https://www.sciencedirect.com/science/article/pii/S2352484723013975 : Non-linear regression

https://www.neuraldesigner.com/blog/electricity\_demand\_forecasting/ : Neural network

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10963184/ : CNN

https://cienciadedatos.net/documentos/py29-forecasting-electricity-power-demand-python.html : LGBM regressor, evaluate using Back testing, tune using Bayesian search or MPC, prediction intervals to find error interval

https://www.sciencedirect.com/science/article/pii/S266616592030034X : ANN but it might overfit

https://onlinelibrary.wiley.com/doi/10.1155/2024/6812425 : LSTM

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10963184/#:~:text=Sujan%20et%20al.,develop%20this%20innovative%20prediction%20model. : CESN CNN network

Baseline: Previous day

**Predicting Energy Demand Using AI**

**1. Data Preprocessing:**

* **Outlier Detection:**
  + Implement time series decomposition to identify anomalies.
  + Apply the Interquartile Range (IQR) method to detect and filter out outliers.
* **Handling Missing and Outlier Values:**
  + **Energy Demand:** Impute missing values with the corresponding values from the previous week.
  + **Weather Data:** Use linear interpolation to address missing values in weather-related features.

**2. Feature Engineering:**

* **Feature Construction:**
  + Include the difference between the current and previous values for energy demand, temperature, and pressure.
  + Add lag features (energy demand at an earlier timestamp)
  + Include moving averages (the average over the last 24 hours).
  + Include indicators for public holidays.
  + Distinguish between weekdays and weekends.
  + Encode dates into cyclic features using sine and cosine transformations to capture seasonal patterns.
* **Normalization:**
  + Normalize all features using Min-Max scaling to ensure they fall within a standard range.

**3. Feature Selection:**

* **Model-Based Selection:**
  + Train an XGBoost model to determine feature importance.
  + Utilize SHAP (SHapley Additive exPlanations) values to identify and select the most impactful features.

**4. Model Development:**

* **Model Candidates:**
  + Primary models: XGBoost, LightGBM (LGBM), and Artificial Neural Networks (ANN).
  + Optimize and evaluate models based on Mean Squared Error (MSE) and compare them with the bas.
* **Advanced Models (if sufficient data is available):**
  + Explore more complex models such as Convolutional Echo State Networks (CESN), Gated Recurrent Units (GRU), or Long Short-Term Memory networks (LSTM).

**4. Baseline Comparison:**

* **Baseline Model:**
  + **Use the energy demand data from the previous day as a baseline for comparison.**
  + **This baseline will provide a simple reference model, allowing us to evaluate the performance of the advanced models by comparing their predictions against the baseline accuracy.**

**5. Deployment:**

* **Frontend:**
  + user interface in React to provide a dynamic and responsive experience.
* **Backend:**
  + Implement the server-side logic and API endpoints using Django for robust and scalable backend management.