Short-Term Electricity Demand Forecasting

Pinky Sherwani, Gagan Viji Sha, Veera Raghavulu Bathula
MSc. Data and Knowledge Engineering
Data Science with Python (Seminar)
Faculty of Computer Science
Otto Von Guericke University Magdeburg

16.01.2026

Prof. Dr.-Ing. Christian Beyer
Supervisor, Faculty of Computer Science (FIN),
Otto von Guericke University Magdeburg

AGENDA

- Introduction
- Objective & Research Questions
- Data Description & Preprocessing
- Exploratory Data Analysis
- Modeling Approaches Used
- Model Performance Results
- Explainability with SHAP
- Website Overview
- Challenges Faced
- Future Work
- Conclusion
- References

Introduction

Problem Statement

- Accurate short-term electricity demand forecasting critical for power system operations,
- With generation scheduling, grid stability, and cost optimization.
- Electricity demand exhibits:
  - Strong daily and seasonal patterns
  - Non-linear behavior driven by weather and human activity
  - Large errors during peak demand hours
- Traditional statistical models struggle with non-linearities, interpretability and stability

Motivation

- Operational risk is highest during peak load periods,
- Studies optimize overall accuracy, but peak-hour performance is still underexplored
- Practical need to:
  - Compare classical and deep learning approaches
  - Improve peak-load forecasting
  - Combining statistical structure with neural flexibility

Ref: Generated by AI

Objectives & Research Questions

Objective: To systematically compare statistical and deep learning models for short-term electricity demand forecasting

RQ1 — Model Comparison

How do classical time-series models (SARIMA) and deep learning models (LSTM) differ for short-term electricity demand forecasting?

- Forecast accuracy
- Interpretability
- Computational and operational characteristics

RQ2 — Peak-Load Performance

Can hybrid SARIMA+LSTM or ensemble models:

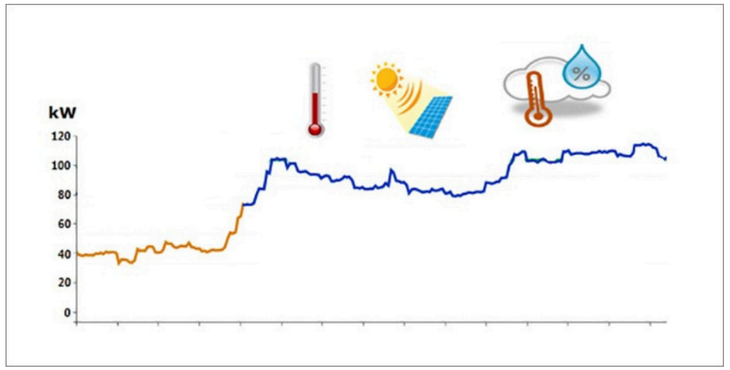
- Reduce forecast error variance
- Improve stability

  during peak-load periods, where forecasting failures are most costly?

RQ3 — Explainability & Trust

Do explainability methods (e.g., SHAP):

- Provide meaningful insights into LSTM and hybrid model behavior?
- Enhance trust and usability of forecasts for grid operators?



Ref: Load Forecasting

Data Description & Preprocessing

Raw Dataset

- Electricity Dataset Source : Open Power System Data
- Weather Dataset : Renewables.ninja
- Original Dimensions: 50,401 rows × 299 columns
- Time Range: 2014-12-31 to 2020-09-30 (hourly data)
- Coverage: ~35 European countries

Processsed Dataset

- Final Dimensions: 32,665 rows × 95 columns
- Time Range: 2015-01-08 to 2018-09-30
- Resolution: Continuous hourly intervals
- Data Quality: 100% complete after preprocessing

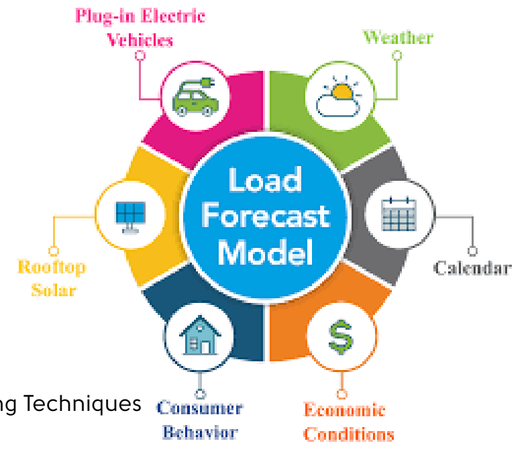


Figure: Load Forecasting Techniques
Ref: [6]

Key Variables

Energy Data

- Target: Electricity load (MW) - actual consumption
- Renewable generation: Solar and wind (actual, offshore, onshore)
- Capacity data: Installed solar and wind capacity
- Control regions: Four German TSO load data

Weather Data (Population-weighted for Germany)

- Temperature (°C) and humidity (%)
- Wind speed (m/s) and solar radiation (W/m²)
- Cloud cover (%) and precipitation (mm)
- Additional: Air density, snow mass, snowfall

Engineered Features

- Calendar: Hour, day, month, season, holidays, weekends
- Temporal: Lag features (1-168 hours), rolling statistics
- Special: Heating/cooling degree days, generation proxies

Data Preprocessing Pipeline

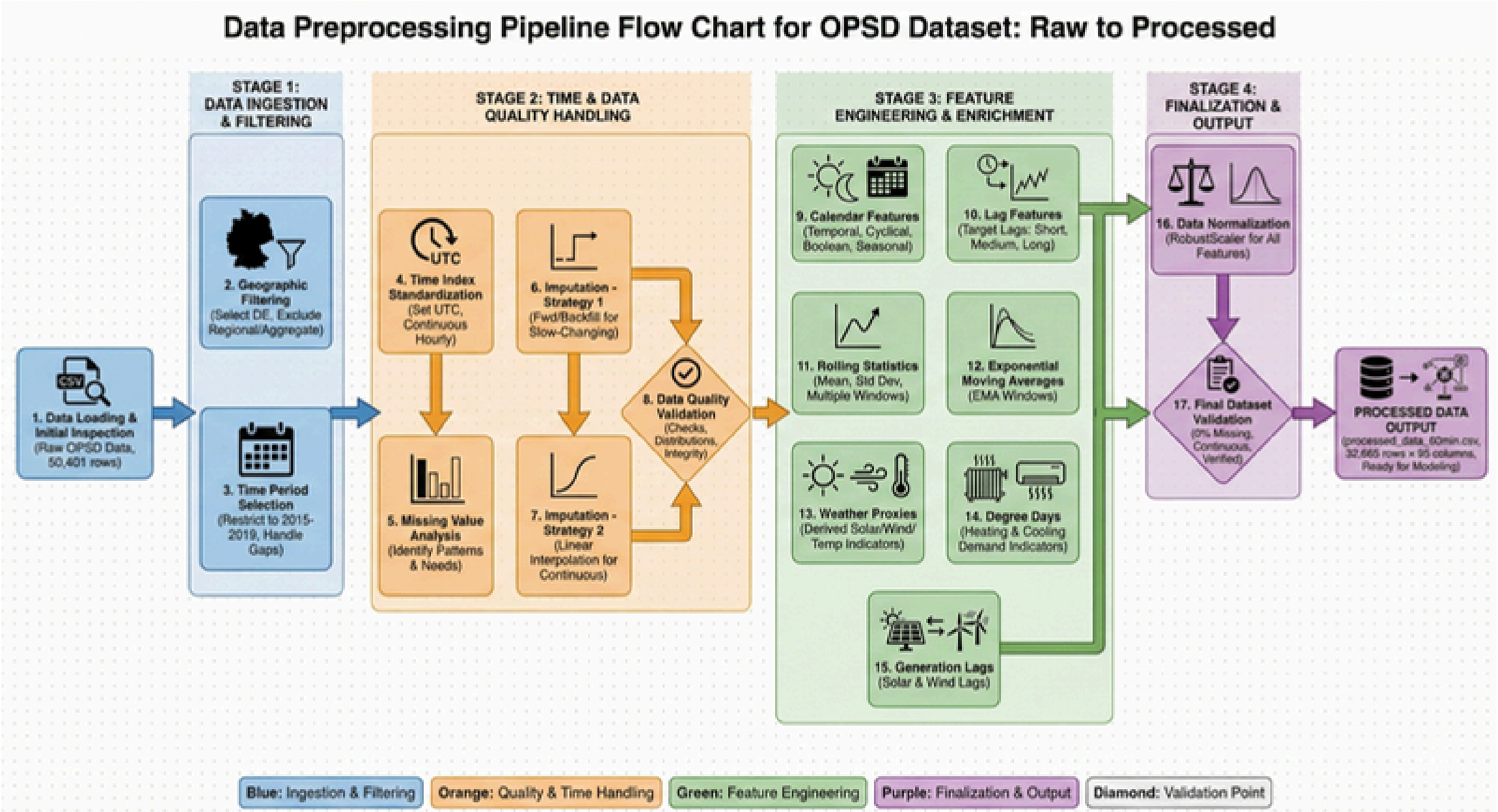


Figure : Preprocessing Steps (Ref: Generated by AI)

Exploratory Data Analysis

Dataset Summary

- Resolution: Hourly (60 min)
- Period: Jan 2015 – Sep 2018 (32,665 obs)
- Features: 95 (load, generation, weather, calendar)
- Data Quality: No missing values

Target Variable (Load)

- Mean: 55,705 MW | Range: 31,307–77,549 MW
- Distribution: Nearly normal (skew -0.04, kurtosis -1.15)
- Outliers: None (IQR method

Seasonality Patterns

- Daily: Peak 10 AM, Min 1 AM (strong diurnal cycle)
- Weekly: Peak Wed, Min Sun (week pattern)
- Monthly: Peak Feb, Min May (annual pattern)
- Weekday vs Weekend: +18.8% higher demand

Weather Impact

- Solar radiation TOA: 0.40 (strongest predictor)
- Surface solar: 0.35 | Wind: 0.28 | Temp: -0.02

Peak Load Analysis

- Threshold: 90th percentile = 68,857 MW (10% of data)
- Peak Timing: 9 AM, Wed, Jan
- Weather: Lower temp (-0.48), higher wind (+0.56)

Feature Importance

- Control region loads: 0.97–0.99
- Lagged loads: 1h (0.966), 168h (0.916)
- Rolling stats: 3h mean (0.9685), 6h EMA (0.9061)
- Calendar: Hour cosine (0.65)

Modeling Approaches

SARIMA

Model Overview

- Model: SARIMA (Seasonal ARIMA)
- Approach: Univariate time series (load only)
- Dataset: Hourly load data (2015–2018)
- Library: statsmodels ARIMA

Data Preparation

- Split: Time-based (no leakage)
  - Train: 2015–2017
  - Test: 2018
- Target: Normalized hourly electricity load

Model Architecture

- SARIMA(1,1,1) × (1,1,1,24)
- Training data: Last 1 year only (8,640 samples)
- Seasonality: Daily (s = 24)

LSTM

Model Overview

- Model: Multivariate LSTM Neural Network
- Dataset: Hourly data (2015–2018)
- Features: 94 engineered features
- Framework: TensorFlow / Keras

Data Preparation

- Time-based Split (no leakage):
  - Train: 2015–2017
  - Validation: 20% of training
  - Test: 2018
- Target: Normalized electricity load (MW)

Model Architecture

- layers: 64 → 32 units (stacked)
- Dense layer: 16 units (ReLU)
- Output: 1 neuron (linear)
- Regularization: Dropout = 0.2
- Parameters: ~30K trainable

Training Strategy

- Optimizer: Adam (lr = 0.001)
- Loss: MSE | Metric: MAE
- Batch size: 64
- Early stopping: Validation-based

Modeling Approaches

HYBRID

Model Overview

- Model: Hybrid SARIMA + LSTM
- Approach: Error-aware modeling for extreme demand hours
  - SARIMA captures linear trend + seasonality
  - LSTM learns nonlinear patterns left in SARIMA residuals
- Two evaluation regimes: All hours and Peak hours only

Data Preparation

- Scaling
  - Residuals normalized using MinMaxScaler
- Peak hours defined as:
  - Load ≥ 90th percentile of historical hourly demand

Model Architecture:

- Target: SARIMA residuals $Residual_t = y_t - y'_t SARIMA$
- Input features: Lagged residual sequences (lookback window)
- Output: Residual correction term
- Final Prediction

$$y'_t Hybrid = y'_t SARIMA + r'_t LSTM$$

ENSEMBLE

Model Overview

- Model: Combines forecasts from multiple models into one prediction to improve accuracy and peak-load stability.
- Dataset: Hourly data (predictions from SARIMA,LSTM,HYBRID)

Data Preparation

- Standardized data using StandardScaler for Ridge-based ensembles and SHAP explainability
- Peak hours defined as load > 90th percentile of evaluation data

Methods:

- **Mean(Simple Average):** Mean of all base models
- **Median**: Median of all base models
- **Stacked Ridge**: Ridge regression meta-learner
- **Weighted Ridge:** Performance-weighted combination
- **NNLS (Non-Negative Least Squares)**: Constrained optimization
- **Residual Corrector**: LSTM-based error correction
- **Switch Models:** Conditional model selection based on peak detection

Overall Model Performance Results

Performance Metrics	SARIMA	LSTM	HYBRID	ENSEMBLE (final_switch_residual)
MAE	11,526 MW	818.79 MW	16475 MW	**577.09 MW**
RMSE	13,362 MW	1,087.63 MW	18460 MW	**740.69 MW**
MAPE	19.56%	1.46%	28.70%	**1.10%**
ME (Bias)	–11,063 MW	–93.88 MW	-16351 MW	**-236.22MW**
MASE	6.04	0.43	8.63	**0.121**
R^2	–0.847	0.9877	-2.53	**0.99**
Directional Accuracy	82.89%	91.13%	80.97%	**92.59%**

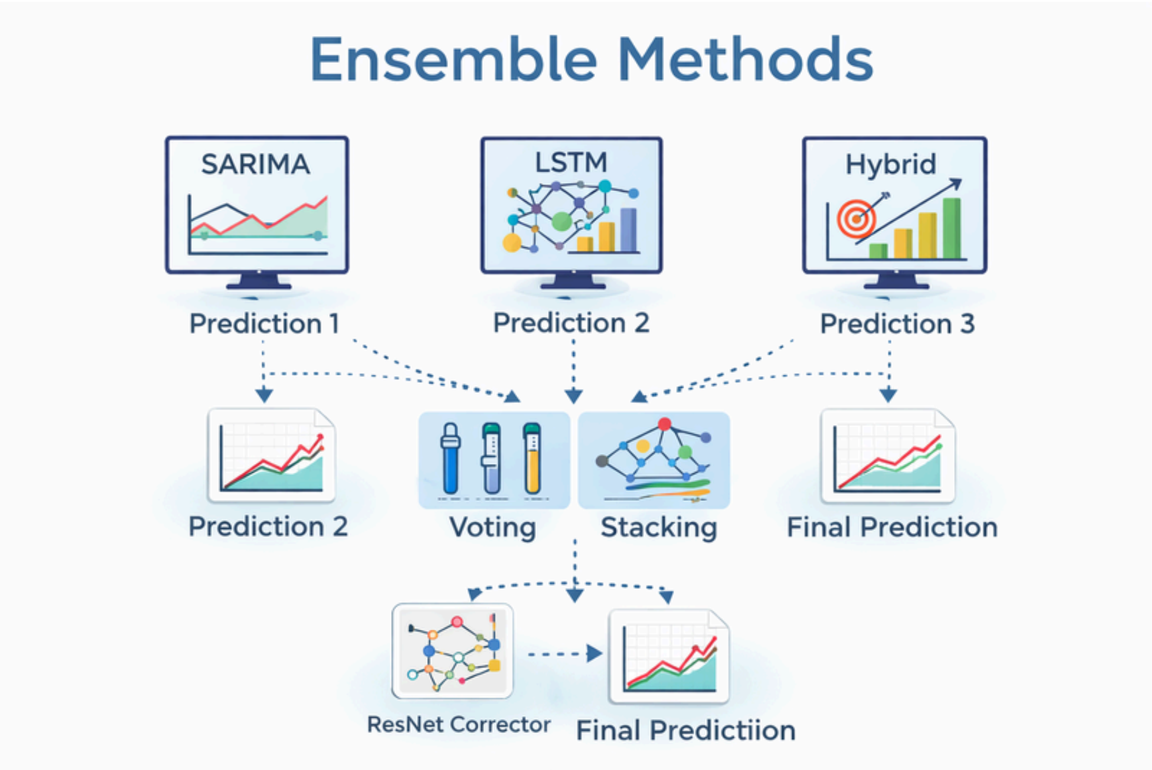
Ensemble Methods Performance Results

Best overall and peak model: final_switch_residual, lowest error overall and during peaks by learning and correcting residual forecast errors.

Strong alternatives: ens_residual_lstm and LSTM/NNLS consistently good, but slightly less accurate than residual correction.

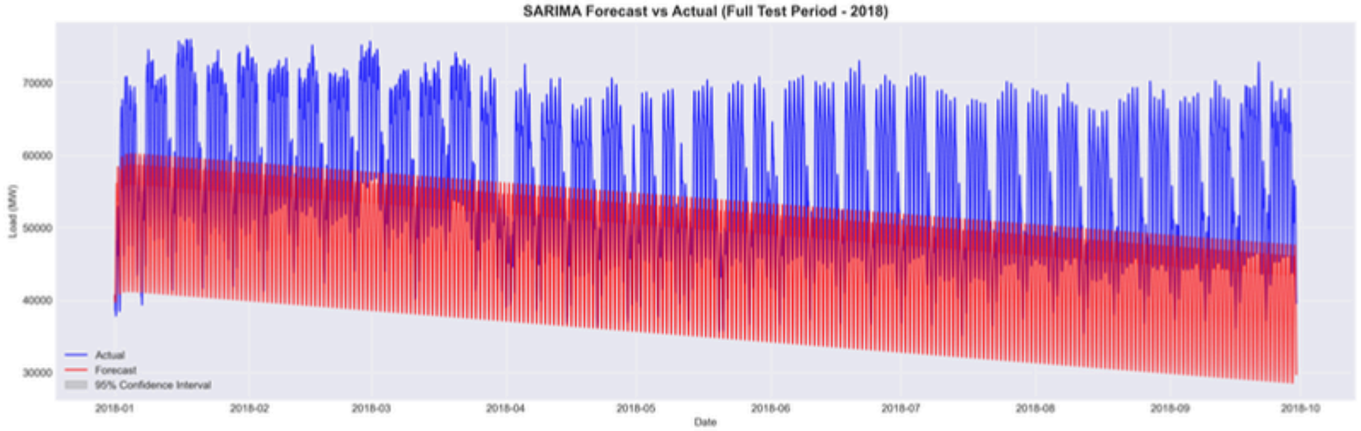
ens_weighted_ridge: strong peak performance by giving higher weight to better base predictors.

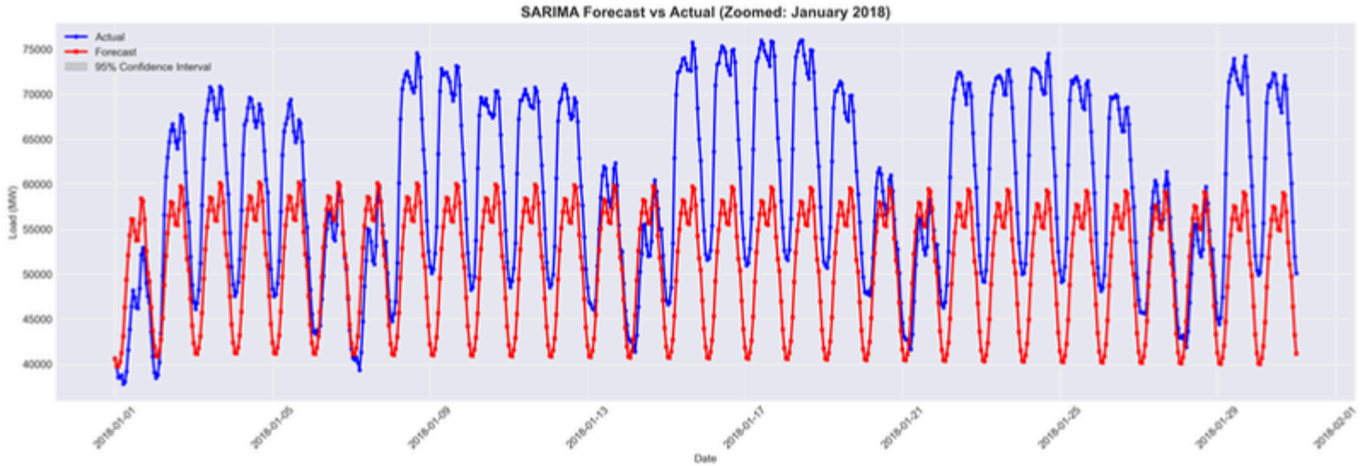
mean/median: mean and median ensembles, weak because poor base forecasts distort the combination and they cannot adapt to spikes.

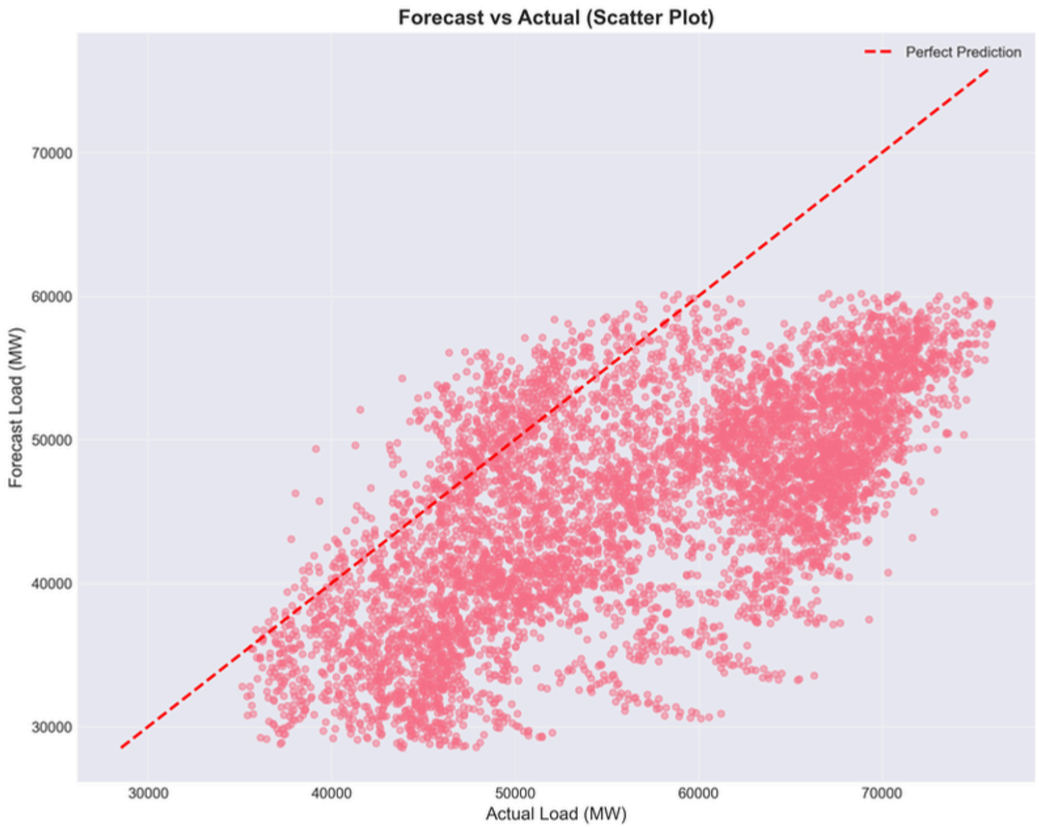


Ref: Generated by AI

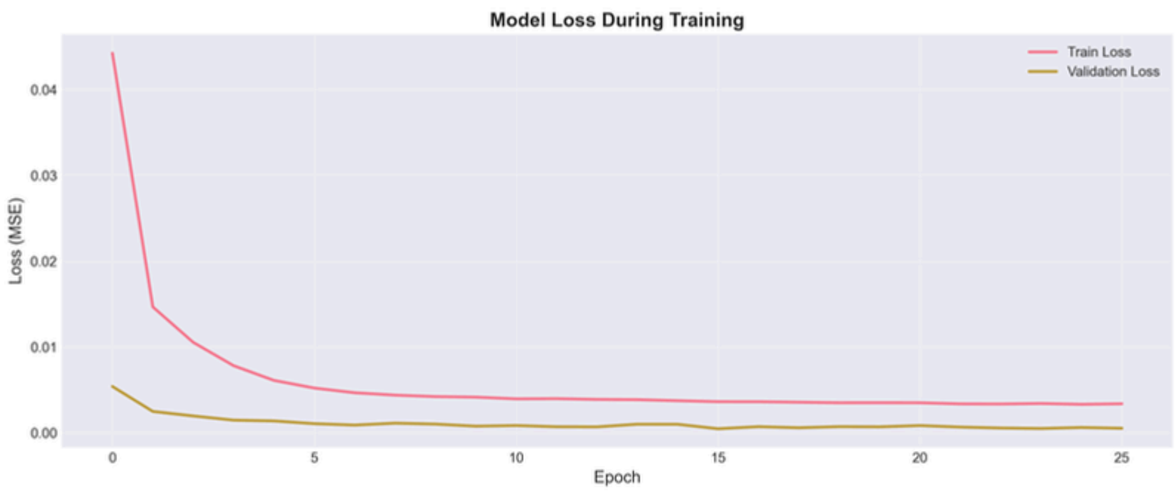
Graphs Analysis & Results: SARIMA

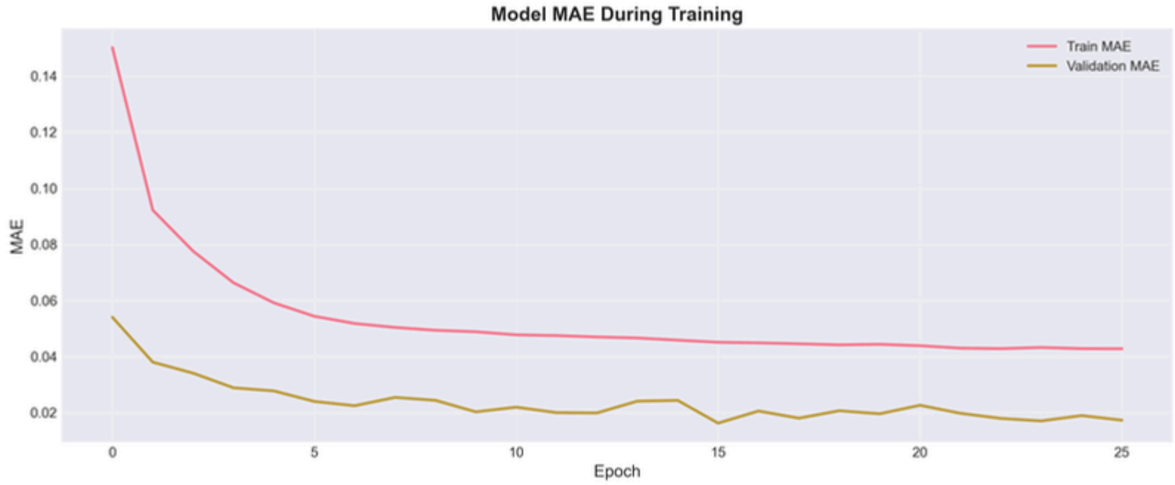


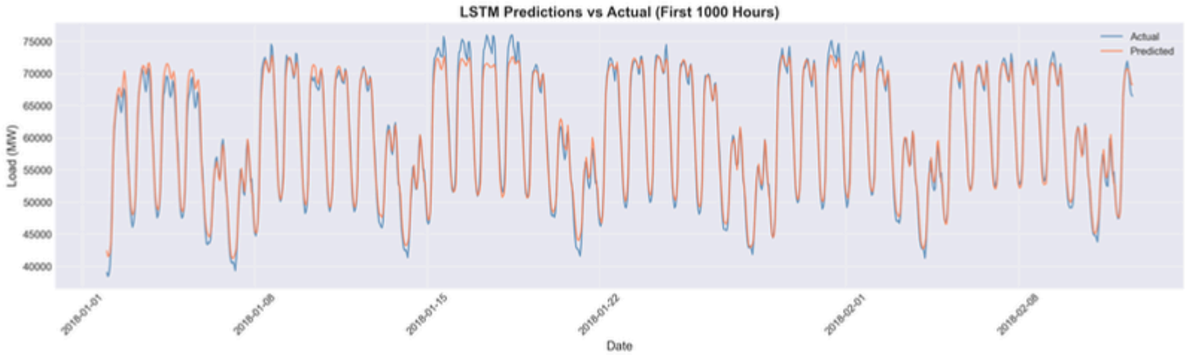


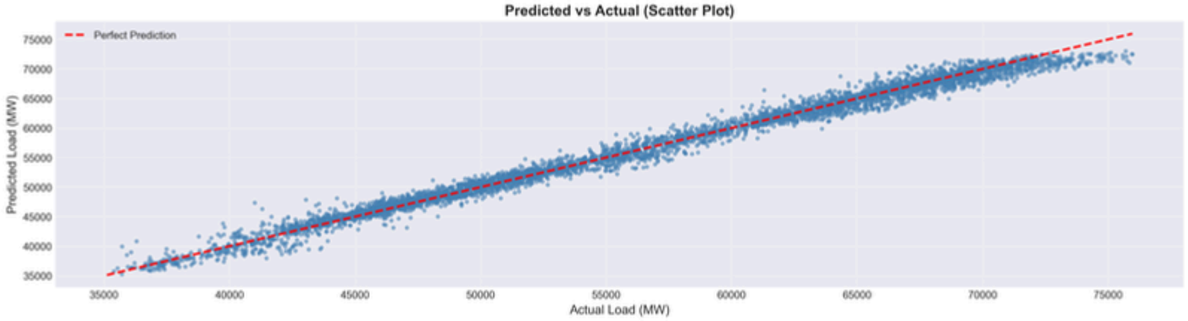


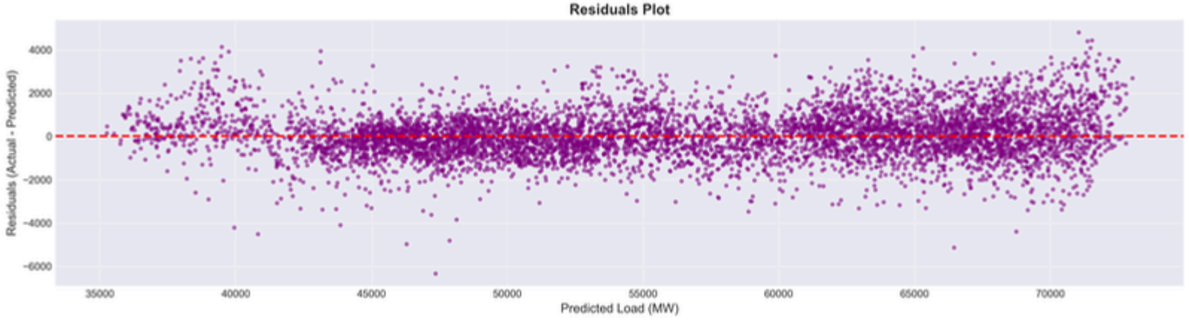
Graphs Analysis & Results: LSTM



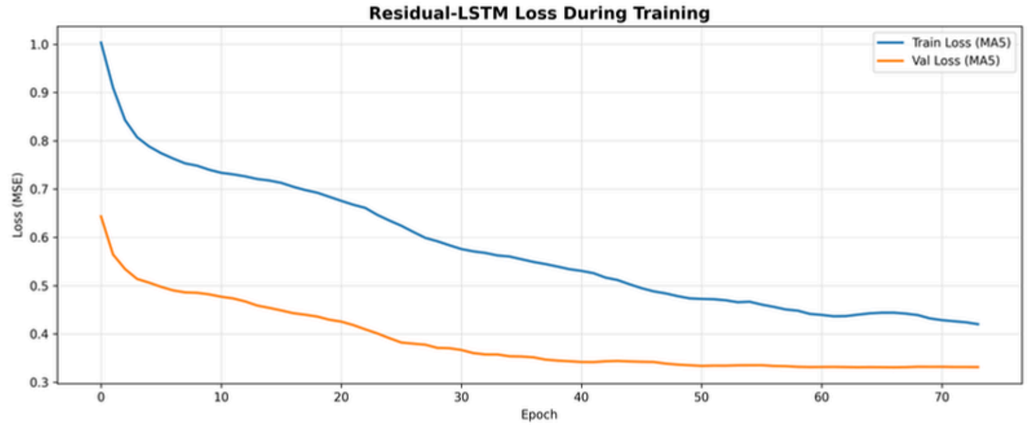


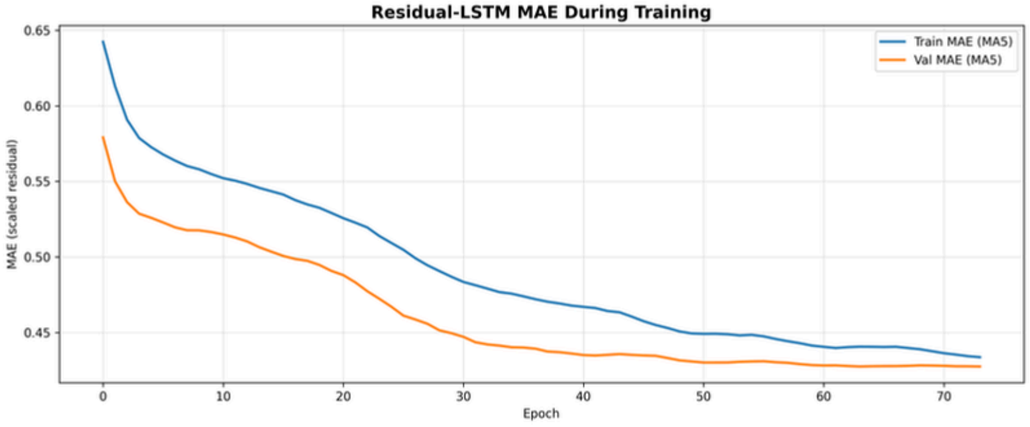


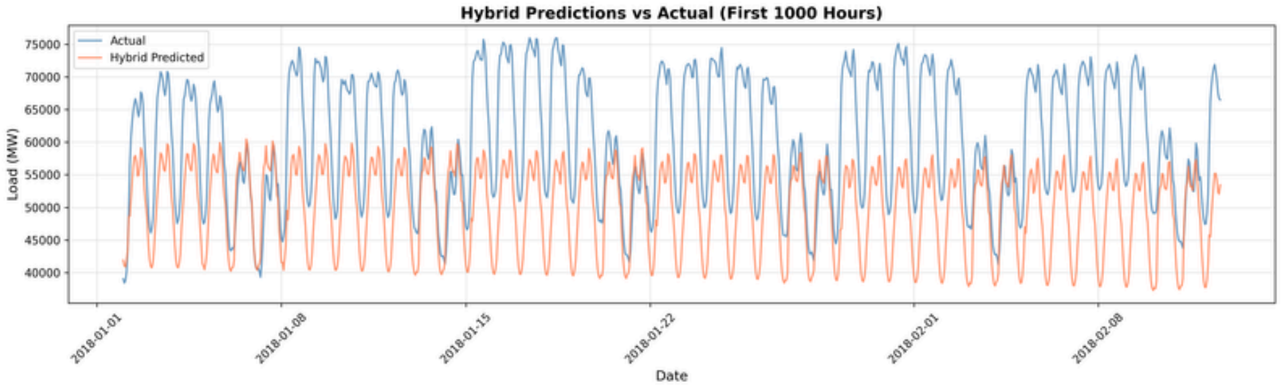


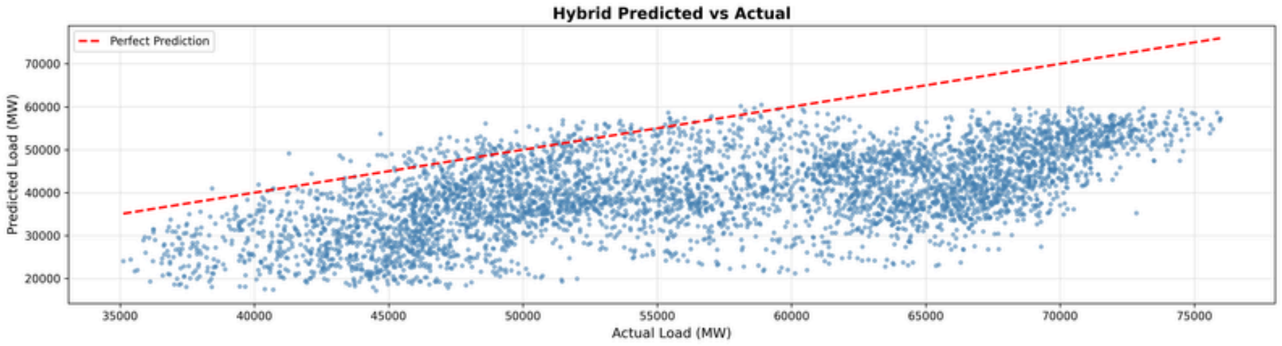


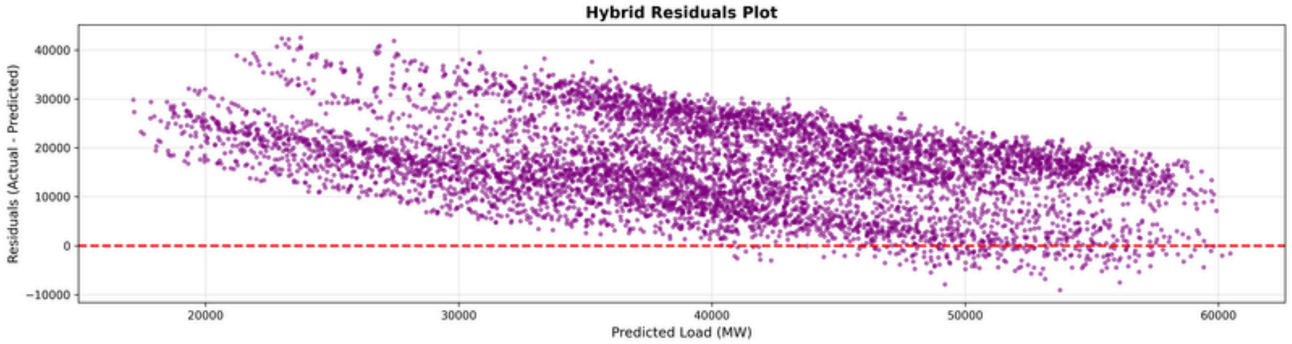
Graphs Analysis & Results: Hybrid



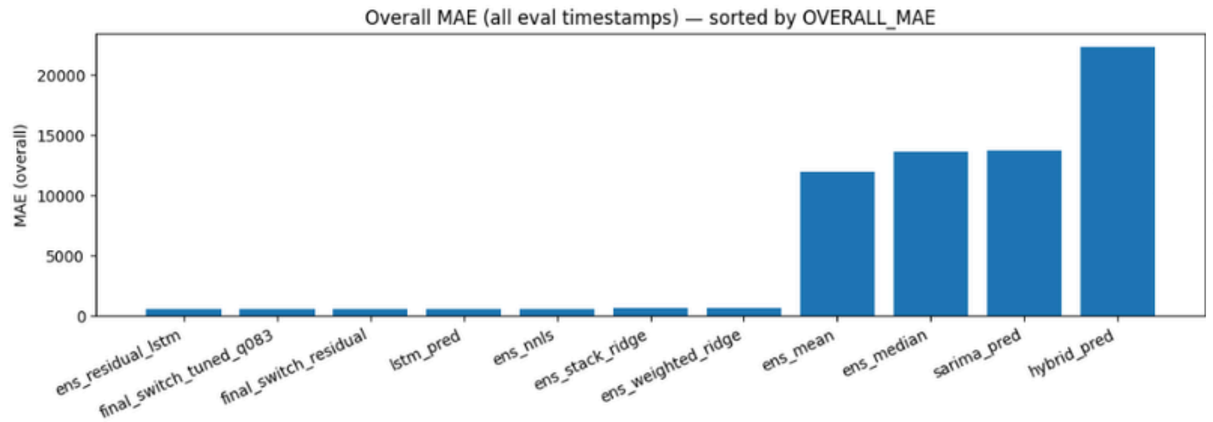


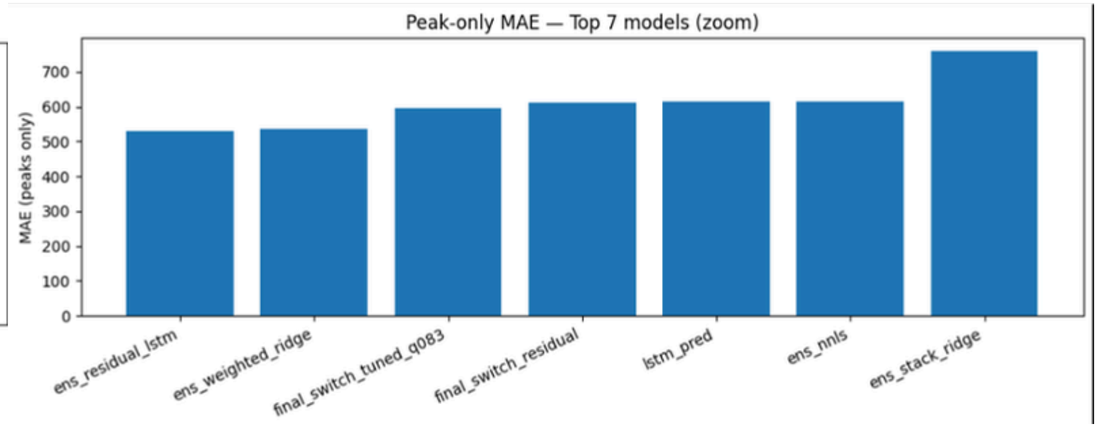


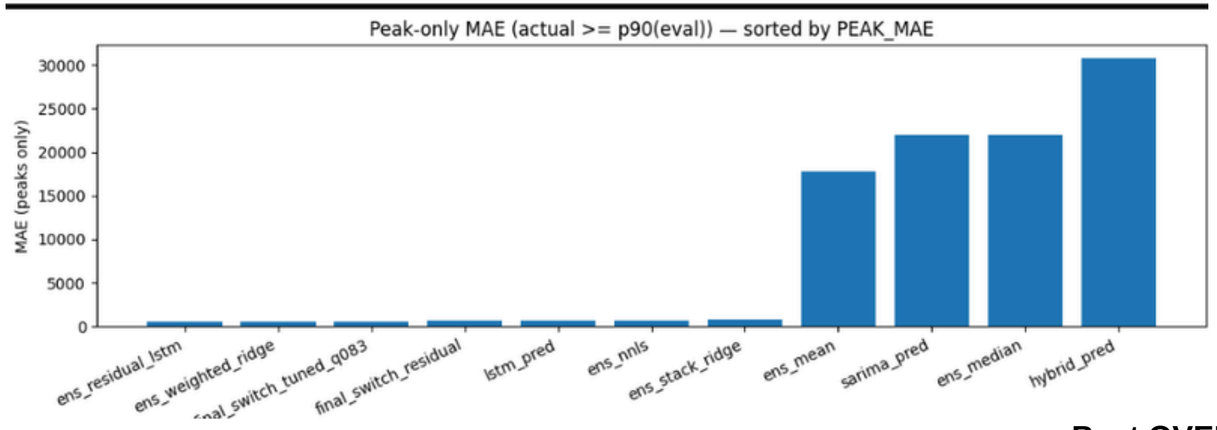




Graphs Analysis & Results: Ensemble Methods



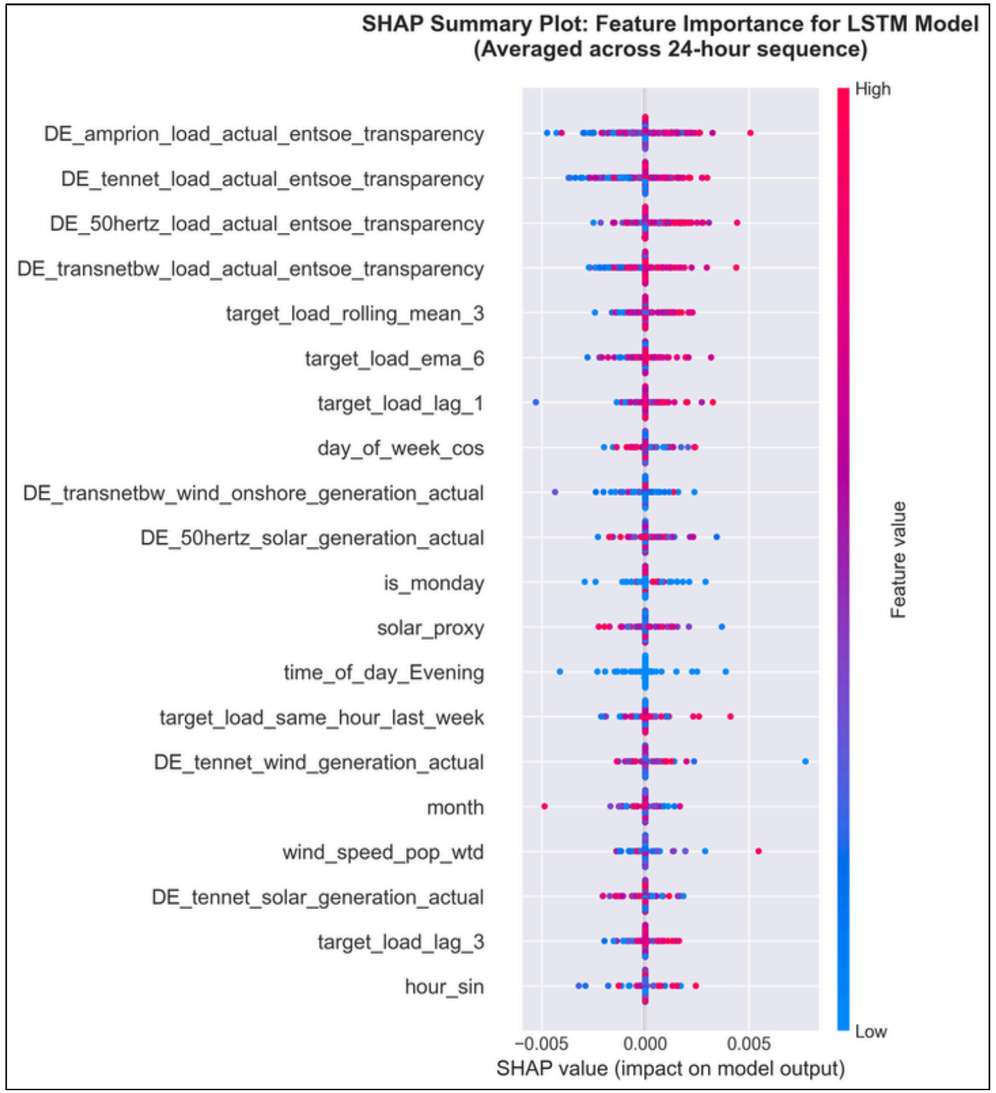




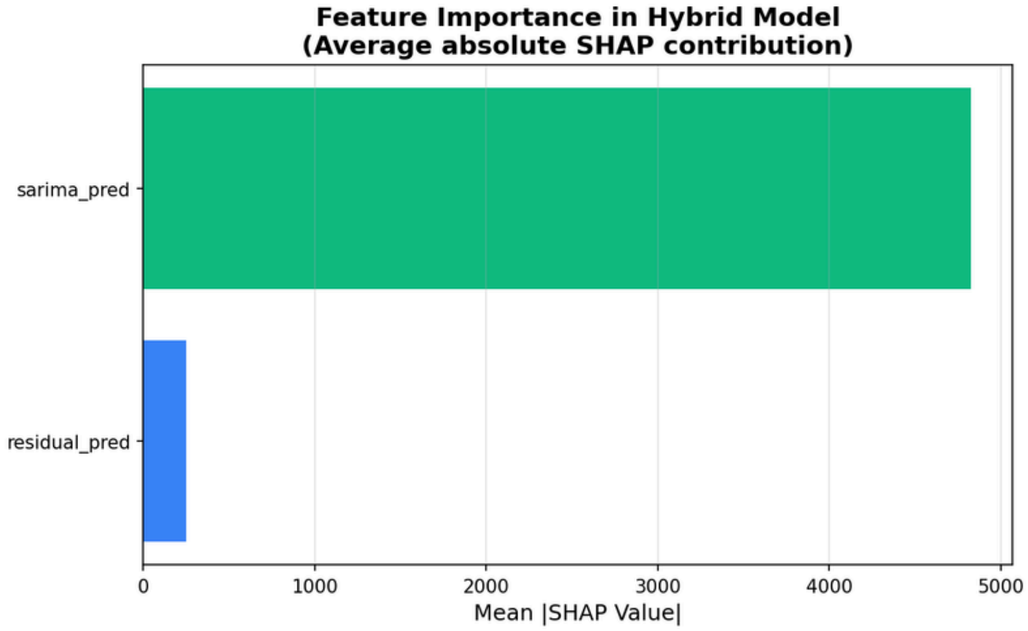
Best OVERALL: **final_switch_residual | OVERALL_MAE: 577.09 | PEAK_MAE: 529.9**

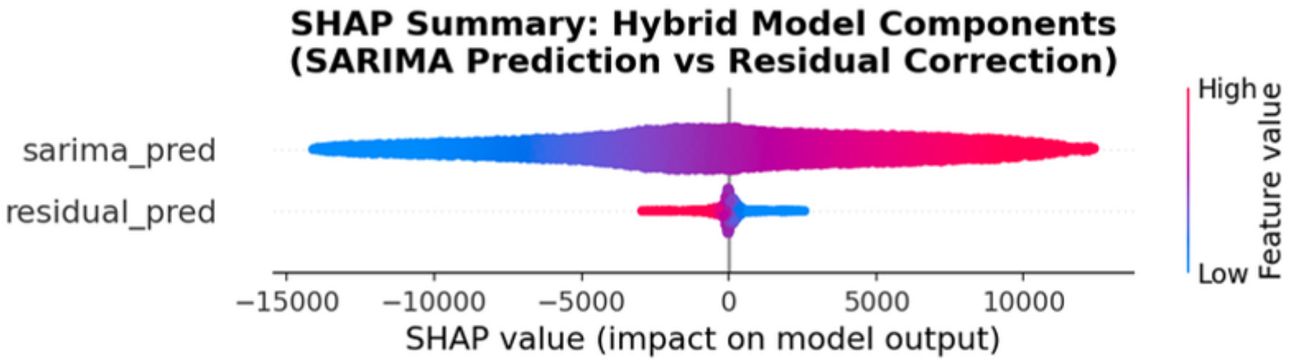
Best PEAK: **final_switch_residual | OVERALL_MAE: 577.08 | PEAK_MAE: 529.9**

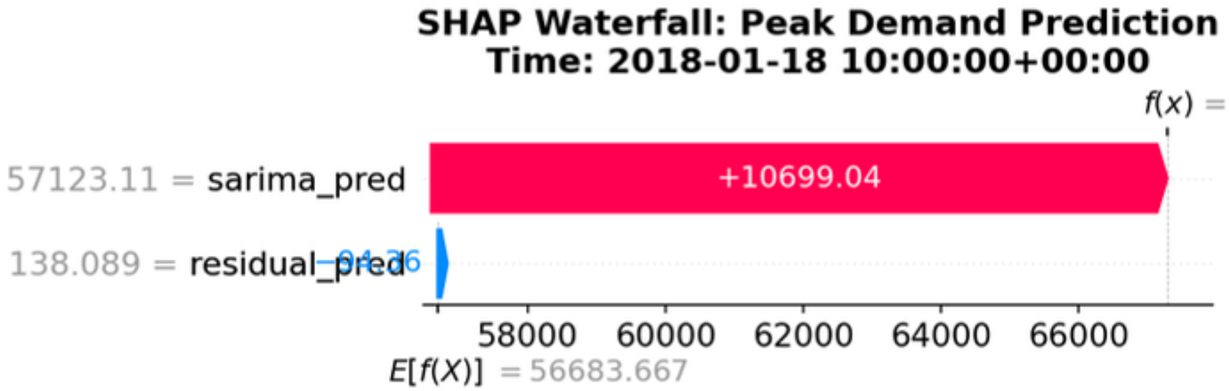
XAI with SHAP: LSTM

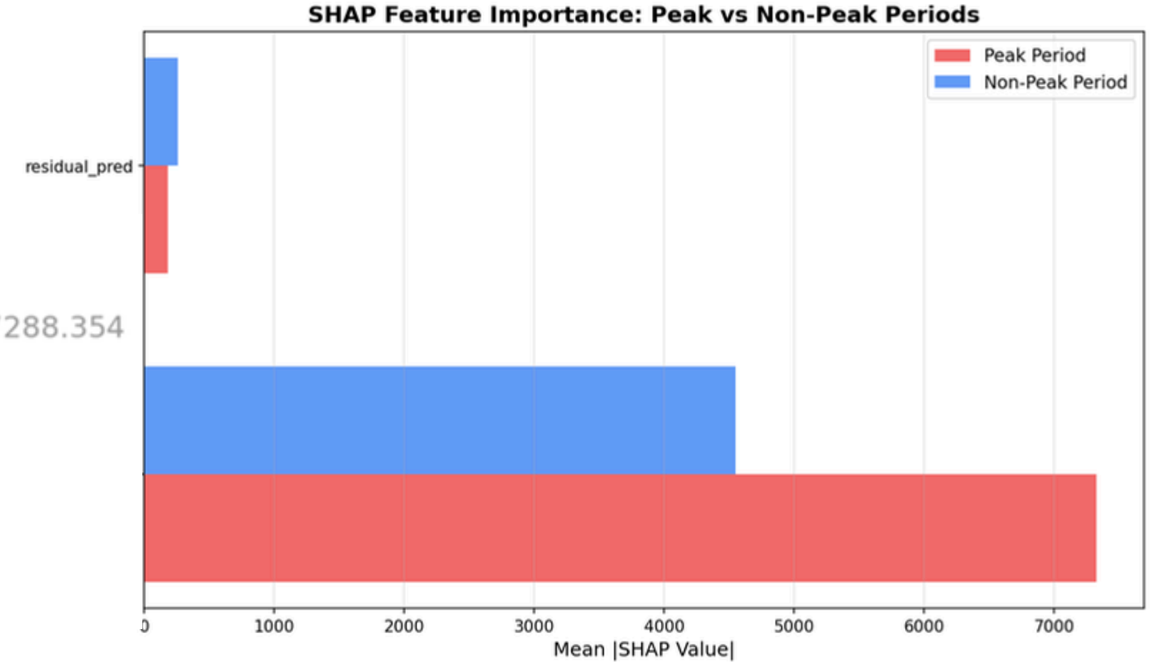


XAI with SHAP: Hybrid

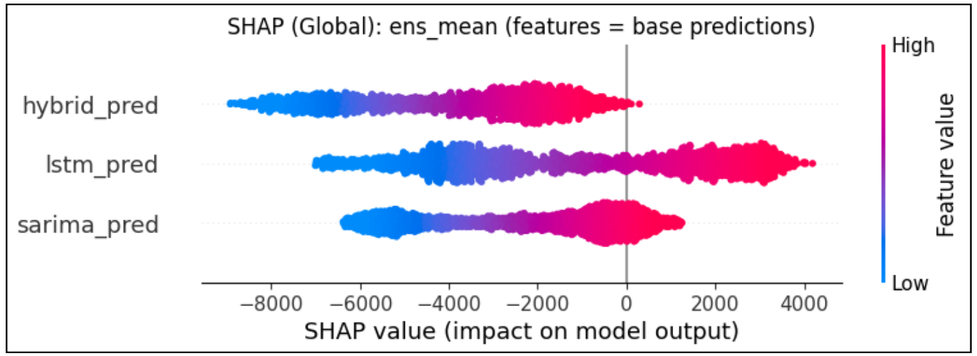


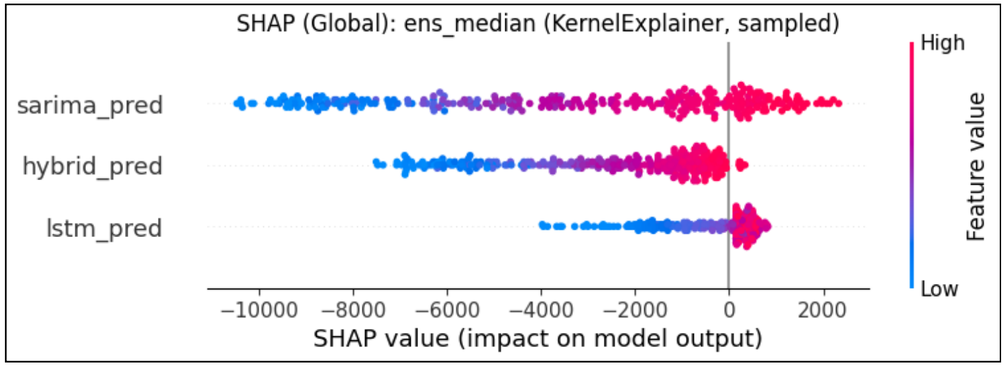


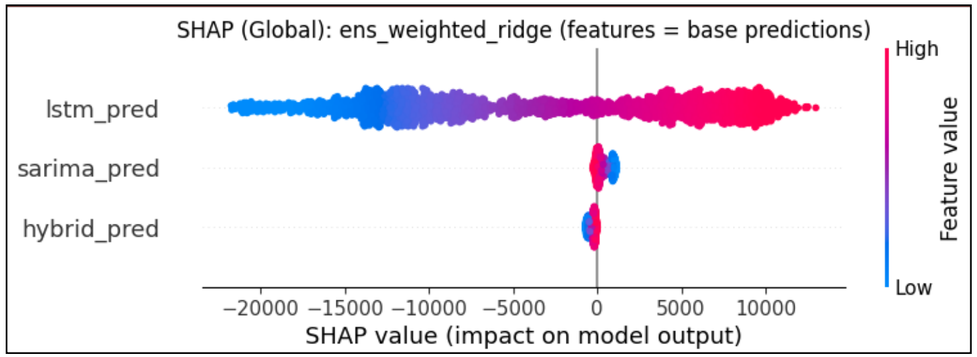


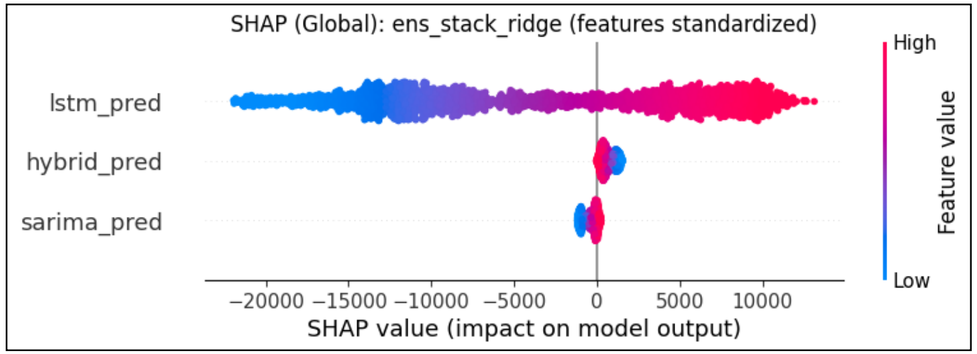


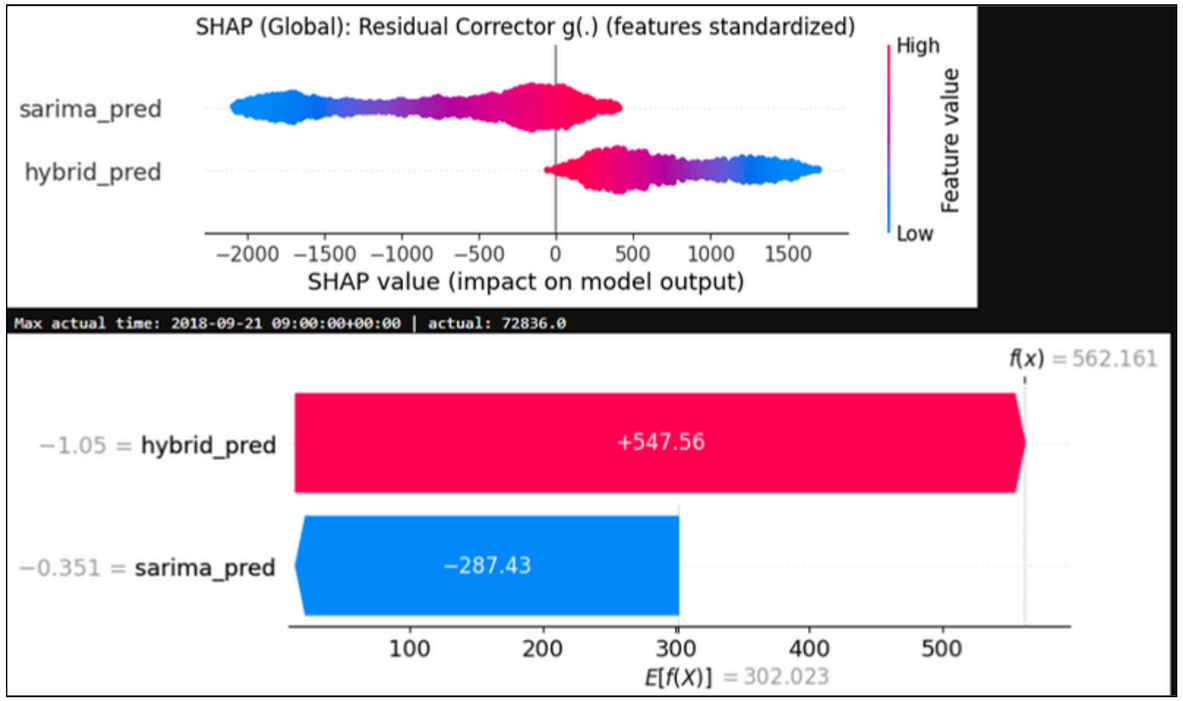
XAI with SHAP: Ensemble Methods





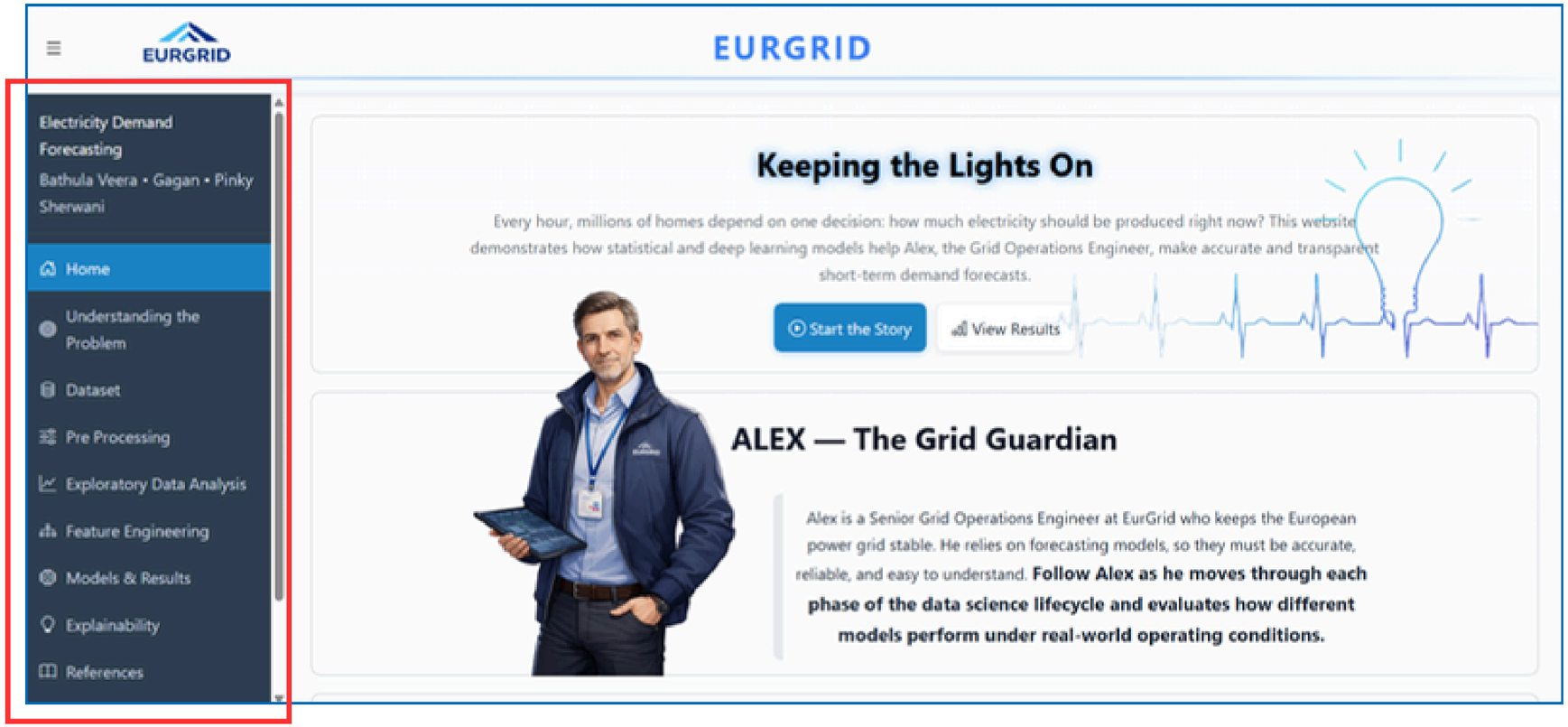






Website Overview

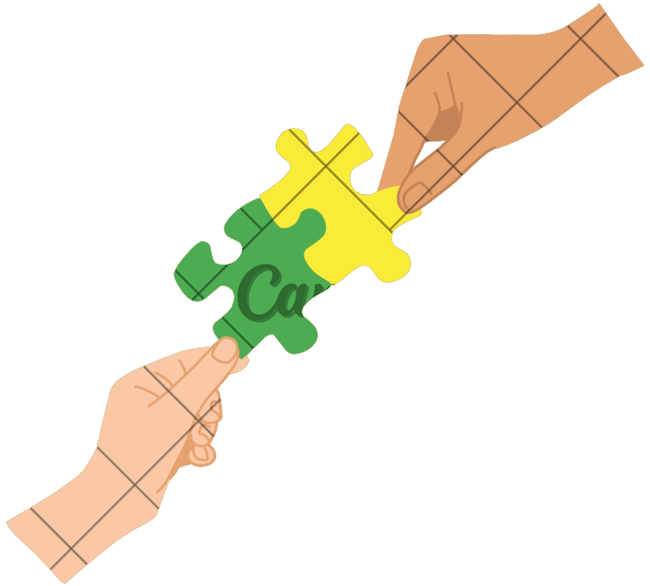
- Public narrative presentation for entire forecasting workflow
- Designed to tell a story from problem to results through interactive content and visuals
- The website uses a fictional grid operations engineer, **"Alex,"** to guide users through stages of the project
- Tailored narrative connects:
  - real grid challenges
  - data science workflow
  - forecasting results
- For audience to see the "why" behind each step



Work Done and Challenges Faced

How Work was Divided?

- Data Preprocessing and EDA - Gagan, Pinky, Raghav
- Modeling Approaches and their XAI Analysis:
  - SARIMA and LSTM: Gagan
  - Hybrid: Pinky
  - Ensemble Methods: Raghav
- Website: Pinky, Raghav, Gagan



Challenges Faced

- Dataset limitations where certain control regions were available after September 2018.
- Inclusion of weather data and creation of proxies.
- Full dataset and weekly seasonality causing memory constraint in Sarima.
- Increasing LSTM layers and bidirectional causing computational constraints.
- Implementation of SHAP for statistical model like Sarima.
- Due to lack of computational power, SARIMA model could not be used with seasonal period of 1 week~168 days

Conclusion &Future Work

Conclusion:

- SARIMA underfits badly and is only a weak seasonal baseline
- LSTM performs very well, capturing non-linear demand patterns
- Hybrid is worst and doesn't generalize here
- Switch-based ensemble is best model in both **Overall and Peak** with the lowest MAE/RMSE and highest R^2

Future Work:

- Use GRU/CNN hybrid ensembles
- Deploy for real-time use and update the model continuously with new data.
- Build a leakage-free setup and use SHAP to clearly explain predictions.

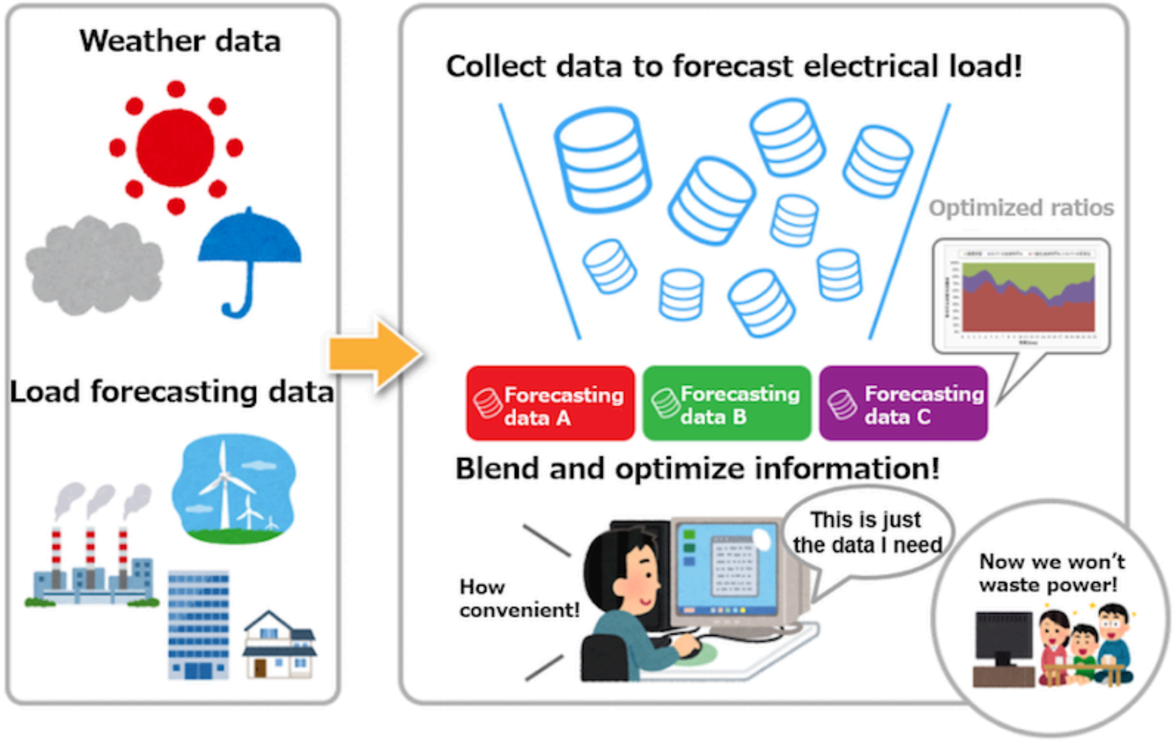


Figure: XAI intio Electricity Load Forecasting
Ref:Toshiba

References

1. Akhtar, S., et al. (2023). Short-Term Load Forecasting Models: A Review of Challenges, Progress, and the Road Ahead. Energies.
2. Tzelepi, M., et al. (2023). Deep Learning for Energy Time-Series Analysis and Forecasting.
3. Aksöz, A., et al. (2024). Analysis of SARIMA Models for Forecasting Electricity Demand.
4. Qureshi, M., et al. (2024). Deep learning-based forecasting of electricity consumption.
5. Open Power System Data. Time Series (Version 2020-10-06). DOI: 10.25832/time_series/2020-10-06
6. Habbak H, Mahmoud M, Metwally K, Fouda MM, Ibrahem MI. Load Forecasting Techniques and Their Applications in Smart Grids. Energies. 2023; 16(3):1480. https://doi.org/10.3390/en16031480

THANK YOU

OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

Q & A