Fortum Energy Consumption Forecasting Solution

An Experimental Exploration of AI in Time Series

Disclaimer: The Future of Data Science?

This project is an experiment to test the applicability of **"Vibe Coding"** to data science–specifically, exploring if a purely AI-driven approach can be the future of model development.

The Challenge: Dual Forecasting Needs

- **Operational (48-Hour):** Need for high-frequency, near-term accuracy for energy trading and grid balancing.
- **Strategic (12-Month):** Need for stable, interpretable long-term planning for budgeting and capacity decisions.

Our Solution: A dual-model approach optimized for each horizon, acknowledging data constraints.

- **Short-Term (48h):** SARIMA (High Precision)
- **Long-Term (12m):** XGBoosted Regression (Price-Aware)

Operational Forecast: SARIMA

- **Model: SARIMA** (Seasonal AutoRegressive Integrated Moving Average).
- **Focus:** Capturing strong, predictable hourly and daily cycles (e.g., morning peaks, weekend dips).
- **Rationale:** SARIMA excels at modeling intrinsic time-series behavior (auto-correlation and seasonality).

Data Scope and Constraints:

- **Training Data Scope:** Only the data of the past year was used for training to significantly improve model training time and iteration speed.
- **Exogenous Constraint:** Due to unavailability of reliable future weather data (temperature, precipitation), the model operates as a pure time-series SARIMA, without exogenous variables.

Strategic Forecast: XGBoost (12-Month)

- **Model: XGBoosted Regression Model** (eXtreme Gradient Boosting).
- **Key Feature:** Takes into account the actual prices of energy at that moment to link consumption patterns with market variables.

Data and Preparation:

- **Available Data:** The long-term model utilized **45 monthly data points** for training and evaluation.
- **Data Quality:** The consumption data for winter 2023 has been specifically investigated to ensure model robustness against anomalous events.

Business Advantage:

- Offers higher non-linear predictive power compared to a simple baseline, leveraging price signals for better strategic planning.

Validation and Business Alignment

Validation Strategy:

- Used rigorous **Time Series Cross-Validation** to ensure chronological integrity.
- Primary Metric: **Mean Absolute Percentage Error (MAPE)**. MAPE provides an intuitive, business-friendly measure of forecast accuracy.

Alignment with Fortum's Objectives:

- **SARIMA (48h):** Supports real-time trading and minimizing imbalance costs.
- **XGBoost (12m):** Supports budgeting and strategic hedging by offering a price-aware model for future consumption risk.

Summary: Robust and Actionable Forecasts

The solution delivers two highly specialized tools:

SARIMA (48h)

- High precision.
- Operational decision support.
- Fast training (1 year of data).

XGBoost (12m)

- Price signal leveraged.
- Strategic financial planning.
- Investigated data anomalies (W23 investigation).