Aim

This project aims to develop a machine-learning model to predict electricity prices. This model will be used to help businesses and individuals make informed decisions about their electricity consumption and production.

Program:

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn.linear\_model import

Linear Regression

from sklearn.model\_selection import

train\_test\_split

# Load the dataset

df =

pd.read\_csv('https://www.kaggle.com/data sets/chakradharmattapalli/electricity-price

-prediction/download')

# Preprocess the data

# Convert the date column to datetime

format

df['Date'] = pd.to\_datetime(df['Date'])

# Create a new column for the time of day df['Time'] = df['Date'].dt.hour

# Create a new column for the day of the week

df['Day of Week'] = df['Date'].dt.dayofweek

# Create a new column for the month df['Month'] = df['Date'].dt.month

# Drop the Date column

df.drop('Date', axis=1, inplace=True)

# Set the index to the Time column df.set\_index('Time', inplace=True)

# Fill in missing values

df.fillna(method='ffill', inplace=True)

# Convert the target variable to a NumPy array

target = df['Electricity Price'].to\_numpy()

# Convert the remaining features to a

NumPy array

features = df.drop('Electricity Price', axis=1).to\_numpy()

# Split the data into training and testing

sets

X\_train, X\_test, y\_train, y\_test =

train\_test\_split(features, target,

test\_size=0.25, random\_state=42)

# Train a linear regression model

model = LinearRegression()

model.fit(X\_train, y\_train)

# Evaluate the model on the test data

y\_pred = model.predict(X\_test)

mse = mean\_squared\_error(y\_test, y\_pred)

rmse = np.sqrt(mse)

r2 = r2\_score(y\_test, y\_pred)

# Print the evaluation metrics

print('MSE:', mse)

print('RMSE:', rmse)

print('R2:', r2)

# Make predictions on the test data

y\_test\_pred = model.predict(X\_test)

# Visualize the predictions

plt.scatter(X\_test[:, 1], y\_test, color='blue',

label='Actual')

plt.scatter(X\_test[:, 1], y\_test\_pred,

color='red', label='Predicted')

plt.legend()

plt.title('Electricity Price Prediction')

plt.xlabel('Time of Day (Hour)')

plt.ylabel('Electricity Price')

plt.show()

The evaluated metrics are:

* Mean squared error (MSE): A measure of the average squared difference between the actual and predicted values.
* Root mean squared error (RMSE): The square root of the MSE.
* R-squared (R²) : A measure of the proportion of the variance in the actual values that is explained by the predicted values.