**Project Proposal:** Electricity Price Prediction

**Problem Statement**

The problem at hand is to develop a predictive model that can forecast future electricity prices based on historical data and relevant influencing factors. The primary goal is to create a tool that assists both energy providers and consumers in making informed decisions regarding electricity consumption and investments. The success of this project relies on effective data preprocessing, feature engineering, model selection, training, and evaluation.

**Design Thinking**

1. **Data Source**

The dataset provided at the following link will serve as our primary data source: [Electricity Price Prediction Dataset](). This dataset contains historical electricity prices along with various features such as date, demand, supply, weather conditions, and economic indicators. These features will be essential in building a robust predictive model.

2. **Data Preprocessing**

Before constructing the predictive model, we must prepare the data for analysis. This step includes:

**Data Cleaning**: Handle missing values, outliers, and any inconsistencies in the dataset.

**Data Transformation**: Convert categorical features (if any) into numerical representations using techniques like one-hot encoding or label encoding.

**Scaling**: Normalize or standardize numerical features to ensure consistent scaling across the dataset.

3. **Feature Engineering**

Feature engineering is crucial for improving the predictive power of the model. We will consider the following techniques:

**Time-Based Features**: Extract relevant time-based features such as day of the week, month, and year, which can capture seasonality and trends.

**Lagged Variables**: Create lagged variables (past electricity prices) to account for autocorrelation in time series data.

Weather and Economic Indicators: Explore how external factors like weather conditions and economic indicators influence electricity prices.

4. **Model Selection**

Selecting the appropriate forecasting model is pivotal for accurate predictions. We will consider various time series forecasting algorithms, including but not limited to:

**ARIMA** (AutoRegressive Integrated Moving Average): A classic model for time series forecasting.

**LSTM** (Long Short-Term Memory): A deep learning model known for its ability to capture long-term dependencies in sequential data.

We will experiment with multiple models and select the one that demonstrates the best predictive performance.

5**. Model Training**

Once the model is chosen, it will be trained using the preprocessed dataset. This step involves splitting the data into training and testing sets, selecting appropriate hyperparameters, and training the model on historical data.

6**. Evaluation**

To assess the model's performance, we will employ suitable time series forecasting metrics, such as:

**Mean Absolute Error (MAE):** Measures the average magnitude of errors between predicted and actual values.

**Root Mean Squared Error (RMSE):** Provides a measure of the model's accuracy while penalizing larger errors more significantly.

**Mean Absolute Percentage Error (MAPE):** Calculates the percentage error between predicted and actual values.

We will use these metrics to gauge the model's effectiveness in forecasting electricity prices.

**Conclusion:**

In summary, this project aims to develop a predictive model for electricity price forecasting. By following a systematic approach involving data preprocessing, feature engineering, model selection, training, and evaluation, we intend to create a valuable tool for energy providers and consumers to make informed decisions. The success of this project will rely on the quality of data preparation and the selection of the most suitable forecasting model.