Electricity Prices Prediction Project Documentation

Problem Statement

The problem addressed in this project is to develop a predictive model that can forecast future electricity prices. The objective is to provide a tool that aids energy providers and consumers in making informed decisions regarding consumption and investments by predicting future electricity prices. The project involves data preprocessing, feature engineering, model selection, training, and evaluation.

Design Thinking Process

Data Source: Utilized a dataset containing historical electricity prices and relevant factors like date, demand, supply, weather conditions, and economic indicators.

Data Preprocessing: Cleaned and preprocessed the data, handled missing values, and converted categorical features into numerical representations.

Feature Engineering: Created additional features like time-based features and lagged variables to enhance the predictive power of the model.

Model Selection: Explored time series forecasting algorithms such as ARIMA, LSTM, Prophet, and more for predicting future electricity prices.

Model Training: Trained the selected model using the preprocessed data.

Evaluation: Assessed the model's performance using appropriate time series forecasting metrics, including Mean Absolute Error, Root Mean Squared Error, and Mean Absolute Percentage Error.

Dataset Description

The dataset used for this project contains historical electricity prices and relevant factors like date, demand, supply, weather conditions, and economic indicators. It serves as the foundation for developing the electricity price prediction model. The data was sourced from Kaggle.

Data Preprocessing

The data preprocessing steps included handling missing values, removing duplicates, converting categorical features into numerical representations, and creating additional time-based features.

Model Training and Selection

We considered various time series forecasting models, including ARIMA, LSTM, and Prophet. The model training process involved splitting the data into training and testing sets, selecting appropriate hyperparameters, and training the models using historical data.

Model Evaluation

The models were evaluated using standard time series forecasting metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). These metrics were used to assess the accuracy of the models in predicting electricity prices.

Conclusion:

In conclusion, this project marks a significant step towards building a predictive tool that can benefit both energy providers and consumers by assisting in making informed decisions regarding electricity consumption and investments. The journey involved a range of data science and time series forecasting techniques, ultimately contributing to a deeper understanding of electricity price dynamics. As future work, we can further refine the model, explore advanced algorithms, and enhance its robustness to create a more accurate electricity price prediction tool.