EX:No.8 221501043

01/04/25

**Program to develop ARIMA model for forecasting time series data.**

**Aim:**

Write a program to create an ARIMA model for time series forecasting for same dataset

**Algorithm:**

1. **Load the Data**:
   * Read the CSV file containing the weather data.
   * Parse the date column as a datetime index.
2. **Clean the Data**:
   * Handle missing values by performing forward and backward filling.
   * Drop any remaining NaN values.
3. **Normalize the Data**:
   * Apply **Min-Max Scaling** to normalize each column's values between 0 and 1.
4. **Add Time-Based Features**:
   * Extract additional features from the datetime index: day, month and year
5. **Visualize the Data**:
   * Plot the time series for a specific column (e.g., temperature T) over time.
6. **Execute the Program**:
   * Sequentially call the functions to load, clean, normalize, add features, and visualize the data.

**Code:**

import pandas as pd

import matplotlib.pyplot as plt

from statsmodels.tsa.arima.model import ARIMA

import warnings

# Suppress warnings

warnings.filterwarnings("ignore")

# Load the dataset

df = pd.read\_csv("daily-minimum-temperatures-in-me.csv")

df.columns = ['Date', 'Temperature']

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

df.set\_index('Date', inplace=True)

# Remove invalid temperature entries and convert to float

df = df[df['Temperature'].str.match(r'^-?\d+(\.\d+)?$', na=False)]

df['Temperature'] = df['Temperature'].astype(float)

# Fit the ARIMA model

model = ARIMA(df['Temperature'], order=(5, 1, 0)) # You can tweak the (p,d,q) values

model\_fit = model.fit()

# Forecast the next 30 days

forecast = model\_fit.forecast(steps=30)

# Plot original data and forecast

plt.figure(figsize=(12, 6))

plt.plot(df.index, df['Temperature'], label='Original Data')

plt.plot(pd.date\_range(start=df.index[-1] + pd.Timedelta(days=1), periods=30, freq='D'),

forecast, label='Forecast (Next 30 Days)', color='red')

plt.title('ARIMA Forecast of Temperature')

plt.xlabel('Date')

plt.ylabel('Temperature')

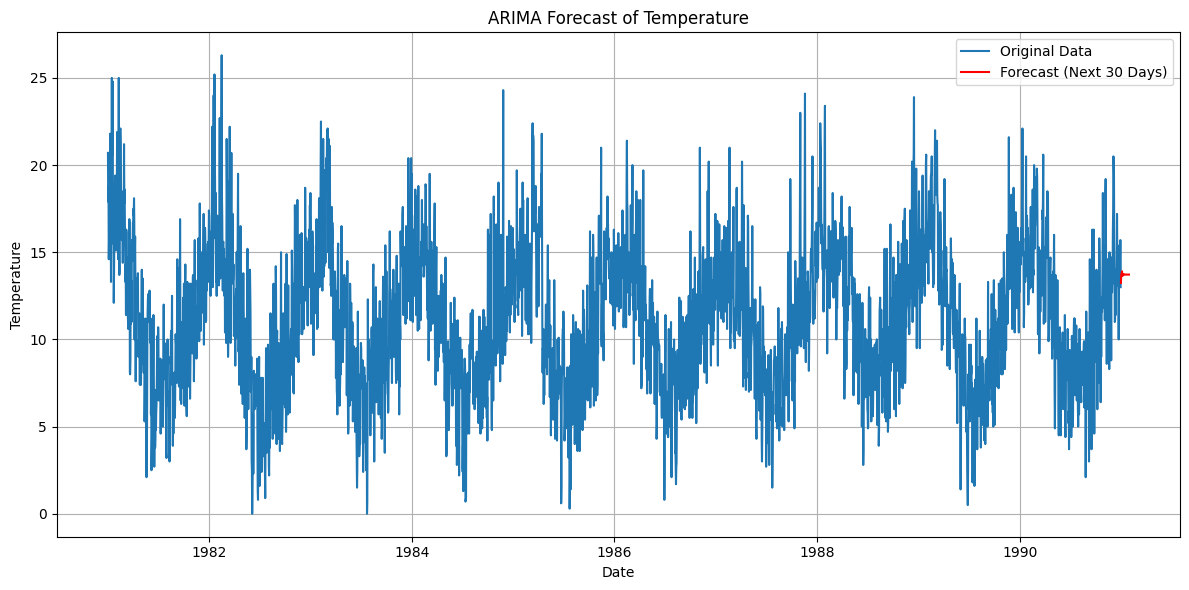
plt.legend()

plt.grid(True)

plt.tight\_layout()

plt.show()

**Output:**



**Result:**

Thus, the program to implement ARIMA in the time series data has been done successfully.