

Project: Electricity Prices Prediction

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 tool that assists both energy providers and consumers in taking informed decisions regarding consumption and investment by predicting future electricity prices. This project involves data processing, feature engineering, model selection, training, and evaluation.

Solution:

Step-1: Data Processing

1. When cleaning the data, it is important to identify and remove any outliers.
2. When handling missing values, you need to decide whether to impute the missing values or drop the rows with missing values.
3. When converting categorical features into numerical representations, you can use a variety of methods, such as one-hot encoding or label encoding.

Step-2: Feature Engineering

1. We can also use feature selection techniques to identify the most important features for predicting electricity prices.
2. This will help to improve the performance of your model and reduce the risk of overfitting.

Step-3: Model Selection

The Suitable time forecasting algorithm would be Autoregressive Integrated Moving Average (ARIMA) for predicting the future electricity prices.

The ARIMA time series model is a statistical model that is used to forecast future values of a time series. The model is based on the assumption that the future values of the time series are a linear function of its past values and its past errors.

The ARIMA model is characterized by three parameters:

p: The order of the autoregressive term. This parameter represents the number of past values of the time series that are used to predict the future value.

d: The order of differencing. This parameter represents the number of times that the time series needs to be differenced in order to make it stationary.

q: The order of the moving average term. This parameter represents the number of past errors that are used to predict the future value.

The ARIMA model is trained by fitting the model to the historical data. The model parameters are selected to minimize the error between the predicted and actual values of the time series.

Once the ARIMA model is trained, it can be used to predict future values of the time series. To predict the future value, the model is given the most recent values of the time series and its past errors. The model then uses these values to predict the future value of the time series.

Step-4: Model training and Evaluation

Once we have selected a model, we need to train it on the data. This involves feeding the model the data and allowing it to learn the relationships between the different variables. Once the model is trained, we need to evaluate its

performance on a held-out test set. This will give us an idea of how well the model will generalize to new data.

Step-5: Model Deployment

Once we are satisfied with the performance of the model, we can deploy it to production. This may involve integrating the model into a software application or making it available as a web service

Conclusion:

This document has outlined the design thinking process for developing an electricity price prediction model using ARIMA time series forecasting algorithms.

The document has also covered the key steps involved in data preprocessing, feature engineering, model selection, training, and evaluation.