**Phase 3: Development Part 1**

In this part you will begin building your project by loading and preprocessing the dataset. Electricity price prediction project by loading and preprocessing the data.

**Coding:**

Import pandas as pd

Import numpy as np

From sklearn.model\_selection import train\_test\_split

From sklearn.linear\_model import LinearRegression

From sklearn.metrics import mean\_squared\_error, r2\_score

Import matplotlib.pyplot as plt

Df = pd.read\_csv(‘Kaggle/input/electricity price-Electricity\_Price.csv’)

Df = df.fillna(0)

Df[‘Date’] = pd.to\_datetime(df[‘Date’])

Df.sort\_values(by=’Date’, inplace=True)

Df.reset\_index(drop=True, inplace=True)

Features = [‘Date’, ‘Consumption’, ‘Temperature’]

Df = df[features]

Train\_size = int(0.8 \* len(df))

Train\_data = df[:train\_size]

Test\_data = df[train\_size:]

X\_train = train\_data[[‘Temperature’]]

Y\_train = train\_data[‘Consumption’]

X\_test = test\_data[[‘Temperature’]]

Y\_test = test\_data[‘Consumption’]

Model = LinearRegression()

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

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

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

R2 = r2\_score(y\_test, y\_pred)

Print(“Mean Squared Error: “, mse)

Print(“R-squared: “, r2)

Plt.scatter(X\_test, y\_test, color=’b’)

Plt.plot(X\_test, y\_pred, color=’r’)

Plt.xlabel(‘Temperature’)

Plt.ylabel(‘Consumption’)

Plt.title(‘Electricity Price Prediction’)

Plt.show()