# Project Proposal: Predicting Electricity Prices using Data Science

## Table of Contents



### **Introduction**:

An electricity price prediction project using data science aims to forecast future electricity prices by analyzing historical data and relevant factors. This project is essential for energy providers, consumers, and traders to make informed decisions. It involves steps like collecting and cleaning data, selecting appropriate models, training and fine-tuning them, and deploying the model for real-time predictions. This project helps stakeholders better understand and anticipate electricity price fluctuations, enabling them to manage their energy usage and trading strategies effectively.

### **Problem Definition:**

The problem is to develop a predictive model that uses historical electricity prices and relevant factors to forecast future electricity prices. The objective is to create a tool that assists both energy providers and consumers in making informed decisions regarding consumption and investment by predicting future electricity prices.

## Design Thinking:

**Data sources:**

Utilize a dataset containing historical electricity prices and relevant factors like -date
-demand and supply
-weather conditions and
-economic indicators.

**Data Preprocessing:**

Clean and preprocess the data, handle missing values, and convert categorical features into numerical representations.

**Feature Selection:**

Create additional features that could enhance the predictive power of the model, such as time-based features and lagged variables.

**Model Selection:**

Choosing the right time series forecasting algorithm is crucial for accurate predictions. We will consider
various forecasting algorithms, including:
-ARIMA(AutoRegressive Integrated Moving Average)
- LSTM(Long Short-Term Memory networks)
The selection will be based on their suitability for the problem and performance during evaluation.

**Model Training:**

Once we have selected an appropriate forecasting algorithm, we will train the model using the preprocessed dataset. This step involves splitting the data into training and testing sets to assess the model's performance.

**Evaluation:**

Evaluate the model's performance using appropriate time series forecasting metrics:

**1.Mean Absolute Error (MAE):**

Measures the average absolute difference between predicted and actual prices.

**2.Root Mean Squared Error (RMSE):**

Measures the square root of the mean squared differences between predicted and actual prices.

**3.R-squared (R^2):**

Indicates the proportion of variance in the target variable that is predictable from the input features.

The model will be fine-tuned and iteratively improved based on the evaluation results.

## Project Timeline:

To effectively manage the project, we will establish a timeline with milestones for each phase, including data acquisition, preprocessing, model development, and evaluation. A detailed project plan will be created to track progress.

## Conclusion:

In conclusion, an electricity price prediction project using data science is a valuable tool for the energy industry. By harnessing historical data and advanced modeling techniques, this project empowers stakeholders to make informed decisions, optimize energy consumption, and enhance trading strategies. It contributes to a more efficient and responsive energy market, ultimately benefiting both providers and consumers.