**FUTURESALESPREDICTION**

**Problem Definition and Design Thinking**

**Problem definition:**

To develop a predictive model that uses historical sales data that improves their demand planning process.

It is used for sales forecasting for a retail company that is predicting the future sales of a product.

Sales forecasting is the process of estimating a company’s sales revenue for a specific time period.

The tool is created for the process of inventory management that is the process of tracking stock levels and the movement of goods and it involves analyzing the benefits.

This project involves data preprocessing,feature engineering,model selection,training,and evaluation.

**Design Thinking**:

**Data source:**

The product that contain the dataset like date,product ID,store ID and the sales quantity, eg:laptop.

The product contain the above dataset and it should be processed.

**Data preprocessing:**

The next step is to preprocess the above data,that involves cleaning, that is identify the error in the data such as missing values,transformation,and convert the product features into numerical representations.

The goal of the data preprocessing is to improve the quality of the data.

**Feature Engineering:**

It is the process for enhance the predictive models and highlighting the features such as time-based features.

For eg:extracting the day of the week,month of the year,or hour of the day.

**Model selection:**

The time-series forecasting algorithm to be choosed.

It is a powerful method for predicting future sales and values in time-series data.

Time series algorithms:

1. ARIMA –autoregressive integrated moving average it uses weighted previous value while the moving average part weights the previously assumed errors of the time series.

2.Exponential smoothing.

**Model Training:**

After the addition of features,selection of model ,the model is to be trained.

It is a method to measure the accuracy of the model. Train the selected model using the preprocessed data.

**Evaluation:**

The three error metrics that are commonly used for evaluating and reporting the performance of the model. They are

1.Mean Squared Error(MSE).

2.Root Mean Squared Error(RMSE).

3.Mean Absolute Error(MAE).