

# **PHASE-2: INNOVATION**

## **PRODUCT SALES ANALYSIS**

### **PREDICTIVE ANALYTICS:**

Implementing the machine learning models to predict future sales trends based on historical data. This can help businesses proactively adjust their strategies. Here are some machine learning models and techniques such as linear regression, time series forecasting, random forest regression, gradient boosting, recurrent neural networks, convolution neural network, ensemble methods, evaluation metrics, hyperparameter tuning, feature engineering.

**Here are some basic algorithms for predicting future sales and trends:**

### **SERIES FORECASTING:**

Time series data is common in sales and trend prediction. Algorithms like ARIMA (Auto Regressive Integrated Moving Average), Exponential Smoothing, and Prophet are designed specifically for forecasting time series data.

### **LINEAR REGRESSION:**

Linear regression models can be used when you have continuous numerical variables and want to predict future sales based on historical data and other features like advertising spend, seasonality.

### **DECISION TREE AND RANDOM FORESTS:**

Decision trees and random forests are useful for capturing non-linear relationships between various factors and sales. They can handle both numerical and categorical data and are interpretable.

### **GRADIENT DESCENT:**

Algorithms like XG Boost, Light GBM, and Cat Boost are powerful ensemble methods that often perform well in sales forecasting tasks. They can handle complex relationships and are robust to outliers.

### **NEURAL NETWORKS:**

Deep learning models, such as feedforward neural networks or recurrent neural networks (RNNs), can be used for more complex and large-scale sales prediction tasks, especially when you have a lot of data.

## **SUPPORT VECTOR MACHINES (SVM):**

SVMs can be used for regression tasks, where the goal is to find a hyperplane that best fits the data. They are particularly useful when dealing with high-dimensional data.

## **K-NEAREST NEIGHBOUR:**

KNN is a simple yet effective algorithm that can be used for sales prediction by finding the "neighbours" of a data point based on similarity and averaging their values.

## **RECURRENT NEURAL NETWORKS(RNNS):**

RNNs are suitable for time series data with sequential dependencies. They can capture trends and patterns in sequential sales data.

## **LSTM (LONG SHORT-TERM MEMORY) NETWORKS:**

LSTM networks are a specialized type of RNN that can handle longer sequences and are well-suited for capturing temporal dependencies in sales data.

## **PROPHET:**

Prophet is an open-source forecasting tool developed by Facebook that is designed for forecasting with daily observations that display patterns on different time scales. It is user-friendly and can capture seasonality and holiday effects.

## **CONVOLUTIONAL NEURAL NETWORK:**

Convolutional Neural Networks (CNNs) can be applied to predictive analytics tasks, especially when dealing with structured grid data such as time series or spatial data.

## **ENSEMBLE METHODS:**

Ensemble methods in predictive analytics refer to techniques that combine multiple machine learning models to create a stronger, more accurate model. These methods aim to improve the overall predictive performance and generalizability of the model by leveraging the strengths of different algorithms.

## **HYPERPARAMETER TUNING:**

Hyperparameter tuning is a crucial step in building machine learning models for predictive analytics. It involves finding the optimal set of hyperparameters for a machine learning algorithm to achieve the best performance on a validation dataset.

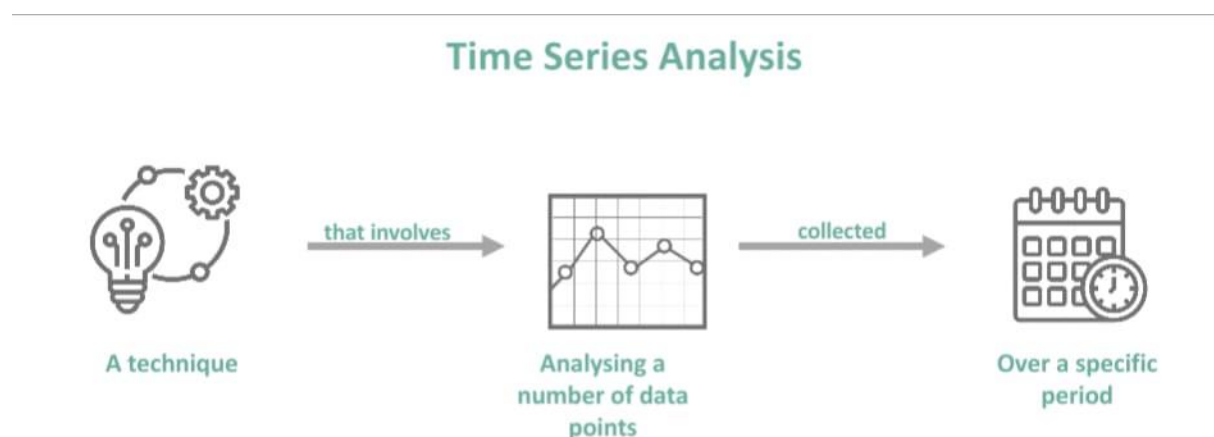
## FEATURE ENGINEERING:

Feature engineering is a crucial step in predictive analytics that involves selecting, transforming, or creating relevant features (variables) from the raw data to improve the performance of machine learning models. Here are some common techniques for feature.

One popularly used machine learning algorithm for predicting future sales and trends is the Time Series analysis.

## TIME SERIES ANALYSIS

Time series analysis is a statistical technique used to analyze and interpret data points collected or recorded at specific time intervals. It is widely used in various fields such as finance, economics, weather forecasting, and signal processing. Time series data consists of observations on a variable or several variables over time.



## KEY COMPONENTS OF TIME SERIES ANALYSIS INCLUDE:

### Trend:

The long-term movement in a time series, which can be upward, downward, or stable.

### Seasonality:

Patterns that occur regularly at specific intervals, often influenced by factors like seasons, months, days, or hours.

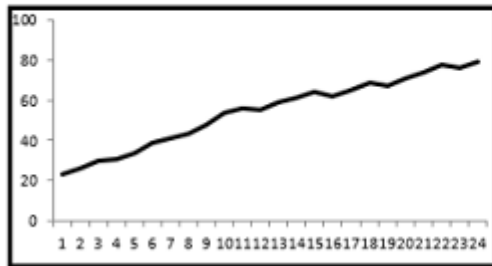
### Cyclical Patterns:

Repeating up and down movements in the data that are not of fixed frequency like seasonality.

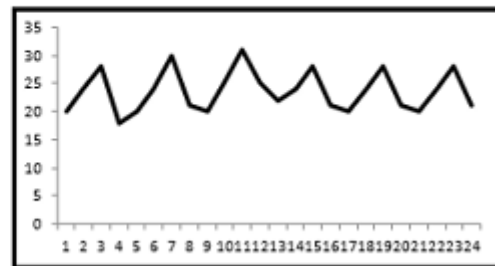
**Noise:**

Random variation or irregular movements in the data that cannot be attributed to the above components.

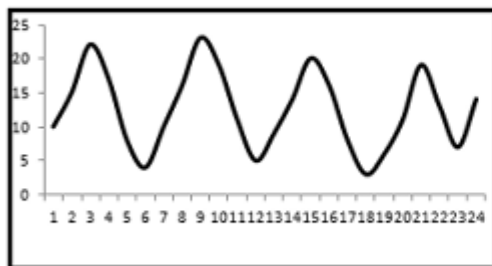
## Components of Time Series



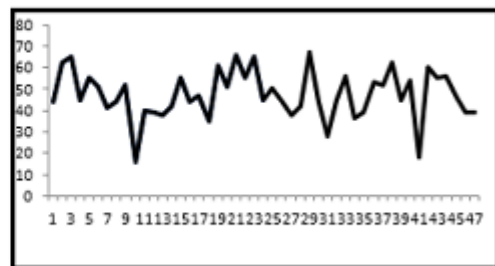
(a) Trend



(b) Seasonality



(c) Cyclicality



(d) Irregular

Common techniques used in time series analysis

### 1.Descriptive Statistics:

Calculating mean, median, and standard deviation to understand the basic properties of the data.

### 2. Smoothing:

Techniques like moving averages to identify trends by averaging out noise.

### 3. Decomposition:

Separating the time series into its components (trend, seasonality, and noise) for detailed analysis.

### 4. Autocorrelation:

Examining the correlation between a time series and its lagged values.

## **5. Forecasting:**

Using historical data to make predictions about future values of the time series, employing methods like ARIMA (Auto Regressive Integrated Moving Average) or machine learning models.

## **FEATURES OF TIME SERIES ANALYSIS:**

- Time ordered data
- Trend
- Seasonality
- Noise
- Auto correlation
- Stationarity
- Forecasting
- Data decomposition
- Visualization

## **APPLICATIONS OF TIME SERIES ANALYSIS:**

- Finance
- Economics
- Business and marketing
- Weather forecasting
- Retail
- Environmental science
- Health care
- Manufacturing
- Traffic and transportation

## **CONCLUSION:**

In conclusion, time series analysis is a powerful statistical technique that plays a vital role in understanding, interpreting, and predicting patterns in sequential data points over time. Its applications are widespread, ranging from finance and economics to healthcare, weather forecasting, and beyond.