# FUTURE SALES PREDICTION

# PHASE 5:FINAL SUBMISSION

## INTRODUCTION

Future sales prediction, often referred to as sales forecasting, is a crucial aspect of business strategy. It involves the use of statistical models and predictive analytics to estimate a company's future sales performance. By examining historical sales data, market trends, and other relevant factors, businesses can make informed decisions regarding inventory management, resource allocation, and overall growth strategies. Sales predictions enable companies to adapt to changing market conditions, optimize marketing efforts, and ensure sustainable success in a competitive business landscape.

## WORKS DONE IN PREVIOUS PHASES

## DEFINITION PHASE

In this phase we defined the problem that develop a predictive model that uses historical sales data to forecast future sales for a retail company. The objective is to create a tool that enables the company to optimize inventory management and make informed business decisions based on data driven sales predictions. These involves data preprocessing, feature engineering, model selection, training, and evaluation.

### DESIGN THINKING

### Data Source:

Utilizing a dataset containing historical sales data, including features like date, product ID, store ID, and sales quantity.

**Data Preprocessing:**

Cleaning and preprocessing the data, handle missing values, and convert categorical features into numerical representations.

**Feature Engineering:**

Creating an additional features that could enhance the predictive power of the model, such as time-based features (e.g., day of the week, month).

**Model Selection:**

Choosing suitable time series forecasting algorithms (e.g., ARIMA, Exponential Smoothing) for predicting future sales.

**Model Training:**

Training the selected model using the preprocessed data.

**Evaluation:**

Validating the model's performance using appropriate time series forecasting metrics (e.g., Mean Absolute Error, Root Mean Squared Error)

## INNOVATION PHASE

In the innovation phase of our future sales prediction project, you can explore advanced techniques and methods to improve the accuracy of your sales forecasting model like Prophet Forecasting model and LSTM

**Prophet Forecasting Model:**

- Facebook's Prophet is a robust time series forecasting tool that is designed for forecasting with daily observations that display patterns on different time scales. It handles holidays, special events, and missing data gracefully. You can integrate Prophet into your project and compare its performance with the initial models you selected in Phase 1.

- Prophet is a forecasting tool developed by Facebook that is designed to handle time series data with daily observations that have seasonality and holidays. It uses a decomposable time series model with components like trend, seasonality, and holiday effects. It's especially

useful when your data exhibits non-linear patterns and has irregular gaps or missing data points.

**Long Short-Term Memory (LSTM) Networks:**

- LSTM is a type of recurrent neural network (RNN) that has shown excellent performance in time series forecasting tasks. You can implement LSTM networks using deep learning libraries like TensorFlow or PyTorch. LSTM networks are capable of capturing complex patterns and long-term dependencies in time series data.

- LSTM networks are a type of recurrent neural network (RNN) architecture that excels at capturing sequential dependencies in time series data. They have memory cells that can retain information over long sequences, making them well-suited for tasks like time series forecasting. Implementing LSTM involves setting up a neural network with LSTM layers, feeding it sequences of historical data, and training it to predict future sales based on past patterns

## DEVELOPMENT PHASE 1

In this phase we loaded the dataset which is provided for us and pre-processed the data by using python library packages and necessary methods to implement it.

These phase can be executed using three parts

- Loading and Pre-processing data
- Training and Testing data
- Model testing and Displaying Output

For loading the data we can use the dataset link

https://www.kaggle.com/datasets/chakradharmattapalli/future-salesprediction

### IMPORTING LIBRARIES

We importing the necessary Python libraries, such as

- Pandas for data manipulation

- NumPy for analysis
- Matplotlib for visualization.

**PREPROCESSING DATA**

We can use the following functions for the data pre-processing. They are

- **Readcsv()**
- **Head()**
- **Tail()**
- **Shape**
- **Columns**
- **Column to list**
- **Isnull**

## DEVELOPMENT PHASE 2

- In this phase we are going to test the model which we are pre-processed by using some of the models and going to evolve those models
- This can be executed by using
    1. Feature engineering
    2. Model testing
    3. Evaluation

**FEATURE ENGINEERING**

We can use the following commands to execute the feature of the sales prediction dataset.

- **Sklearn**
- **KNearestNeighbor**

**MODEL TESTING**

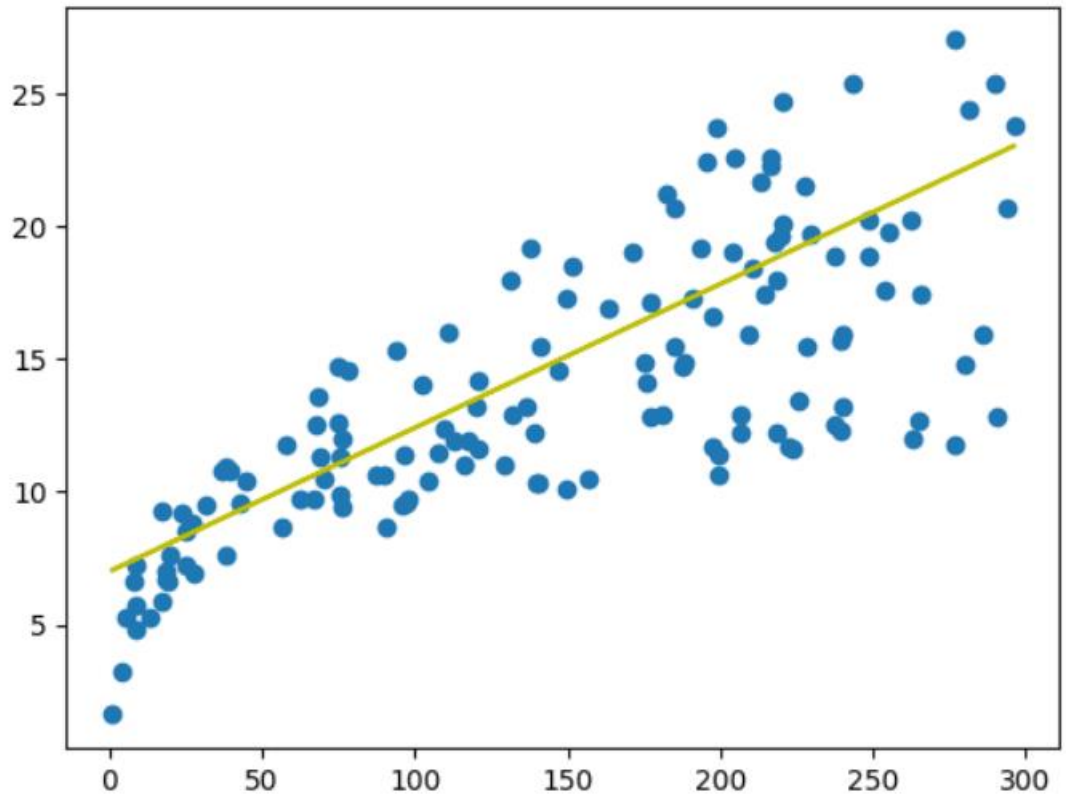
- **Support Vector Regression**
- **Root Mean Squared**
- **Mean Value Error**

These are the commands used for the testing of the model in sales prediction dataset.

## EVALUATION

The following libraries are used for the visualization of the data output by using graphical data structure.

- **Matplotlib**
- **Seaborn**




The above mentioned graph is the output of the values which provided in the dataset

**NOTE: Graph may vary if the value of the data is changed**