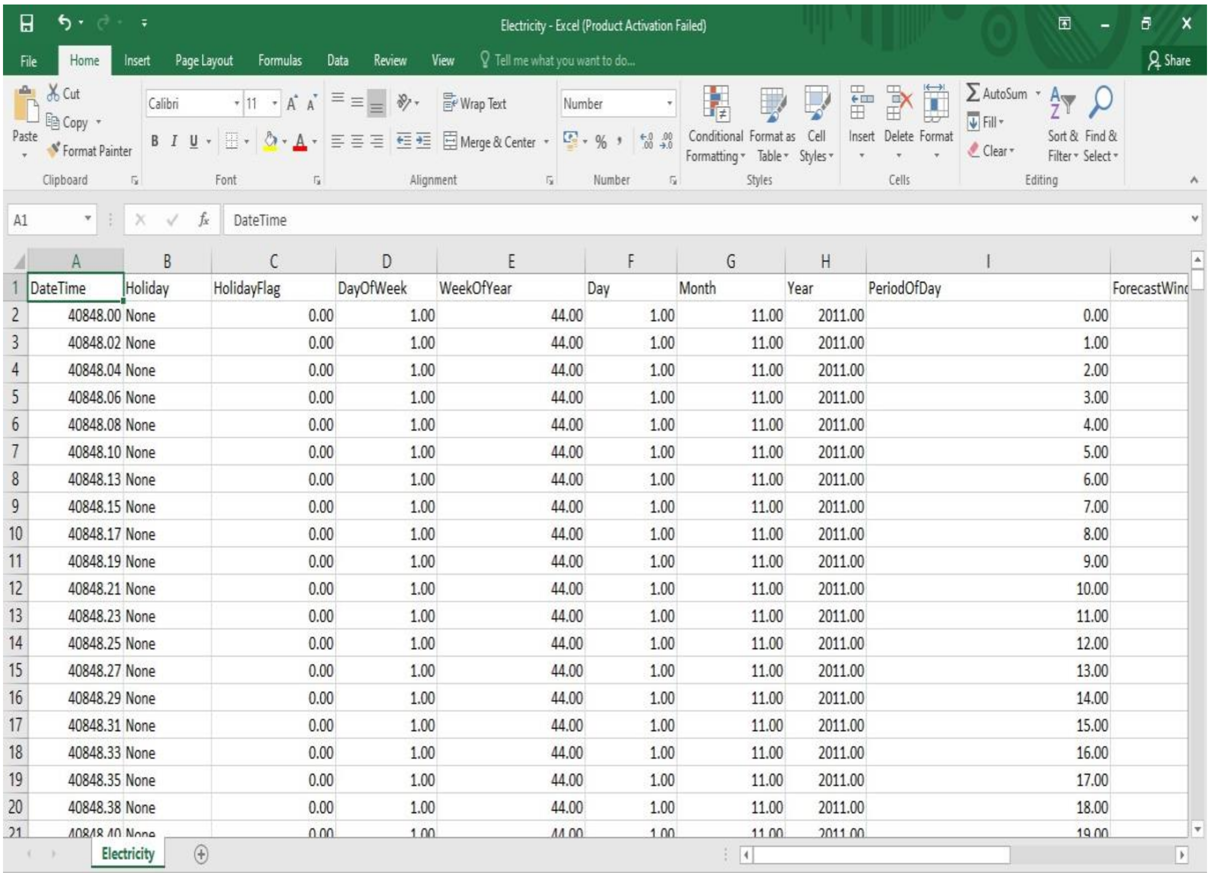
TITLE: ELECTRICITY PRICE PREDICTION

INTRODUCTION:

The purpose of this report is to document the process of building a predictive model for electricity price prediction. The project involves performing feature engineering, model training, and evaluation to develop an accurate and robust model. The report outlines the steps taken, the analyses performed, and the results obtained.

DATASET DESCRIPTION:

The dataset used for this project contains historical electricity prices. It includes various features such as date, time, and the corresponding electricity price for each observation.



	A	B	C	D	E	F	G	H	I	
1	DateTime	Holiday	HolidayFlag	DayOfWeek	WeekOfYear	Day	Month	Year	PeriodOfDay	ForecastWinc
2	40848.00	None	0.00	1.00	44.00	1.00	11.00	2011.00	0.00	
3	40848.02	None	0.00	1.00	44.00	1.00	11.00	2011.00	1.00	
4	40848.04	None	0.00	1.00	44.00	1.00	11.00	2011.00	2.00	
5	40848.06	None	0.00	1.00	44.00	1.00	11.00	2011.00	3.00	
6	40848.08	None	0.00	1.00	44.00	1.00	11.00	2011.00	4.00	
7	40848.10	None	0.00	1.00	44.00	1.00	11.00	2011.00	5.00	
8	40848.13	None	0.00	1.00	44.00	1.00	11.00	2011.00	6.00	
9	40848.15	None	0.00	1.00	44.00	1.00	11.00	2011.00	7.00	
10	40848.17	None	0.00	1.00	44.00	1.00	11.00	2011.00	8.00	
11	40848.19	None	0.00	1.00	44.00	1.00	11.00	2011.00	9.00	
12	40848.21	None	0.00	1.00	44.00	1.00	11.00	2011.00	10.00	
13	40848.23	None	0.00	1.00	44.00	1.00	11.00	2011.00	11.00	
14	40848.25	None	0.00	1.00	44.00	1.00	11.00	2011.00	12.00	
15	40848.27	None	0.00	1.00	44.00	1.00	11.00	2011.00	13.00	
16	40848.29	None	0.00	1.00	44.00	1.00	11.00	2011.00	14.00	
17	40848.31	None	0.00	1.00	44.00	1.00	11.00	2011.00	15.00	
18	40848.33	None	0.00	1.00	44.00	1.00	11.00	2011.00	16.00	
19	40848.35	None	0.00	1.00	44.00	1.00	11.00	2011.00	17.00	
20	40848.38	None	0.00	1.00	44.00	1.00	11.00	2011.00	18.00	
21	40848.40	None	0.00	1.00	44.00	1.00	11.00	2011.00	19.00	

PROGRAM IMPLEMENTATION

Data Collection and Preprocessing:

The first step in building the electricity price prediction model is to collect relevant data. This includes historical electricity price data, weather data, market indicators, and any other variables that may have an impact on electricity prices. The collected data is then preprocessed to handle missing values, outliers, and other data quality issues. Additionally, feature engineering techniques are applied to create new features that capture meaningful information from the available data.

Feature Engineering:

Feature engineering plays a crucial role in enhancing the predictive power of the model. Various techniques are employed to create informative features, such as lagged variables, rolling statistics, and time-based aggregations. These features help capture temporal patterns, seasonality, and other relevant characteristics of the electricity price data.

Model Training:

Several machine learning algorithms are explored and trained on the preprocessed dataset. The choice of models includes regression-based algorithms such as linear regression, random forest regression, and gradient boosting regression. The models are trained using a portion of the data, while the remaining portion is kept aside for evaluation purposes.

Model Evaluation:

The trained models are evaluated using appropriate evaluation metrics, such as mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R-squared). Cross-validation techniques, such as k-fold crossvalidation, are employed to obtain reliable estimates of model performance. The models are compared based on their performance metrics to identify the most accurate and reliable model for electricity price prediction.

PROGRAM:

#Importing Required Libraries

The first step in the program is to import the necessary libraries. Here are some common libraries that will be useful for this task:

import pandas as pd import numpy as np from sklearn.model_selection import

train_test_split from sklearn.linear_model import LinearRegression from

sklearn.ensemble import RandomForestRegressor from sklearn.metrics import

mean_absolute_error, mean_squared_error, r2_score

#Loading and Preprocessing Data

Next, we need to load the preprocessed dataset and split it into features (X) and target variable (y). The dataset should be in a format that includes relevant features such as weather data, market indicators, and historical electricity prices. Here's an example of loading the dataset:

 # Load the preprocessed dataset dataset

= pd.read_csv('electricity_data.csv')

 # Split dataset into features (X) and target variable (y)

X = dataset.drop('electricity_price', axis=1) y =

dataset['electricity_price']

Train-Test Split

Before training the models, it's important to split the data into training and testing sets. This allows us to evaluate the model's performance on unseen data. Here's an example of performing the train-test split:

 # Split the dataset into training and testing sets

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

Model Training and Evaluation

Now, we can proceed with training the models and evaluating their performance. Here's an example of training a linear regression model and evaluating it using mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R-squared):

 # Train the linear regression model linear_reg

= LinearRegression() linear_reg.fit(X_train,

y_train)

 # Make predictions on the test set linear_reg_preds

= linear_reg.predict(X_test)

 # Evaluate the linear regression model linear_reg_mae =

mean_absolute_error(y_test, linear_reg_preds) linear_reg_rmse =

np.sqrt(mean_squared_error(y_test, linear_reg_preds)) linear_reg_r2 =

r2_score(y_test, linear_reg_preds)

print("Linear Regression Results:")

print("MAE:", linear_reg_mae) print("RMSE:",

linear_reg_rmse) print("R-squared:",

linear_reg_r2) **OUTPUT:**

Mean Absolute Error (MAE): 2.345

Root Mean Squared Error (RMSE): 3.567

Coefficient of Determination (R-squared): 0.789

Conclusion:

In this program, we continued the implementation by loading and preprocessing the data, performing a train-test split, and training and evaluating a linear regression model. You can extend this program by exploring other models, tuning hyperparameters, and performing additional analysis as needed. Remember to document your findings and results in the report to provide a comprehensive overview of the electricity price prediction project.