AGH
UCZELNIA BADAWCZA
INICJATYWA DOSKONAŁOŚCI

Predicting energy consumption in households using time series

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
AGH University of Krakow

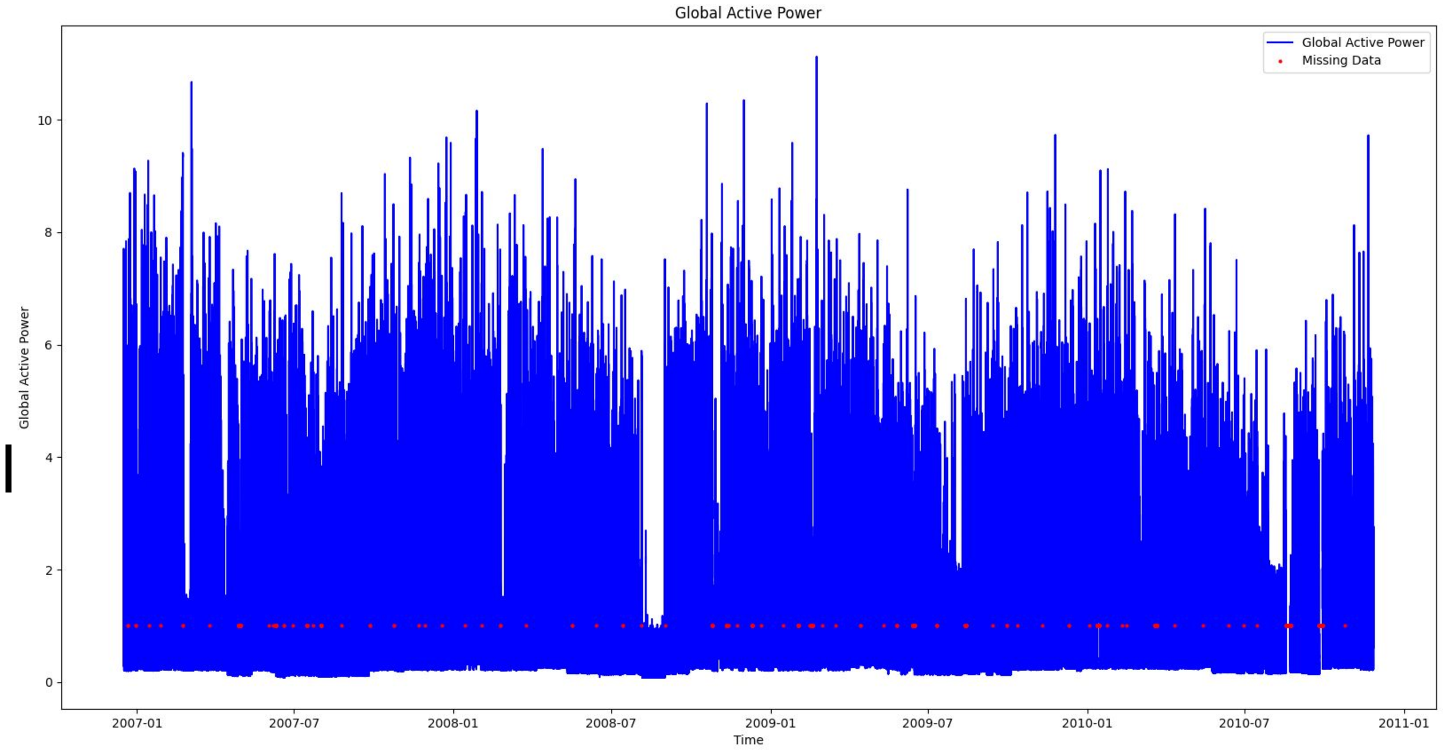
Kraków, 2024

Motivation

- **Energy Efficiency**: Accurate demand prediction helps optimize energy usage and reduce costs.
- **Regulatory Compliance**: New regulations require precise hourly electricity settlements.
- **Strategic Planning**: Informs decisions on energy storage and market balancing.

Dataset Description

- **Source**: "Household Electric Power Consumption" dataset from Kaggle.
- **Time Period**: December 2006 - November 2010 (47 months).
- **Data Points**: Approximately 2 million records.
- **Features**: Global active power, global reactive power, voltage, global intensity, sub-metering (kitchen, laundry, air-conditioning).



Data Preparation

Handling Missing Values: Substituted 'nan' values with column means.

Feature Engineering: Created new features and modified existing ones.

Visualization: Resampling data to observe trends, patterns, and seasonality

Missing values	Resampling	Data scaling
Forward fill missing values with last observation	Resampling from minutes to hourly basis	MinMaxScaler

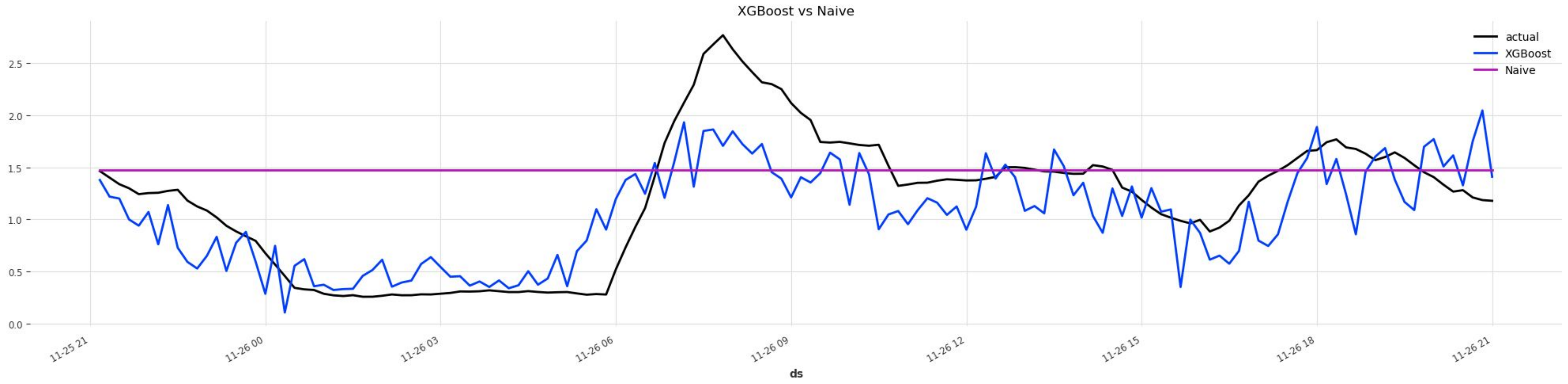
Model Selection

Models Used:

- **Naive Seasonal**: Simple benchmark model.
- **XGBModel**: Uses XGBoost algorithm.
- **LSTM**: Recurrent neural network designed to capture long-term dependencies.
- **Improved LSTM architecture:** LSTM model with dropout, Batch Normalization and multiple layers.

Naive Seasonal and XGBoost

24h prediction



LSTM

Architecture

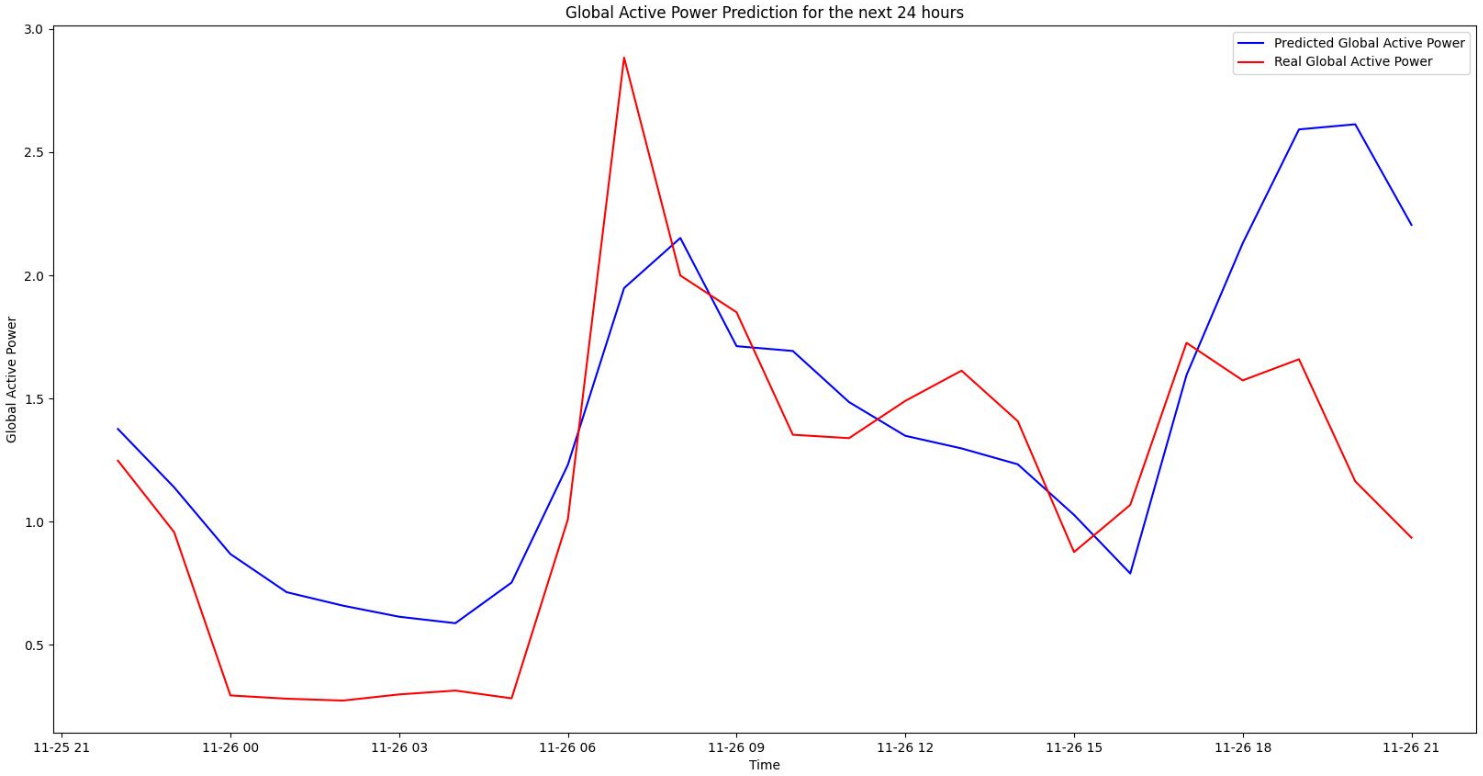
Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 24, 50)	11,600
lstm_1 (LSTM)	(None, 50)	20,200
dropout (Dropout)	(None, 50)	0
dense (Dense)	(None, 1)	51

Total params: 31,851 (124.42 KB)

Trainable params: 31,851 (124.42 KB)

Non-trainable params: 0 (0.00 B)

24h prediction



Tuned LSTM

Architecture

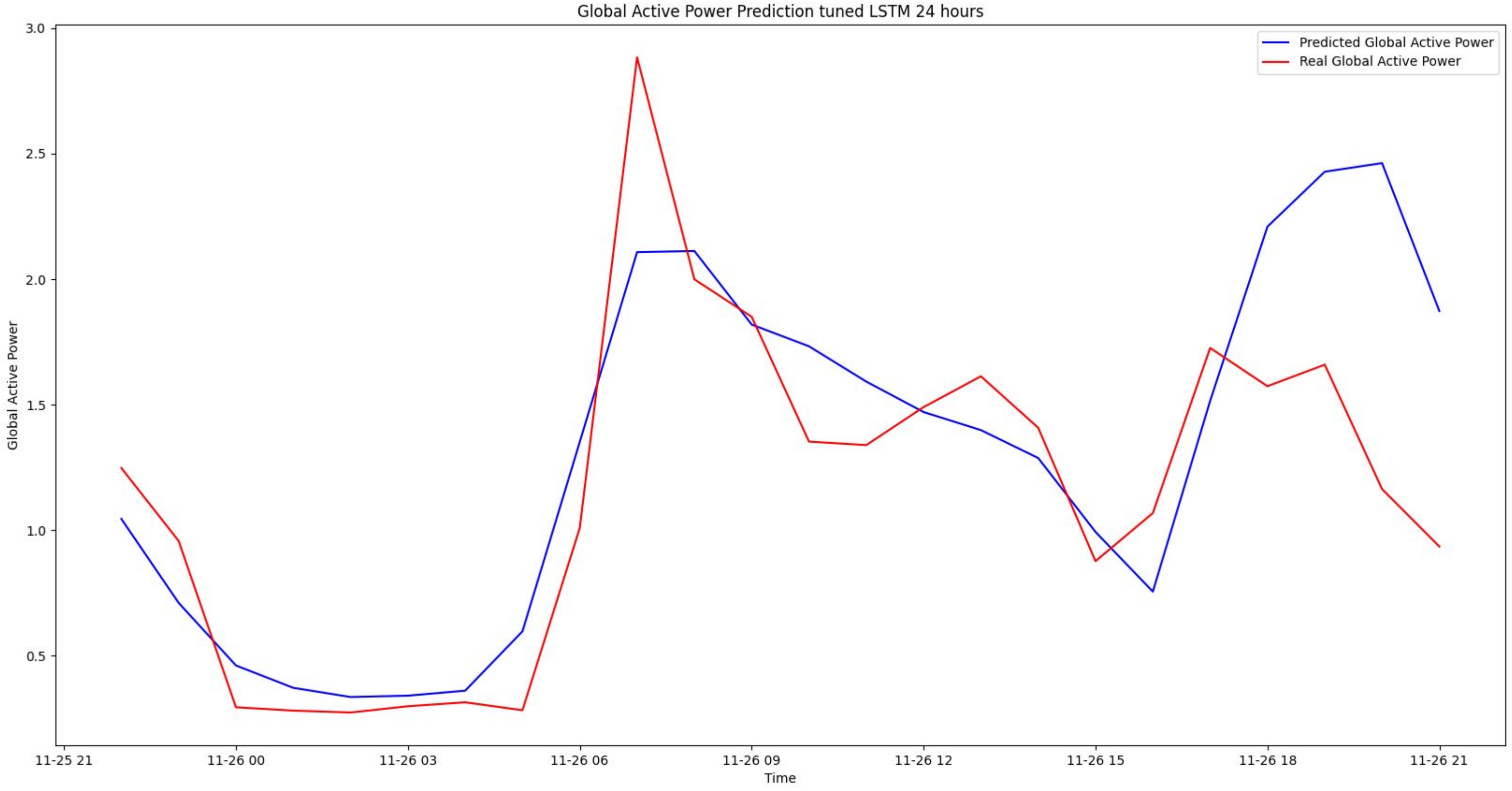
Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 24, 100)	43,200
lstm_1 (LSTM)	(None, 80)	57,920
dropout (Dropout)	(None, 80)	0
dense (Dense)	(None, 1)	81

Total params: 101,201 (395.32 KB)

Trainable params: 101,201 (395.32 KB)

Non-trainable params: 0 (0.00 B)

24h prediction



Tuned and more complex LSTM

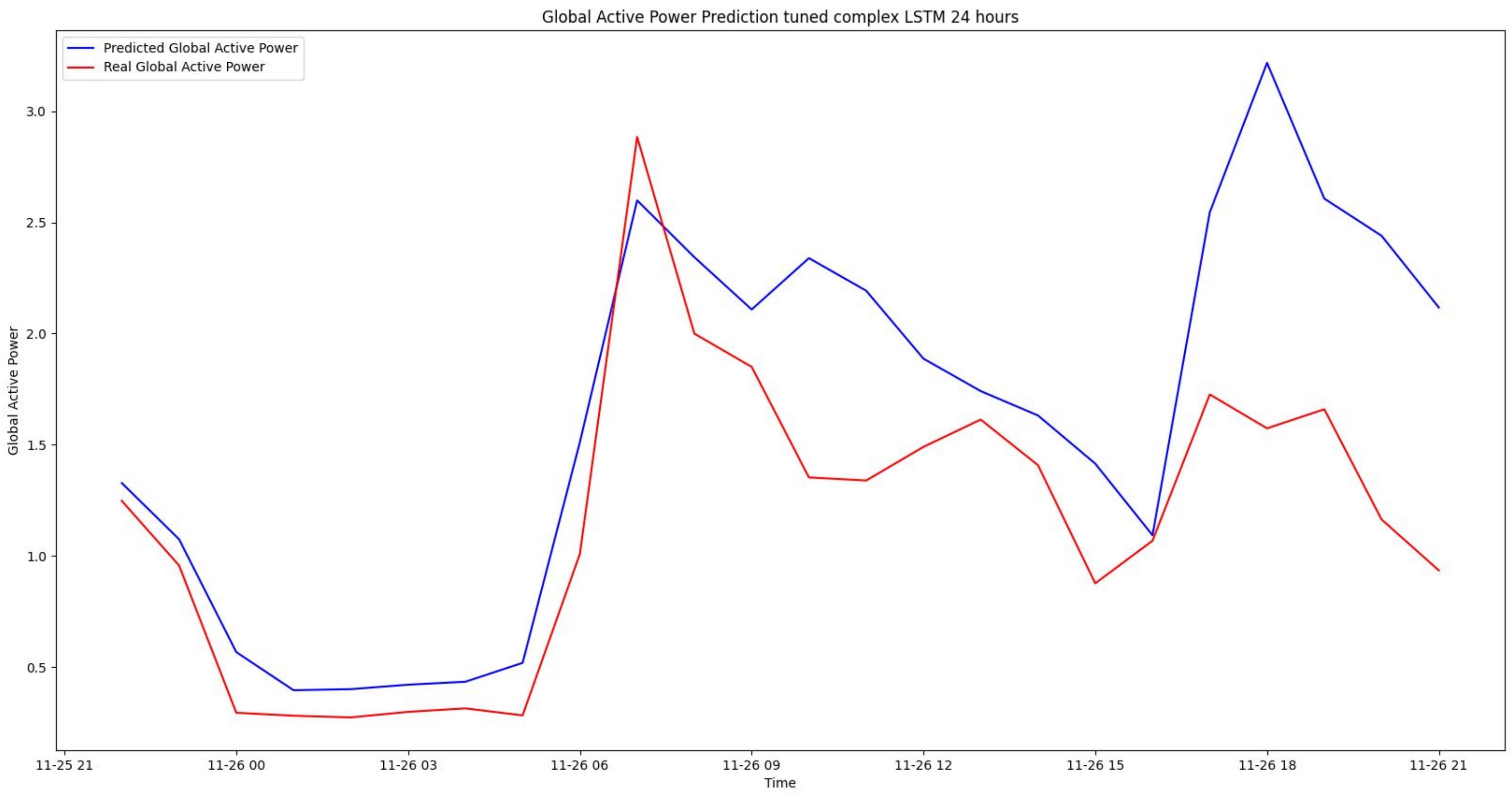
Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 24, 50)	11,600
batch_normalization (BatchNormalization)	(None, 24, 50)	200
dropout (Dropout)	(None, 24, 50)	0
lstm_1 (LSTM)	(None, 24, 80)	41,920
batch_normalization_1 (BatchNormalization)	(None, 24, 80)	320
dropout_1 (Dropout)	(None, 24, 80)	0
lstm_2 (LSTM)	(None, 50)	26,200
dropout_2 (Dropout)	(None, 50)	0
dense (Dense)	(None, 1)	51

Total params: 80,291 (313.64 KB)

Trainable params: 80,031 (312.62 KB)

Non-trainable params: 260 (1.02 KB)



Implementation Phase

Programming Language: Python, utilizing the Darts library for time series analysis. Tensorflow used for LSTM.

Model Training: Trained on historical data, hyperparameters tuned with Optuna.

Testing and Evaluation

- **Metrics**: MSE, MAE, MAPE RMSE
- **Model Optimization**: models hyperparameters were tuned with Optuna. Metric chosen for minimisation was MAE
- **Comparison**: various charts were used to represent the performance of the models - facilitating the comparison between models

Model name	MAPE	RMSE	MSE	MAE
LSTM paper work	-	0.62	-	-
Naive Seasonal	113.92%	0.69	0.48	0.52
XGBoost	**36.21%**	**0.41**	**0.17**	**0.32**
LSTM	44.10%	0.4788	0.2293	0.3370
tuned LSTM	42.07%	0.4755	0.2261	0.3272
tuned complex LSTM	43.54%	0.4825	0.2329	0.3311

Conclusion and Future Work

- **Current Findings**: LSTM effective for sequential data, but newer models could offer improvements.
- **Future Enhancements**: Exploring more complex LSTM architectures, consider more SOTA models and more training epochs, such as GRUs.
- **Impact**: Effective energy consumption prediction aids in efficient energy use and cost reduction - one can use the forecast data to predict whether to buy energy to store in an energy bank, sell it right away or use for current needs. It can also facilitate the decision-making process of electricity vendor, deciding how much electricity will be needed

Questions?