurately forecasted considering the high variability of the data of the provided customers and the **limited amount of historical data**
- AutoML struggled to achieve competitive results and model combinations resulted in an improvement to DL models only in a limited number of cases

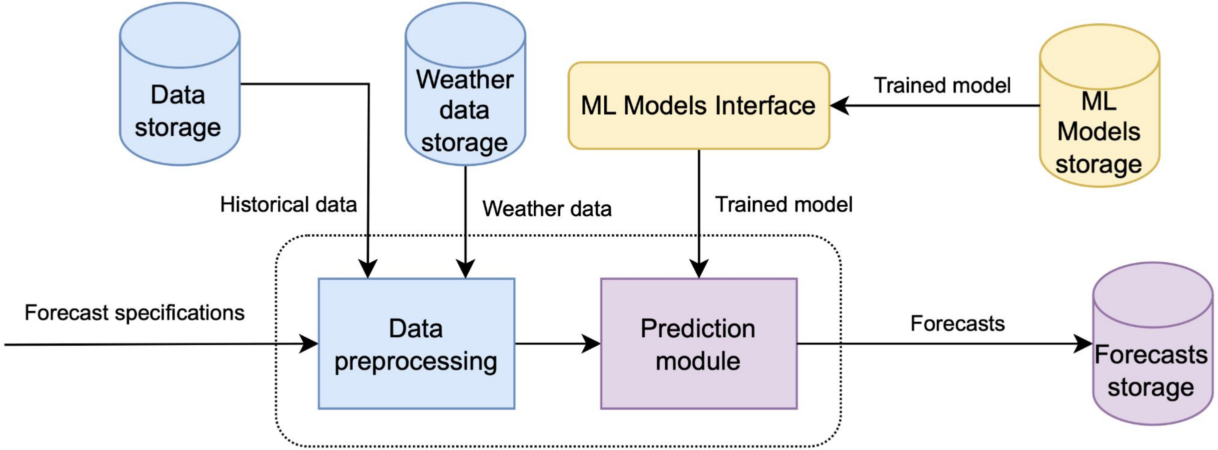
Future Directions

- **Ensemble learning** techniques
- Additional features (socio-demographic, household composition, …)
- Consumption disaggregation
- Full-fledged **SaaS solution implementation**

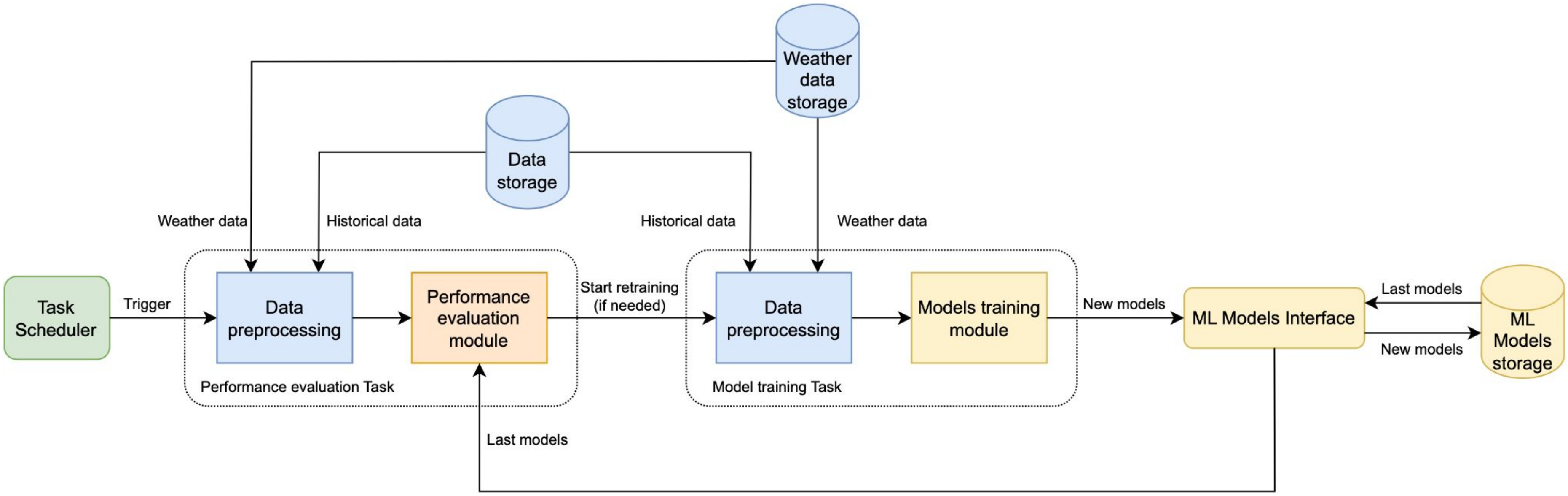
Model Training



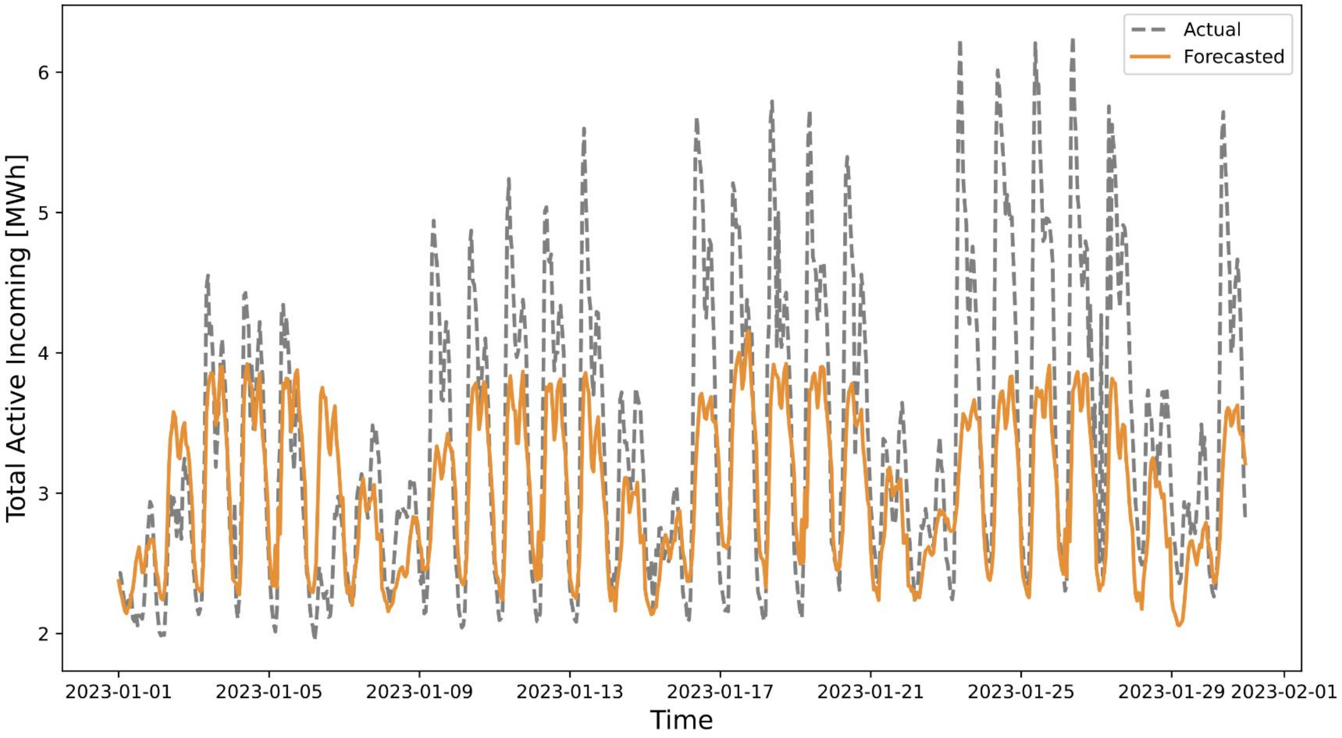
Forecasting



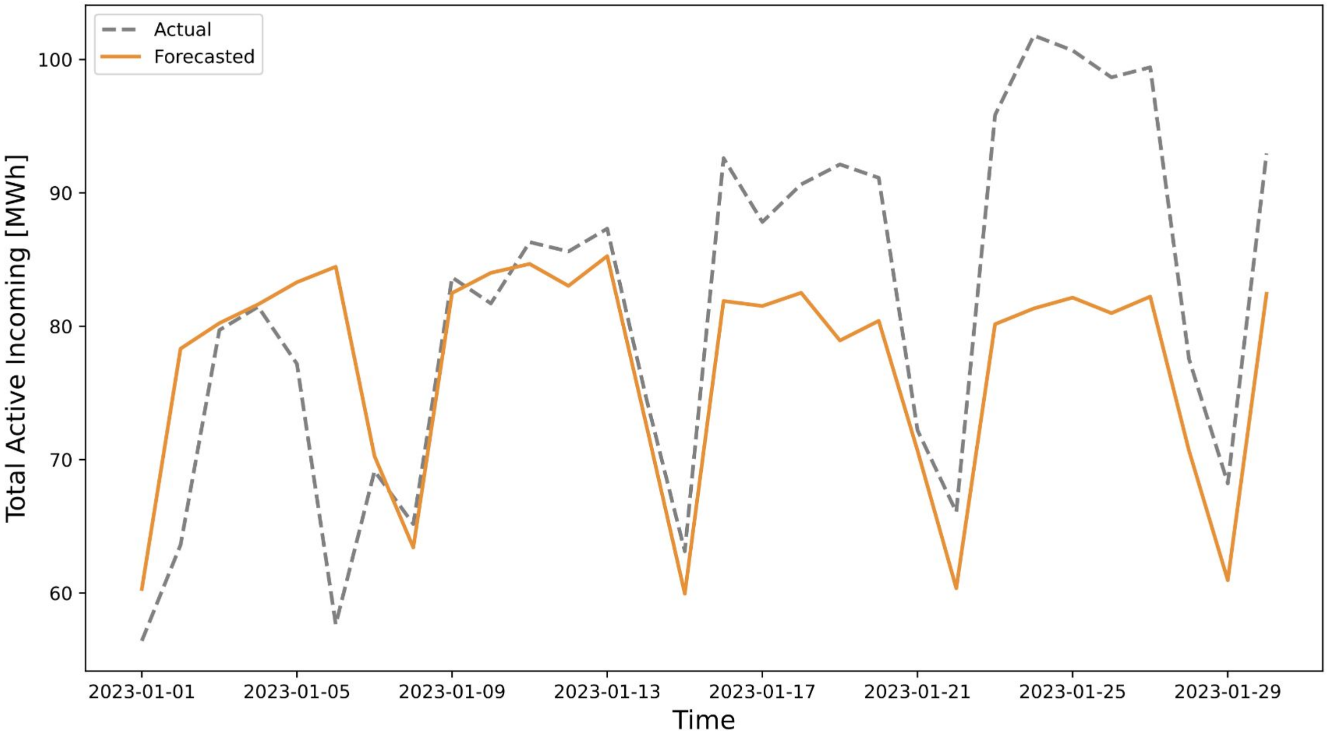
Task Scheduler



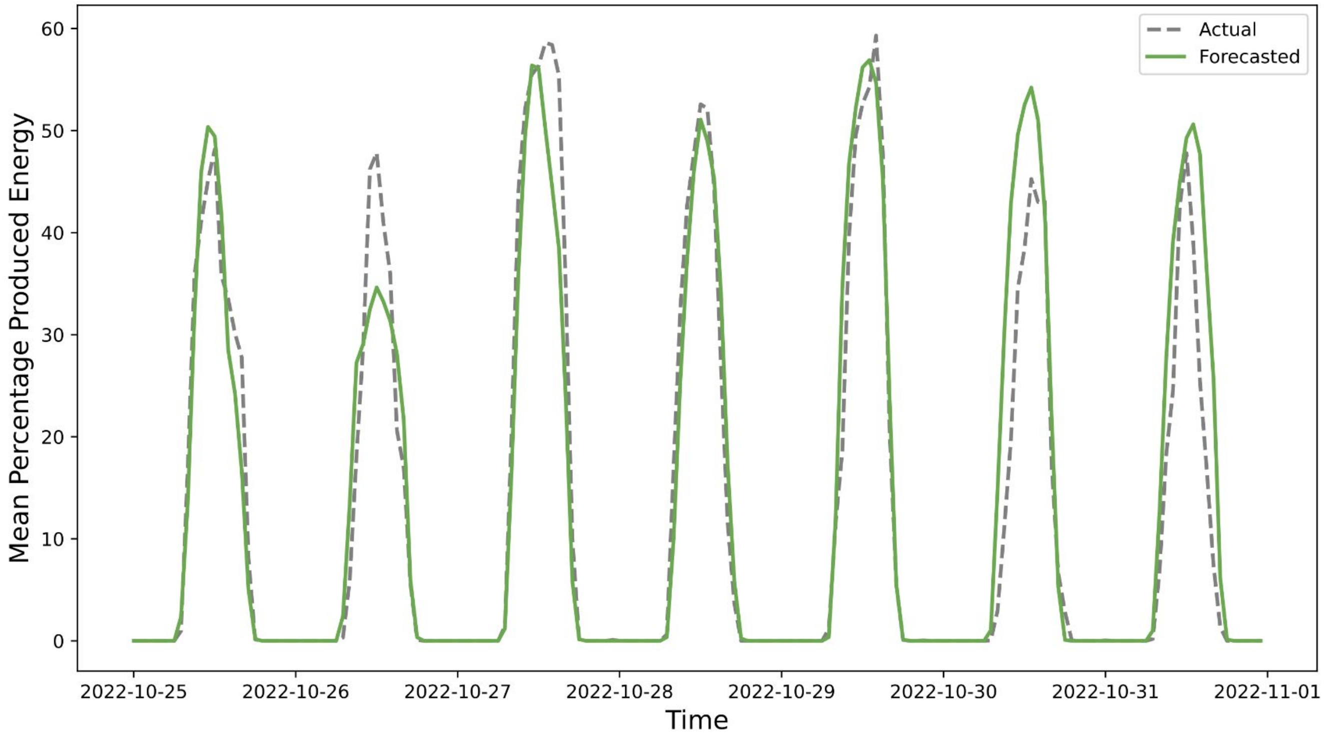
One week baseline combined with GRU Forecasts



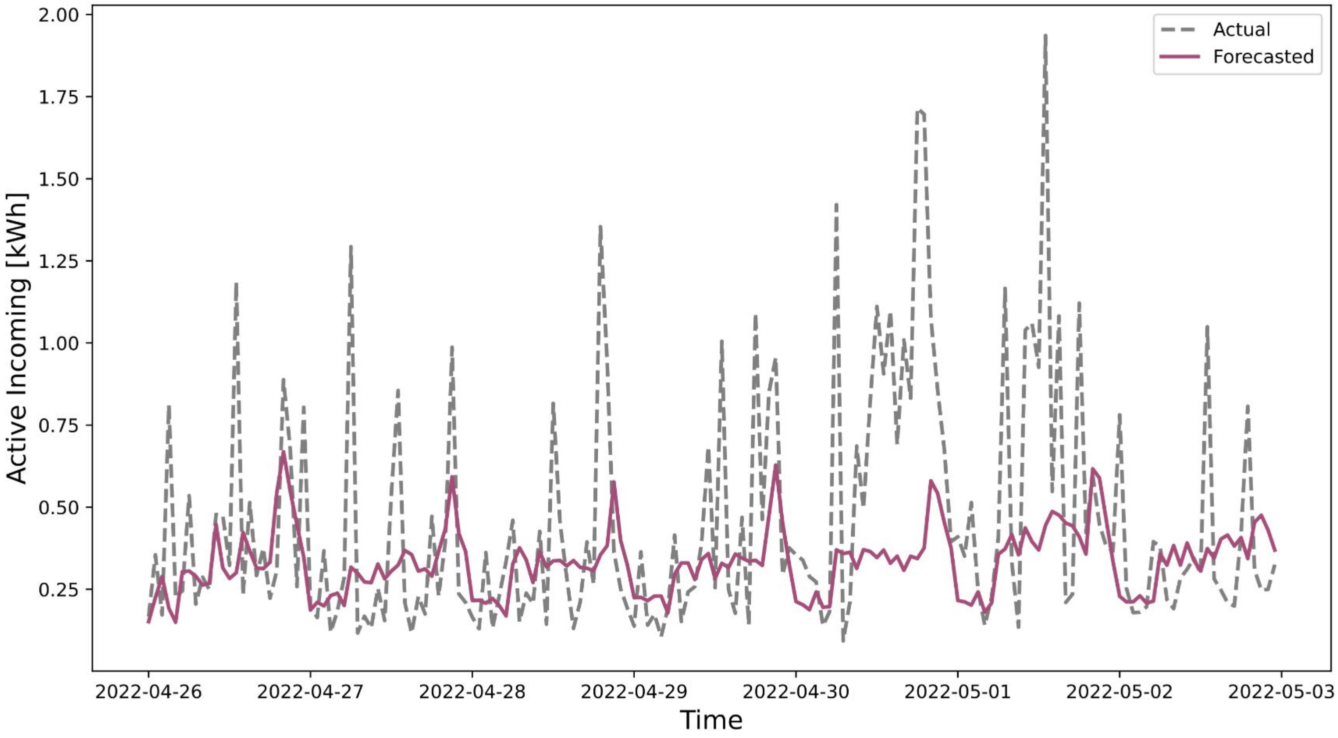
HistGradientBoostingRegressor Forecasts



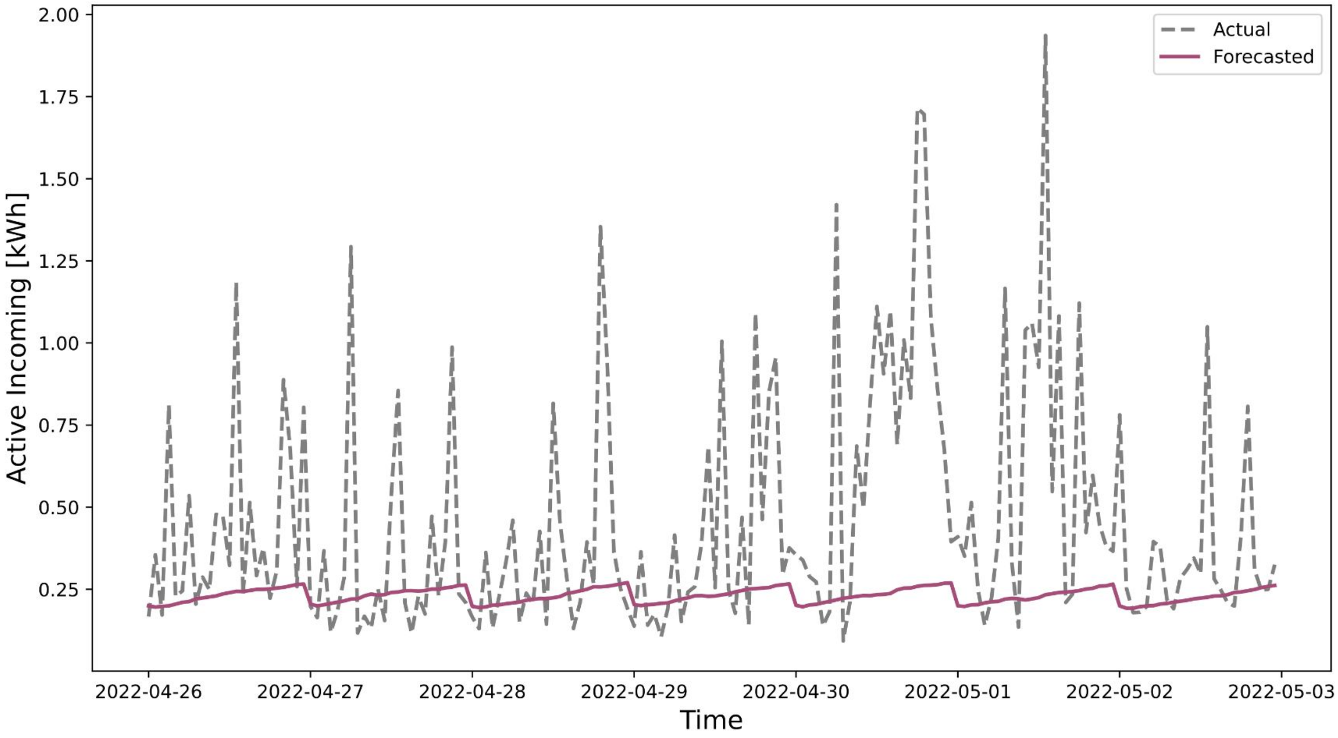
HistGradientBoostingRegressor Forecasts



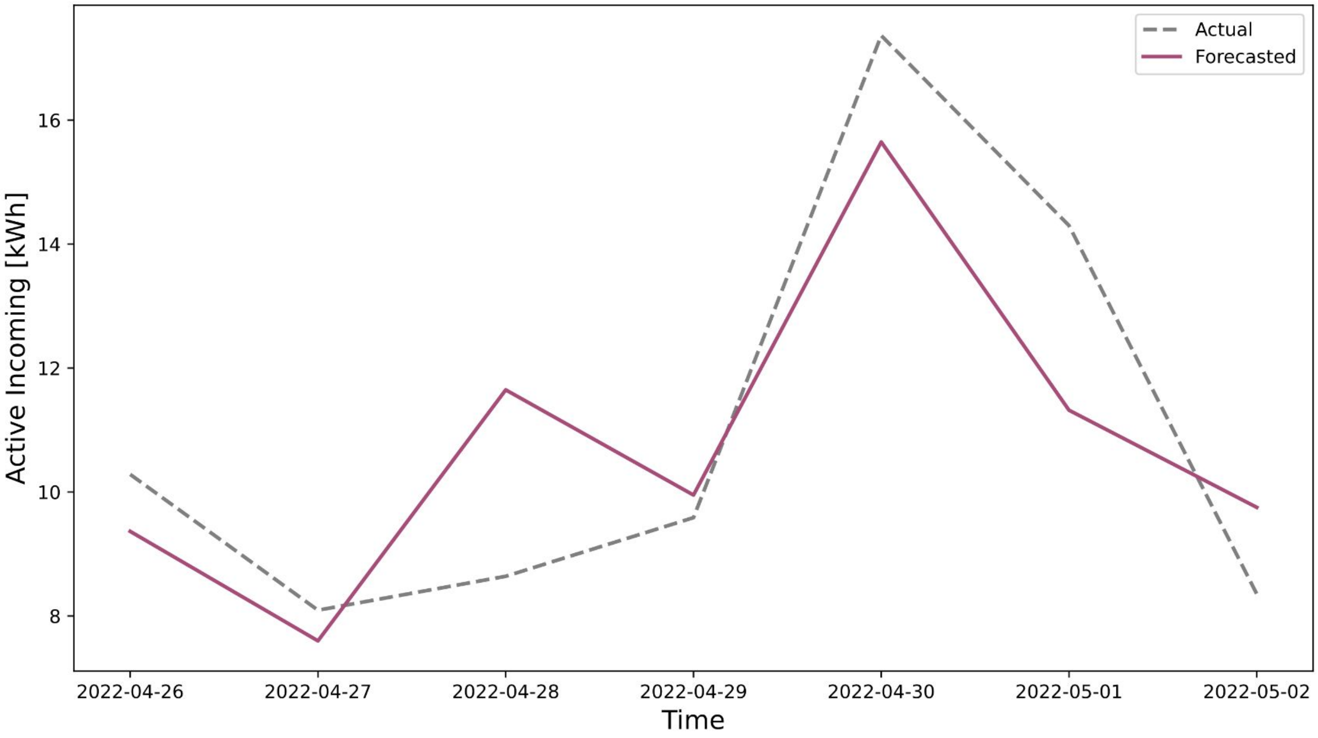
HistGradientBoostingRegressor Forecasts



GRU Forecasts



Baseline one week Forecasts



CNN Forecasts