Project Documentation: Future Sales Prediction

Phase 2:

**2.1 Short Explanation (Future Sales Prediction)**

The central problem of this project is to develop an innovative predictive model for a retail company, leveraging historical sales data to accurately forecast future sales. The primary goal is to create a predictive tool to optimize inventory management and make data-driven decisions that enhance the company's operational efficiency and profitability.

**2.2 Data Set Details:**

**Data Set Link**: https://www.kaggle.com/datasets/chakradharmattapalli/future-sales-prediction

**2.3 Details about Columns**

The dataset will include the following essential columns for analysis:

* **TV**: The TV of the sales Count.
* **Radio**: A Radio Count for each product.
* **Newspaper**: A Newspaper Count for each product.
* **Sales**: The quantity of each product Sales.

**2.4 Libraries and Tools**

We will use the following libraries and tools for data analysis, model building, and evaluation:

* **Python**: The programming language for data analysis and modeling.
* **Pandas**: For data manipulation and preprocessing.
* **NumPy**: For numerical operations and array handling.
* **Matplotlib**: For data visualization and plotting.
* **Prophet**: A time series forecasting library developed by Facebook.
* **Keras**: A deep learning library for implementing LSTM networks.
* **Scikit-Learn**: For additional machine learning tools.

**2.5 Data Preprocessing, Training, and Testing**

**Data Preprocessing**: We will preprocess the data, including data cleaning to remove errors and inconsistencies. We'll also handle missing values and engineer relevant features, such as time-based features and lagged variables.

**Model Training and Testing**:

1. **Prophet Model**: Implement the Prophet model for time series forecasting. Train the model using the preprocessed data and evaluate its performance using time series forecasting metrics like Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).
2. **LSTM (Long Short-Term Memory) Networks**: Preprocess the data for LSTM training, including scaling and reshaping. Design and train an LSTM model to predict future sales. Evaluate its performance using appropriate evaluation metrics, including MAE and RMSE.

**2.6 Metrics for Accuracy Check**

The following metrics will be used to evaluate the accuracy of the forecasting models:

* **Mean Absolute Error (MAE)**: This metric measures the average absolute difference between predicted and actual values.
* **Root Mean Squared Error (RMSE)**: RMSE provides a measure of the square root of the average squared differences between predicted and actual values, which penalizes larger errors more heavily.

**2.7 Conclusion**

This project focuses on innovative sales forecasting for a retail company. We will explore advanced time series forecasting techniques like Prophet and LSTM networks to provide the company with improved accuracy in predicting future sales. The primary objectives are to optimize inventory management, make informed business decisions, and enhance operational efficiency and profitability.

**2.8 Program Link:** <https://github.com/ShariqueShaqub/ADS_Phase1>