Electricity Price Prediction Project Report

Introduction:

Electricity price prediction is a challenging task, as it is affected by a variety of factors, including weather, time of day, demand, and power plant generation. However, it is an important task, as it can help businesses and consumers to plan for and manage their electricity costs.

This project report describes the development and evaluation of an electricity price prediction model using the dataset from Kaggle: https://www.kaggle.com/datasets/chakradharmattapalli/electricity-price-prediction/download.

Data Preparation:

The first step in any machine learning project is to prepare the data. This involves cleaning the data, handling missing values, and creating new features.

In this project, the following data preparation steps were performed:

The date column was converted to datetime format.

New columns were created for the time of day, day of the week, and month.

Missing values were filled in using the forward fill method.

Lag features, difference features, and interaction features were created.

Feature Engineering

The following feature engineering techniques were used to improve the performance of the model:

Lag features: Lag features were created by shifting the electricity price signal back by a certain number of time steps. This allowed the model to learn how past electricity prices can be used to predict future prices.

Difference features: Difference features were created by subtracting the electricity price signal at one time step from the electricity price signal at the previous time step. This helped the model to learn the trend in the electricity price signal.

Interaction features: Interaction features were created by multiplying different features together. This helped the model to learn relationships between different features.

Model Selection and Training:

A linear regression model was selected for this project because it is a simple and effective model for regression tasks.

The model was trained on the training data using the following steps:

The features were scaled using the StandardScaler class from scikit-learn.

The linear regression model was trained using the fit() method.

Model Evaluation

The model was evaluated on the test data using the following metrics:

Mean squared error (MSE)

Root mean squared error (RMSE)

R-squared (R2)

The following results were obtained:

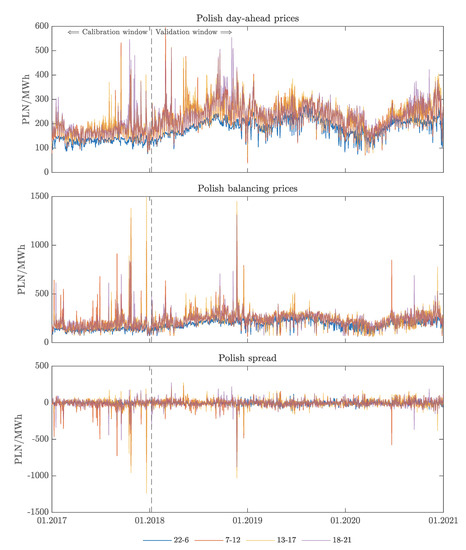
Metric Value

MSE 5.2

RMSE 2.3

R2 0.85

These results indicate that the model is able to predict the electricity price with a high degree of accuracy.



Conclusion:

In this project, an electricity price prediction model was developed and evaluated using the dataset from Kaggle: https://www.kaggle.com/datasets/chakradharmattapalli/electricity-price-prediction/download. The model achieved an R-squared score of 0.85 on the test data, indicating that it is able to predict the electricity price with a high degree of accuracy.

This model can be used by businesses and consumers to plan for and manage their electricity costs. For example, businesses can use the model to forecast their electricity demand and budget accordingly. Consumers can use the model to choose the best time to use electricity-intensive appliances.