---2---

### Time Series Challenge

### **Walmart Sales Forecast**

#### **Overview of the Task**

This data science challenge focuses on forecasting sales for Walmart using historical sales data. The dataset, available at [Walmart Sales Forecast Dataset](https://www.kaggle.com/datasets/aslanahmedov/walmart-sales-forecast), includes detailed records of sales transactions at various Walmart stores. Participants will need to preprocess the data, build forecasting models, and evaluate their performance to predict future sales. The final solution should also include a deployment strategy to make the forecasting model practical and accessible.

#### **Problem Statement**

Given the historical sales data from Walmart, your task is to build a robust time series forecasting model to predict future sales for each store. You need to handle the complexities of time series forecasting, such as seasonality, trends, and potential external factors influencing sales. Additionally, your solution should be deployable, allowing users to generate forecasts for future periods.

#### **Objective**

1. **Data Exploration and Preprocessing**: Clean and preprocess the dataset to prepare it for analysis. This includes handling missing values, encoding categorical variables, and addressing any inconsistencies or anomalies in the data.
2. **Feature Engineering**: Develop additional features that can enhance the forecasting models. Consider temporal features (e.g., day of the week, month), store-specific attributes, and external factors (e.g., holidays, promotions).
3. **Model Building and Evaluation**: Create and evaluate various forecasting models. Explore traditional time series models (e.g., ARIMA, SARIMA) as well as more advanced methods (e.g., Prophet, LSTM). Assess model performance using appropriate evaluation metrics.
4. **Forecasting**: Generate sales forecasts for a specified future period and validate the accuracy of these forecasts.
5. **Deployment**: Develop a deployment strategy for your forecasting model. This may involve creating a web application or an API that allows users to input data and receive sales forecasts.

#### **Data Description**

The dataset includes the following features:

* **Date**: The date of the sales transaction, formatted as YYYY-MM-DD.
* **Store**: The unique identifier for each Walmart store.
* **Dept**: The unique identifier for each department within the store.
* **Weekly\_Sales**: The total sales for the week, given in USD.
* **IsHoliday**: A boolean indicator of whether the week includes a holiday.

The data spans multiple years and includes weekly sales figures for each store and department, with indicators for holidays that may affect sales.

#### **Evaluation Criteria**

|  |  |
| --- | --- |
| **Data Collection, Connectivity, and Cleaning** | **30%** |
| **Logic and Results** | **40%** |
| **Presentation** | **5%** |
| **Code Quality** | **5%** |
| **Deployment / Running App** | **20%** |

We look forward to your innovative solutions and insights. Good luck with your forecasting models and deployment strategy!