**PROJECT TITLE:**

PRODUCTION DEMAND PRDICTION USING MACHINE LEARNING

**PROBLEM DEFINITION:**

* A product company plans to offer discounts on its product during the upcoming holiday season. The company wants to find the price at which its product can be a better deal compared to its competitors. For this task, the company provided a dataset of past changes in sales based on price changes. You need to train a model that can predict the demand for the product in the market with different price segments.

**DESIGN THINKING:**

Data Collection:

The Dataset contain the 5 columns. The columns are ID, STORE ID, TOTAL PRICE, BASE PRICE, UNITS SOLD

Data Preprocessing:

Data Cleaning: Clean the data to remove duplicates, missing values, and outliers.

Feature Engineering:

The feature engineering means, how to select the features for the build the model like chi-square, f-test ,etc..

Model Selection:

The product demand prediction project build by the ML models like supervised and unsupervised algorithms. The algorithms are LOGISTIC REGRESSION, LINEAR REGRESSION, XGBoost ,SVM, RANDOM FOREST , and, etc…

Model Training:

The dataset is split into training and testing set for the model evaluation . the training set size is 80% and testing size 20%.

Model Evaluation:

The all models measure the accuracy score, f1- score, pression and recall and etc.. the high accurate model select the testing data.

**PRE-PROCESSING:**

**STEPS:**

1. **DATA CLEANING**
2. **HANDLE MISSING VALUES**
3. **CATEGORICAL TO NUMERICAL REPRESENTATIONS.**

**DATA CLEANING:**

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset. When combining multiple data sources, there are many opportunities for data to be duplicated or mislabeled.

**HANDLE MISSING VALUES:**

1. Deleting Rows with missing values
2. Impute missing values for continuous variable
3. Impute missing values for categorical variable
4. Other Imputation Methods
5. Using Algorithms that support missing values
6. Prediction of missing values
7. Imputation using Deep Learning Library — Datawig

**CATEGORICAL TO NUMERICAL REPRESENTATIONS:**

1. cat.codes Attribute
2. replace
3. Label Encoder

**ALGORITHM:**

1. **ARIMA**
2. **Prophet**

**ARIMA:**

ARIMA stands for Autoregressive Integrated Moving Average Model. It belongs to a class of models that explains a given time series based on its own past values -i.e.- its own lags and the lagged forecast errors. The equation can be used to forecast future values. Any ‘non-seasonal’ time series that exhibits patterns and is not a random white noise can be modeled with ARIMA models.

So, ARIMA, short for AutoRegressive Integrated Moving Average, is a forecasting algorithm based on the idea that the information in the past values of the time series can alone be used to predict the future values.

ARIMA Models are specified by three order parameters: (p, d, q),

where,

p is the order of the AR term

q is the order of the MA term

d is the number of differencing required to make the time series stationary

AR(p) Autoregression – a regression model that utilizes the dependent relationship between a current observation and observations over a previous period. An auto regressive (AR(p)) component refers to the use of past values in the regression equation for the time series.

I(d) Integration – uses differencing of observations (subtracting an observation from observation at the previous time step) in order to make the time series stationary. Differencing involves the subtraction of the current values of a series with its previous values d number of times.

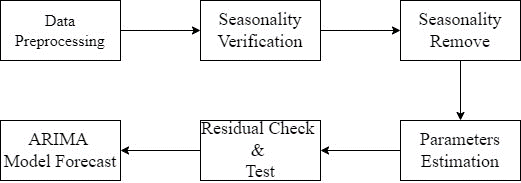
MA(q) Moving Average – a model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations. A moving average component depicts the error of the model as a combination of previous error terms. The order q represents the number of terms to be included in the model.

Types of ARIMA Model

ARIMA : Non-seasonal Autoregressive Integrated Moving Averages

SARIMA : Seasonal ARIMA

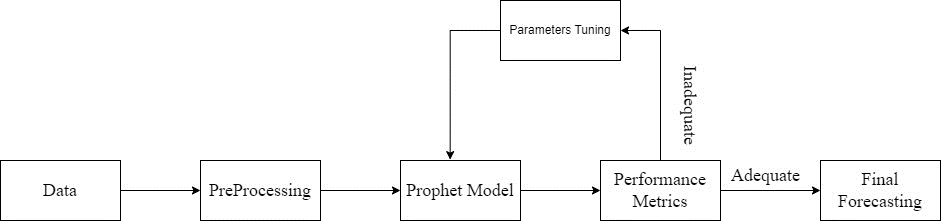
SARIMAX : Seasonal ARIMA with exogenous variables



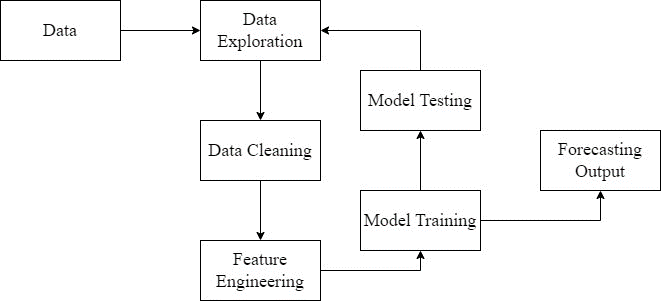
**Prophet:**

Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data. Prophet is robust to missing data and shifts in the trend, and typically handles outliers well.

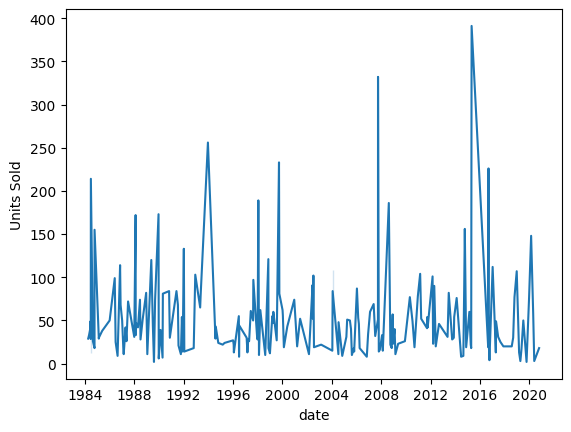
Y(t) = g(t) + s(t) + h(t) + E(t)



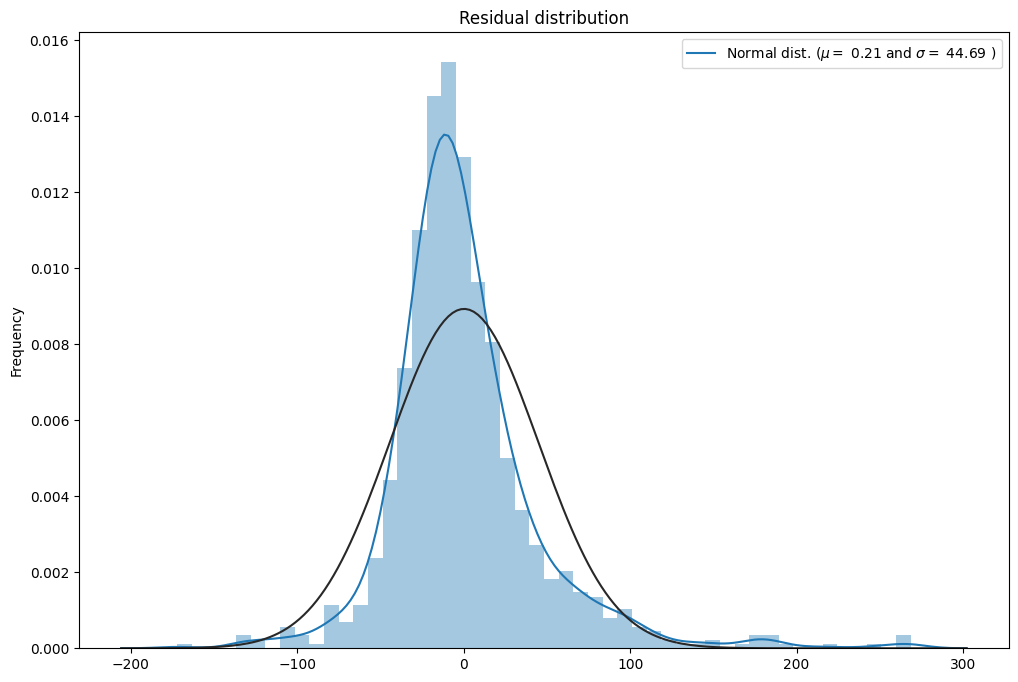
**PROJECT WORKFLOW:**

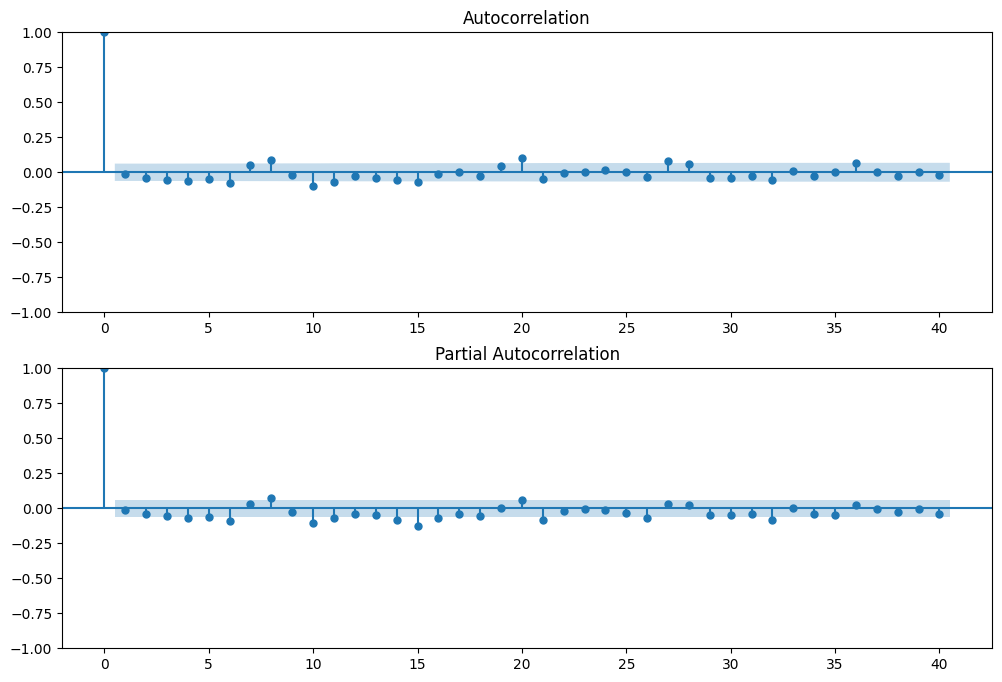


The product sales in the year range of 1984 to 2020

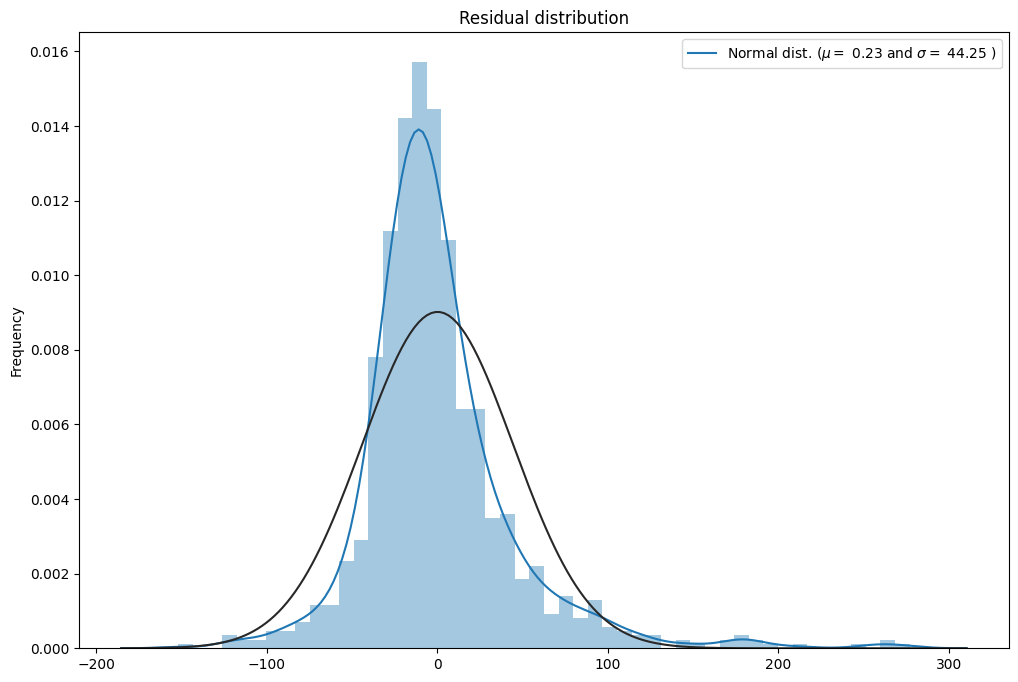


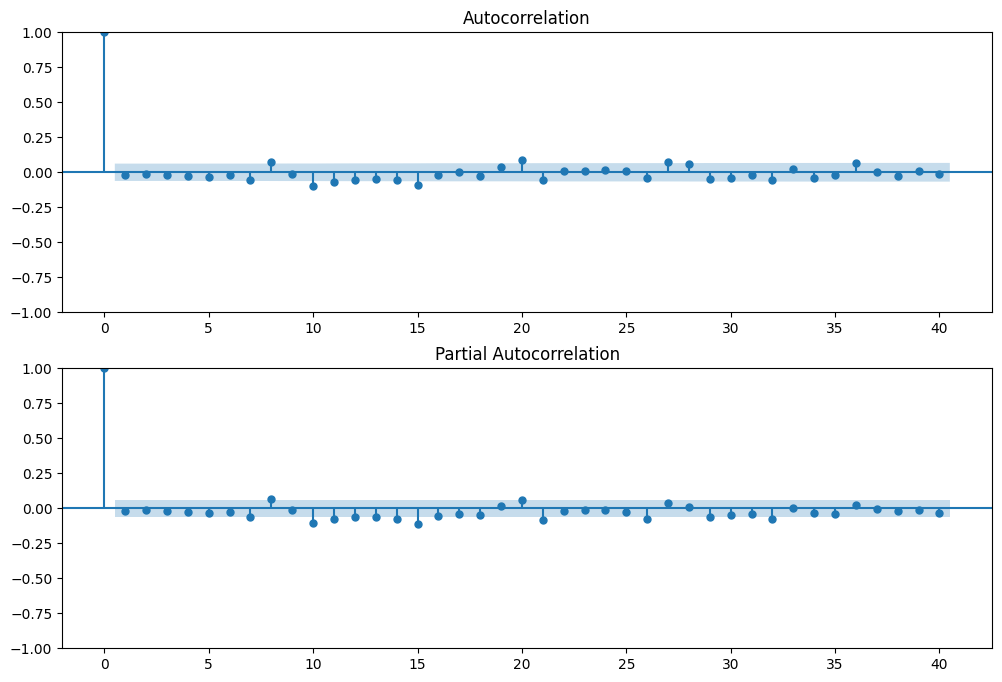
The ARIMA model prediction:



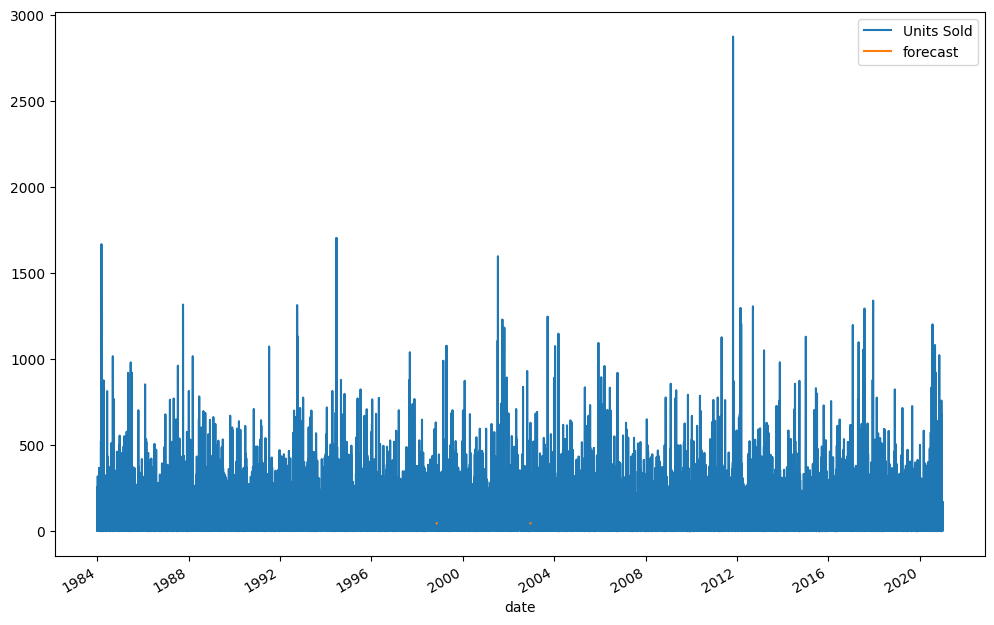


**The SARIMA model:**

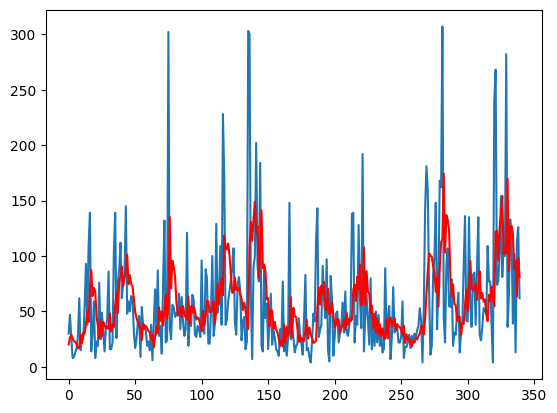




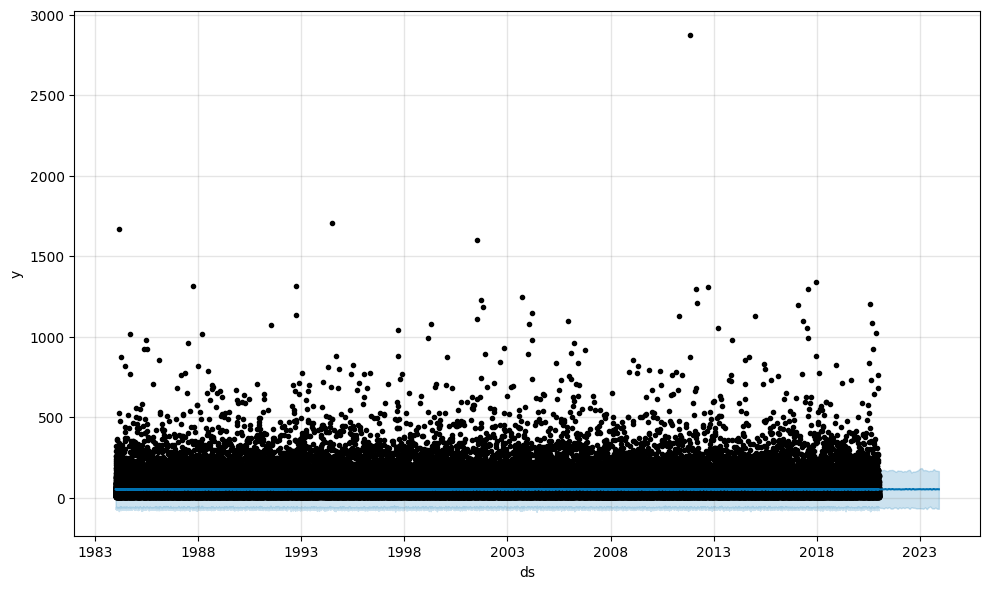
**The forecasting is:**



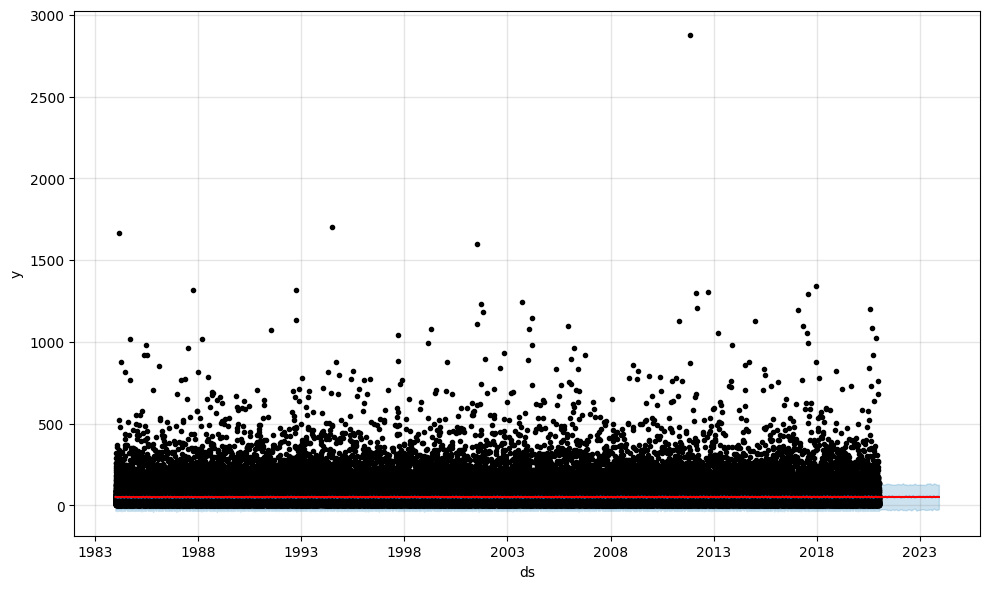
**Model predicted is ,**



**The prophet model is**



**The final prophot:**



**CONCLUSION:**

The product demand prediction using machine learning models, specifically ARIMA, SARIMA, and Prophet, has provided valuable insights for businesses .It is noteworthy to mention that in this analysis, ARIMA and SARIMA models produced similar forecasting results, which indicates that both models can be effectively employed for time series forecasting.

The ARIMA and SARIMA models are well-suited for capturing the temporal patterns and trends in historical demand data, making them reliable choices when dealing with stationary time series data. These models utilize the auto-regressive, integrated, and moving average components to provide accurate predictions, especially when seasonality and other cyclic patterns are present**.**

Prophet, on the other hand, is a versatile forecasting tool, capable of handling irregularly spaced data and capturing both seasonality and holiday effects. Its ability to automatically detect changepoints in the time series data makes it a valuable addition to the forecasting toolkit.

In summary, the integration of ARIMA, SARIMA, and Prophet models in product demand prediction empowers businesses to make informed decisions, optimize inventory management, and enhance customer satisfaction by ensuring the right products are available when needed. These models, when appropriately chosen and fine-tuned, can contribute to more accurate demand forecasting, leading to improved operational efficiency.

**REFERENCES:**

* Forecasting new product demand using machine learning To cite this article: P S Smirnov and V A Sudakov 2021 J. Phys.: Conf. Ser. 1925 012033 <https://www.researchgate.net/publication/352614910_Forecasting_new_product_demand_using_machine_learning>
* Demand Prediction using Machine Learning Methods and Stacked Generalization Resul Tugay and S¸ule Gund ¨ uz¨ Og¸ ¨ ud¨ uc¨ u¨ Department of Computer Engineering, Istanbul Technical University, Istanbul, Turkey <https://www.scitepress.org/papers/2017/64316/64316.pdf>
* AN EFFECTIVE AND REAL TIME PRODUCT DEMAND FORECASTING MODEL USING ML Niveditha M Shaini S, Shwetha T, Sneha H K, Dr. Natesh M Student, Department Of Computer Science And Engineering, Vidyavardhaka College OfEngineering, Mysore, Karnataka, India. Associate Professor, Department Of Computer Science And Engineering, Vidyavardhaka College Of Engineering, Mysore, Karnataka, India. <https://www.irjmets.com/uploadedfiles/paper/issue_4_april_2022/21003/final/fin_irjmets1650120394.pdf>
* Machine Learning in Predicting Demand for Fast-Moving Consumer Goods: An Exploratory Research Elcio Tarallo ,Getúlio K. Akabane, Camilo I. Shimabukuro, Jose Mello Douglas Amancio <https://www.sciencedirect.com/science/article/pii/S240589631931153X>