EX:No.4 221501039

25/01/25

Program to Develop a linear regression model for forecasting time series data.

Aim:

Write a program to implement time series data for import library, load data, Preprocessing and visualising.

Algorithm:

- Read the CSV file.
- Convert the `date` column to datetime and set it as the index.

Clean the Data

- Handle missing values using **forward fill** and **backward fill**.
- Drop any remaining NaN values.

Normalize the Data

- Apply **Min-Max Scaling** to normalize all numeric columns between 0 and 1.

Add Time-Based Features

- Extract additional features from the `date` column:
    - **Day**
    - **Month**
    - **Year**

Visualize the Data

- Plot a time series of the `price` column over time.

Execute the Program

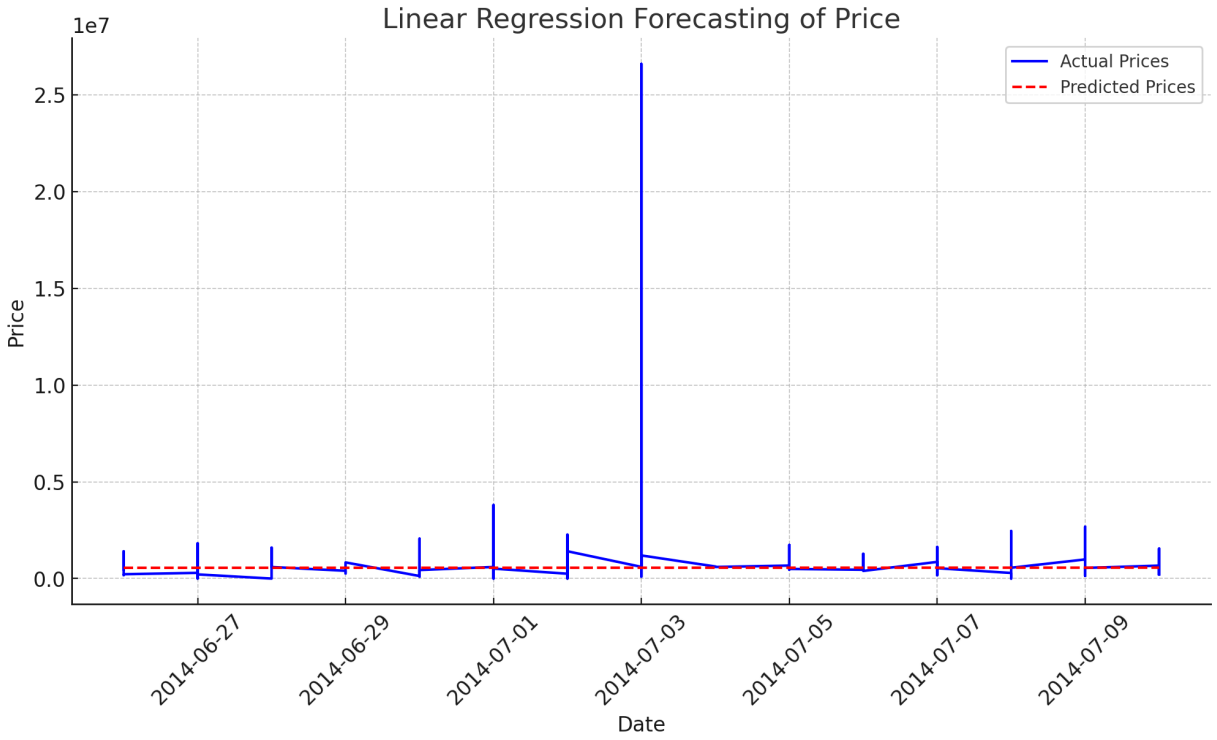
- Call each function in sequence to perform all the steps

Code:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import MinMaxScaler
```

```
def load_data(file_path):
    df = pd.read_csv(file_path, parse_dates=['date'])
    df.set_index('date', inplace=True)
    return df


def clean_data(df):
    df.fillna(method='ffill', inplace=True)  # Forward Fill
    df.fillna(method='bfill', inplace=True)  # Backward Fill
    df.dropna(inplace=True)  # Drop remaining NaNs
    return df


def normalize_data(df):
    scaler = MinMaxScaler()
    df[df.columns] = scaler.fit_transform(df[df.columns])
    return df

def add_time_features(df):
    df['day'] = df.index.day
    df['month'] = df.index.month
    df['year'] = df.index.year
    return df

def visualize_data(df, column='price'):
    plt.figure(figsize=(12, 6))
    sns.lineplot(x=df.index, y=df[column], label=column)
    plt.xlabel("Date")
    plt.ylabel("Normalized Price")
    plt.title(f"Time Series Plot of {column}")
    plt.grid(True)
    plt.legend()
    plt.show()

file_path = "/mnt/data/data.csv"
df = load_data(file_path)
df = clean_data(df)
df = normalize_data(df)
df = add_time_features(df)
visualize_data(df)
```

Output:


Linear Regression Forecasting of Price
1e7
2.5
2.0
1.5
1.0
0.5
0.0
Price
Actual Prices
Predicted Prices
2014-06-27
2014-06-29
2014-07-01
2014-07-03
2014-07-05
2014-07-07
2014-07-09
Date

Result:

Thus, the program using the time series data implementation has been done successfully.