Time Series Forecasting Project-Building ARIMA Model in Python

Build a time series ARIMA model in Python to forecast the use of arrival rate density to support staffing decisions at call centres.

Project template outcomes

* Introduction to Time series
* Understand the basics of time series
* Importing the dataset and required libraries
* Exploratory Data Analysis (EDA)
* White Noise detection
* Random Walk detection
* Stationarity test
* Seasonality plot
* Holt Winter Exponential Smoothing model
* ARIMA model
* ACF plots
* Log-likelihood and AIC test
* ARIMAX model
* SARIMAX model

Description

**Business Objective**

A time series is simply a series of data points ordered in time. In a time-series, time is often the independent variable, and the goal is usually to make a forecast for the future.

Time series data can be helpful for many applications in day-to-day activities like:

* Tracking daily, hourly, or weekly weather data
* Monitoring changes in application performance
* Medical devices to visualize vitals in real-time

Auto-Regressive Integrated Moving Average (ARIMA) model is one of the more popular and widely used statistical methods for time-series forecasting. ARIMA is an acronym that stands for Auto-Regressive Integrated Moving Average. It is a class of statistical algorithms that captures the standard temporal dependencies unique to time-series data.

The model is used to understand past data or predict future data in a series. It’s used when a metric is recorded in regular intervals, from fractions of a second to daily, weekly or monthly periods.

ARIMAX (Auto-Regressive Integrated Moving Average Exogenous) is an extension of the ARIMA model, and similarly, SARIMAX (Seasonal Auto-Regressive Integrated Moving Average with Exogenous factors) is also an updated version of the ARIMA model. We will see how to implement these two models as well.

We have already covered the concepts of Time Series Project to Build an Autoregressive Model in Python and Build a Moving Average Time Series Forecasting Model in Python

In this project, we will be implementing the ARIMA model on the given dataset.

(ARIMA modelling is the third project in our list of time series projects, you can refer to the previous project through this link : [Build a Moving Average Time Series Forecasting Model in Python](https://www.projectpro.io/project-use-case/moving-average-time-series-project)[)](https://www.projectpro.io/project-use-case/time-series-projects-autoregressive-model-python)

**Data Description**

The dataset is “Call centers” data. This data is at month level wherein the calls are segregated at domain level as the call centre operates for various domains. There are also external regressors like no of channels and no of phone lines which essentially indicate the traffic prediction of the inhouse analyst and the resources available.

The total number of rows are 132 and number of columns are 8:

* Month, healthcare, telecom, banking, technology, insurance, no of phonelines and no of channels.

**Aim**

This project aims to build an ARIMA model on the given dataset.

**Tech stack**

* Language - Python
* Libraries - pandas, numpy, matplotlib, seaborn, statsmodels, scipy

**Approach**

1. Import the required libraries and read the dataset
2. Perform descriptive analysis
3. Exploratory Data Analysis (EDA) -
   * Data Visualization
4. Check for white noise
5. Check for Random Walk
6. Perform Stationarity tests
   * Augmented Dickey-Fuller test
   * KPSS test
7. Seasonal decomposition plot
8. Holt Winter Exponential Smoothing
   * Create and fit the model
   * Make predictions on the model
   * Plot the results
9. ARIMA model
   * Create models with varying lag values
   * Compare these models using log-likelihood and AIC values
   * Check with the LLR test
   * ACF Plots of residuals
10. ARIMAX model
    * Create a model
    * ACF plots of residuals
11. SARIMAX model
    * Create a model
    * ACF plots of residuals