Forecasting Household Electricity Usage

This project combines time series forecasting and pattern recognition to analyze residential electricity use. Using machine learning, we develop models for short-term consumption prediction and detect daily usage patterns.

A **by Anish Karnik**

Problem Overview & Opportunity

Smart meters produce massive datasets daily from household consumption.

Most households lack visibility into consumption patterns and usage drivers.

Energy dynamics vary due to weather, appliances, and seasonal behaviors.

Data science offers predictive insights to help save costs by understanding these complex factors.

Data Science Vision

Data Integration

Combining minute-level electricity data with hourly weather metrics

Pattern Exploration

Spotting inefficiencies and unusual consumption trends

Advanced Analytics

Clustering similar usage days and forecasting with time and weather variables

Detecting spikes and anomalies in consumption

Unlock the power of data - uncover hidden patterns, spot inefficiencies, and find fresh ways to slash household electricity waste and costs.

1 Time Series Forecasting

Develop accurate short-term usage models using ARIMA

2 Sesonal Pattern Recognition

Identify seasonal consumption trends through Prophet

3 Clustering

Categorize days by consumption patterns for weekdays, weekends, through KMeans

4 Sub-metering Analysis

Analyze appliance-level usage to pinpoint high-consumption metrics and weather influence through Regression

What Can This Solution Achieve?

Households

Save 10–20% on energy bills through behavior change.

Gain visibility on which appliances and conditions drive high usage.

Utilities

Better demand forecasting reduces grid strain significantly.

More effective time-of-use pricing and strategic planning capabilities.

Environmental

Reduce unnecessary electricity consumption across residential sectors.

Support sustainable, low-carbon lifestyles through data-driven insights.

Understanding the Data

UCI Dataset

Individual household electric power consumption (France, 2006–2010)

- 2M+ rows of 1-minute resolution data, with 6 features
- 3 Sub-metering power features: kitchen, laundry, HVAC systems
- Average Voltage, Average Current and Total Power usage (Target) used for the House

Weather Dataset

Historical hourly weather data from Open-Meteo API which has features such as Temperature, Relative Humidity, Pressure, Cloud Cover, Wind Speed etc. A total of 12 features

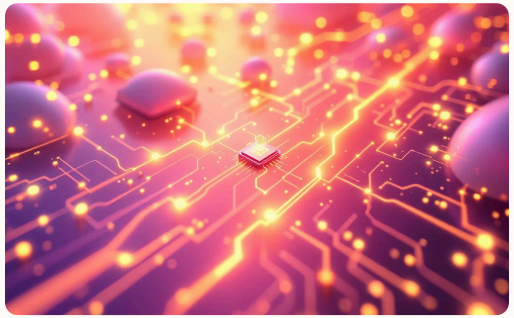
Data Quality Notes

- Missing values and timestamps (e.g., blackout periods)
- Sub-meterings do not fully cover all consumption(eg. lights)
- Weather data has lower time resolution (hourly vs minute)
- Interdependencies like Power= Current * Voltage, Active vs Reactive Power

Early Basic EDA

- Clear daily/weekly usage patterns for Power consumption
- High Power use in colder months — likely heating-related
- Voltage, Current dips/highs correspond to power usage spikes

Next Steps for Data Processing, Feature Engineering, and Modeling



Data Processing

Normalize and scale features weatherdataset features with UCI dataset features   for consistent model input.

Feature Engineering

Create time-based features, rolling averages, and lag variables to boost model accuracy.

Modeling

- ARIMA & Prophet for seasonality forecasting
- KMeans clustering to identify usage patterns
- Baseline prediction via Linear Regression on weather and appliances metrics

Evaluation

Measure regression accuracy using MAE and RMSE metrics.

Thank You!

Comments or Suggestions?