# Electricity Demand and Price Forecasting

## Project Overview

The "Electricity-Demand-and-Price-Forecasting" project is a crucial initiative aimed at enhancing our understanding of electricity markets through advanced predictive analytics. In an era where energy consumption is continuously rising and renewable energy sources are being integrated into the grid, accurate forecasting of electricity demand and prices becomes essential. This project employs machine learning techniques, including neural networks, to predict electricity prices for the day-ahead markets, thereby enabling stakeholders to make informed decisions.

The significance of this project extends beyond mere prediction; it plays a vital role in optimizing electricity storage and consumption. By accurately forecasting electricity prices, consumers and energy providers can adjust their usage patterns, reducing consumption during peak price periods and potentially increasing it during times of lower prices. This behavior not only leads to cost savings but also promotes a more sustainable energy ecosystem by minimizing waste and facilitating better utilization of resources.

Furthermore, the correlation between electricity prices and carbon intensity highlights the environmental benefits of effective forecasting. By reacting to price fluctuations, energy consumers can reduce their carbon footprint, contributing to broader climate change mitigation efforts. The ability to forecast demand and prices accurately can improve demand-side flexibility, allowing buildings and facilities to operate more efficiently and economically.

In summary, the "Electricity-Demand-and-Price-Forecasting" project serves as a foundational effort to enhance the efficiency of energy markets. Through sophisticated modeling and predictive techniques, it aims to empower users to optimize their electricity consumption while supporting sustainability in energy production and use.

## Data Collection

The success of the Electricity Price Forecasting project largely depends on the quality and relevance of the data collected. For this initiative, data was sourced from multiple avenues, primarily focusing on electricity prices, commodity prices, and weather conditions, which are crucial for accurate predictive modeling.

Hourly electricity prices were obtained from Nordpool, a significant player in European electricity markets. The data encompassed a timeframe from 2013 to 2019, providing a robust dataset for analysis. This extensive period allowed for the identification of trends and patterns in electricity pricing, essential for developing effective predictive models. The data was collected through automated scripts that interfaced with Nordpool's API, ensuring precise and timely retrieval of the latest price information.

In addition to electricity prices, external factors influencing these rates were also considered. Daily commodity prices for coal, natural gas, uranium, and oil were gathered, as these resources significantly impact electricity production costs. The data collection for these commodities involved scraping various financial and market data websites, ensuring comprehensive coverage of the market dynamics affecting electricity prices.

Weather data was another critical component of the dataset. Collected using the DarkSky API, this information included hourly temperature readings that were vital for understanding the correlation between weather conditions and electricity demand. The integration of weather data into the model allowed for a more nuanced analysis of how temperature variations influence electricity consumption patterns.

Overall, the data collection phase was meticulously executed to ensure a diverse and representative dataset, setting a solid foundation for the predictive modeling efforts that followed. By leveraging historical data from reliable sources, the project aimed to enhance the accuracy of electricity price forecasts and support better decision-making in energy management.

## Methodology

The methodology employed in the Electricity Price Forecasting project encompasses a range of statistical and machine learning techniques designed to maximize the accuracy and reliability of predictions regarding electricity prices. This multifaceted approach integrates various methodologies, including inferential statistics, machine learning algorithms, deep learning methods, and rigorous data preprocessing steps.

### Inferential Statistics

Inferential statistics played a pivotal role in the initial stages of the project. By analyzing historical data, we were able to derive insights and establish correlations between various factors affecting electricity pricing. This statistical groundwork helped to formulate hypotheses that guided the selection of features for subsequent modeling.

### Machine Learning Techniques

Several machine learning techniques were employed to develop predictive models. Among these, Random Forests and XGBoost were highlighted for their robustness and effectiveness in handling complex datasets. Random Forests, an ensemble learning method, helped capture nonlinear relationships and interactions among variables, while XGBoost utilized a gradient boosting framework to enhance predictive performance. Both methods were pivotal in establishing baseline models that informed further refinements.

### Deep Learning Methods

Deep learning techniques, particularly Neural Networks, were employed to harness the power of large datasets and complex feature interactions. Utilizing libraries such as Keras and TensorFlow, we built models capable of processing vast amounts of input data, including time-series data from the previous week. The neural networks were designed to predict future electricity prices based on historical trends, effectively capturing the temporal dependencies inherent in the data.

### Time-Series Analysis

Given the nature of the data, time-series analysis was a critical component of the methodology. This involved reshaping the data to suit the requirements of different models, ensuring that temporal patterns were adequately captured. The analysis included techniques for handling seasonality and trends, which are essential in forecasting electricity prices accurately.

### Data Preprocessing Steps

Data preprocessing was crucial in ensuring the quality and usability of the dataset. This phase included extensive data cleaning and wrangling, addressing issues such as missing values, outliers, and inconsistencies. The data was transformed and normalized to facilitate effective model training. Additionally, exploratory data analysis (EDA) was conducted to uncover underlying patterns and relationships, which informed feature selection and engineering.

In summary, the methodology of this project combines statistical analysis, machine learning, and deep learning techniques, alongside robust data preprocessing, to create a comprehensive framework for predicting electricity prices in the day-ahead markets.

## Model Development

In the pursuit of accurate electricity price predictions, various models were developed, ranging from simple baseline methods to sophisticated machine learning and deep learning techniques. Each model was trained and evaluated to determine its effectiveness in forecasting future prices.

### Naive Methods

The naive method served as a baseline model. This straightforward approach predicted future prices based solely on the most recent price data available. Specifically, it assumed that the electricity price 24 hours ahead would be equal to the current price. While simple, this method provided a reference point against which more complex models could be evaluated. The evaluation metric for this model was Mean Absolute Percentage Error (MAPE), which highlighted its limitations when faced with fluctuations in electricity prices.

### ARIMA (AutoRegressive Integrated Moving Average)

ARIMA is a traditional statistical approach widely used for time series forecasting. The model captures the temporal dependencies in the data by using past values (autoregression), differencing to make the time series stationary (integration), and lagged forecast errors (moving average). For this project, the ARIMA model was trained using historical electricity price data from Nordpool. The model parameters were selected through grid search techniques, and validation was performed using the 2018 dataset, with results assessed using MAPE to evaluate forecasting accuracy.

### XGBoost

XGBoost, or Extreme Gradient Boosting, is an advanced machine learning technique that excels in predictive accuracy and efficiency. This model was particularly useful for capturing complex interactions among features. Training involved utilizing the historical electricity price data along with external variables such as commodity prices and weather data. XGBoost's hyperparameters were optimized through cross-validation, ensuring robust performance. The model's effectiveness was assessed through various metrics, including MAPE and RMSE (Root Mean Square Error), allowing for a comprehensive evaluation of its predictive capability.

### Neural Networks

Deep learning methods, particularly neural networks, were employed to enhance the model's ability to capture intricate patterns within the data. The architecture involved multiple layers with numerous nodes, enabling the model to learn complex relationships from a vast amount of input data. Training was conducted using historical data, where the model ingested time-series data from the past week to predict prices for the subsequent 24 hours. The use of Keras and TensorFlow facilitated the training process on Google Colab, leveraging TPU for faster computations. Evaluation metrics such as MAPE were utilized to assess the accuracy of the predictions, allowing for direct comparison with other models.

Overall, this diverse array of models provided a comprehensive framework for predicting electricity prices. Each model's unique strengths contributed to the project's goal of accurate forecasting, ultimately enabling stakeholders to make data-driven decisions in the energy market.

## Analysis of Results

The results from various predictive models showcased significant variations in forecasting accuracy, highlighting the nuances inherent in electricity price prediction. Among the methods employed, the Mean Absolute Percentage Error (MAPE) emerged as a principal measure of accuracy. MAPE provided a clear percentage-based indicator of the average error, facilitating an intuitive understanding of model performance. However, challenges arose due to the presence of zero values within the dataset, which complicated the MAPE calculations. In instances where actual prices dipped to zero, the division by zero issue led to undefined values, necessitating additional strategies to manage these anomalies.

In terms of effectiveness, neural networks outperformed traditional methods like ARIMA and XGBoost in capturing complex patterns in the data. The neural network model, which processed a week’s worth of hourly data to predict the upcoming 24 hours, demonstrated remarkable accuracy, particularly in volatile market conditions. Its ability to learn nonlinear relationships and interactions from vast datasets allowed it to adapt more effectively to fluctuations in electricity prices.

The comparative analysis revealed that while ARIMA provided a solid foundation in time-series forecasting, it struggled with the non-stationarity of electricity prices. XGBoost, characterized by its gradient boosting framework, excelled in handling feature interactions but still fell short when addressing the temporal dependencies critical in this context. In contrast, neural networks showcased their robustness in learning from historical data, allowing them to make more informed predictions.

Despite the successes of the neural network approach, it was not without challenges. The model's complexity required substantial computational resources and careful tuning of hyperparameters. Moreover, the necessity for extensive data preprocessing was crucial to ensure the integrity of inputs, particularly when dealing with outliers and missing values. Overall, the integration of various modeling techniques provided invaluable insights, illustrating the strengths and limitations of each method in the pursuit of accurate electricity price forecasting.

## Battery Function Model

The development of a simple battery function model is a pivotal component of the Electricity Price Forecasting project, designed to leverage the predictive capabilities achieved through advanced modeling techniques. This battery function operates on the premise of optimizing energy storage and consumption based on real-time electricity prices, ultimately providing additional value for both consumers and energy suppliers.

The primary purpose of the battery function model is to simulate the behavior of a battery system that stores electricity during periods of low prices and discharges it when prices are high. By accurately predicting the day-ahead electricity prices, the model enables stakeholders to make informed decisions about when to charge or discharge their battery systems. This time-sensitive approach maximizes the economic value derived from energy storage, allowing users to take advantage of price fluctuations in the electricity market.

To build this model, the team utilized the predictive outputs from the neural network and other machine learning models developed during the project. The neural network's ability to capture complex patterns in electricity pricing, informed by historical data, serves as the foundation for the battery function's operational logic. For instance, when the model forecasts that electricity prices will be lower during the upcoming hours, it signals the battery to charge, effectively storing energy for later use. Conversely, when a rise in predicted prices occurs, the model prompts the battery to discharge, allowing users to benefit from selling stored energy back to the grid at a higher price.

The implementation of this battery function model not only enhances economic efficiency but also contributes to a more sustainable energy ecosystem. By optimizing energy storage based on price signals, the model supports demand-side management, helping to balance supply and demand in real-time. Additionally, this aligns with broader environmental goals, as utilizing stored energy reduces reliance on fossil fuel generation during peak demand periods, thus lowering greenhouse gas emissions.

Overall, the battery function model stands as a practical application of the project's predictive analytics, showcasing how advanced data-driven insights can translate into tangible benefits for energy consumers and providers alike.

## Conclusions

The Electricity Price Forecasting project has yielded significant insights and findings regarding the prediction of electricity prices, which hold crucial implications for energy management and policy. The project utilized advanced machine learning techniques, particularly neural networks, to effectively forecast day-ahead electricity prices. The results demonstrated that these models can capture complex patterns and fluctuations in electricity pricing, outperforming traditional methodologies such as ARIMA and XGBoost.

Key findings indicate that accurate electricity price predictions can optimize energy consumption, allowing both consumers and utility providers to adjust their usage patterns according to price signals. This optimization can lead to considerable cost savings and increased revenue opportunities, particularly during periods of high price volatility. Moreover, by facilitating demand-side flexibility, the project highlights the potential for reducing carbon emissions associated with electricity generation. Consumers can strategically lower their energy use during peak pricing periods, thus contributing to environmental sustainability efforts.

Future improvements in this domain may include exploring more sophisticated modeling techniques, such as ensemble methods that combine various predictive models to enhance accuracy further. Additionally, incorporating real-time data streams, such as instant market fluctuations and emerging renewable energy sources, could refine predictions and enable more responsive energy management strategies.

Further research could also focus on the integration of these predictive models with energy storage systems, as demonstrated by the battery function model developed in this project. This integration could lead to innovative solutions that maximize the economic and environmental benefits of energy consumption.

Overall, the project has established a strong foundation for future developments in electricity price forecasting and its applications within energy management, highlighting the importance of continued research and technological advancement in this critical area.

## Appendices

This section includes supplementary materials and relevant resources that were utilized throughout the Electricity Price Forecasting project. The documentation provided here serves as a reference for code repositories, presentation documents, and data processing scripts that were essential for the project's execution and analysis.

### Data Processing Scripts

* **Data Collection Scripts**: The scripts responsible for gathering hourly electricity prices, commodity prices, and weather data are included in the repository. These scripts utilize APIs and web scraping methods to ensure accurate and timely data retrieval.
* **Data Cleaning and Transformation**: Within the repository, you will find notebooks dedicated to data cleaning and preprocessing. These notebooks detail the methodologies used to handle missing values, outliers, and other data inconsistencies to prepare the dataset for modeling.

### Featured Notebooks

* **Initial EDA**: The Initial\_eda.ipynb notebook details the early exploratory data analysis conducted to understand the dataset better and inform feature selection.
* **Model Analysis**: The ResultAnalysis.ipynb notebook provides a comprehensive analysis of the results from various models, including performance metrics and visualizations.
* **Battery Function Model**: The Battery.ipynb notebook outlines the development and implementation of the battery function model, demonstrating its application in optimizing energy storage based on predicted electricity prices.