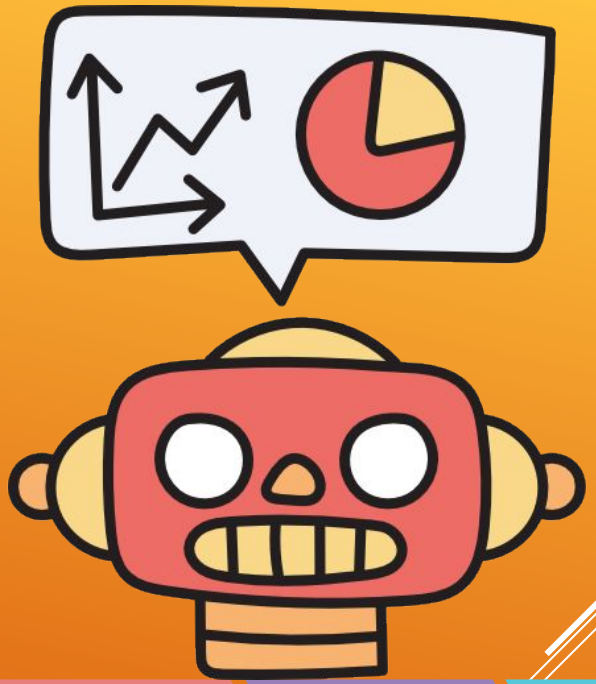
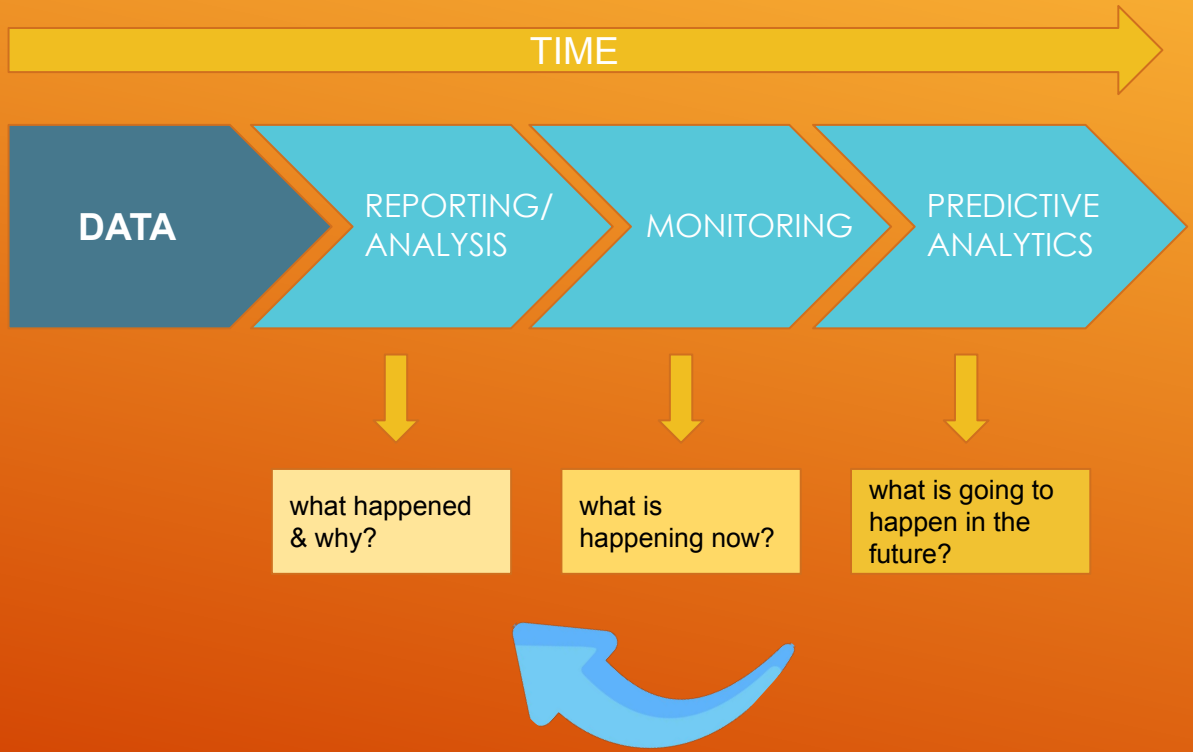
Predictive Analytics



Predictive Analytics

Predictive analytics is the branch of the advanced **analytics** which is used to make predictions about unknown future events. **Predictive analytics** uses many techniques from data mining, statistics, modeling, machine learning, and artificial intelligence to analyze current data to make predictions about future.

Predictive Analytics Process



Business process and features on Predictive Modelling

Business process on Predicting modelling

- Creating the model
- Testing the model
- Validating the model
- Evaluating the model

Features in Predicting modelling

- Data analysis and manipulation
- Visualization
- Statistics
- Hypothesis testing

How the model work

In predictive modeling, data is collected for the relevant predictors, a statistical model is formulated, predictions are made and the model is validated (or revised) as additional data becomes available. The model may employ a simple linear equation or a complex neural network, mapped out by sophisticated software.

How the model work

Here you will learn what a predictive model is, and how, by actively guiding marketing campaigns, it constitutes a key form of business intelligence. we'll take a look inside to see how a model works-

- Predictors Rank Your Customers to Guide Your Marketing
- Combined Predictors Means Smarter Rankings
- The Computer Makes Your Model from Your Customer Data
- A Simple Curve Shows How Well Your Model Works
- Conclusions

Why Predictive Modelling

Nearly every business in competitive markets will eventually need to do predictive modeling to remain ahead of the curve. Predicting Modeling (also known as Predictive Analytics) is the process of automatically detecting patterns in data, then using those patterns to foretell some event. Predictive models are commonly built to predict:

- Customer Relationship Management
- the chance a prospect will respond to an ad
- Mail recipients likely to buy
- when a customer is likely to churn
- if a person is likely to get sick
- Portfolio or Product Prediction
- Risk Management & Pricing

Applications of Predictive Modelling

- Analytical customer relationship management (CRM)
- Health Care
- Collection Analytics
- Cross-sell
- Fraud detection
- Risk management

Industry Applications

Predictive modelling are used in insurance, banking, marketing, financial services, telecommunications, retail, travel, healthcare, oil & gas and other industries.

Predictive Models in Retail industry

Campaign Response Model – this model predicts the likelihood that a customer responds to a specific campaign by purchasing a products solicited in the campaign. The model also predicts the amount of the purchase given response.

- Regression models
- Customer Segmentation
- Cross-Sell and Upsell
- New Product Recommendation
- Customer Retention/Loyalty/Churn
- Inventory Management

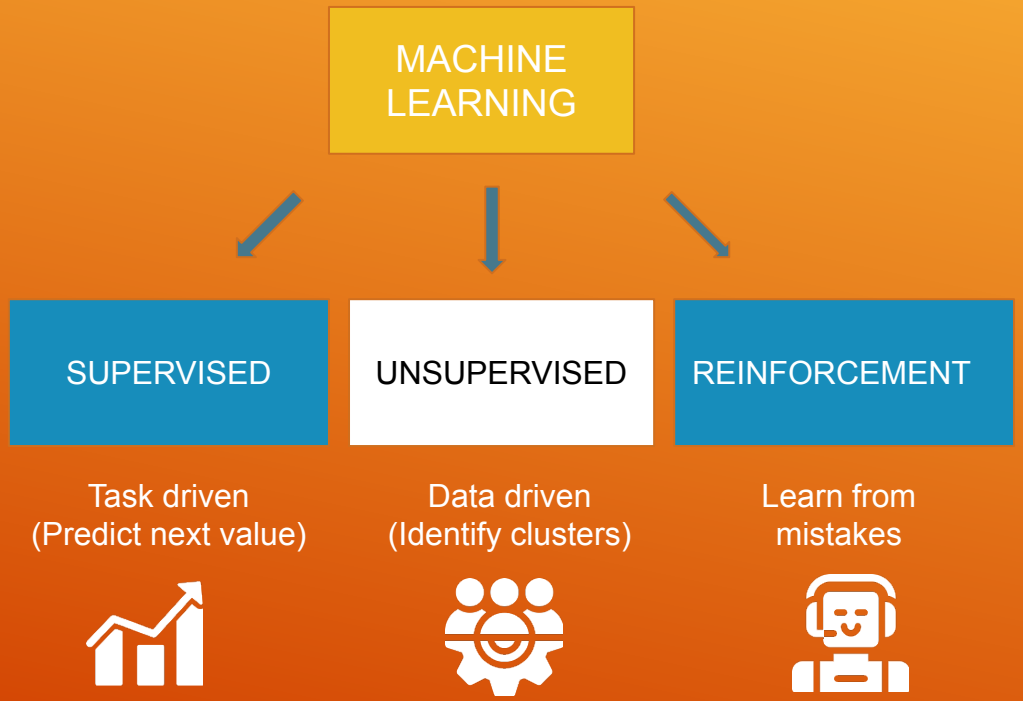
What is Machine Learning

- Machine learning is an application of artificial intelligence that involves algorithms and data that automatically analyse and make decision by itself without human intervention.
- It describes how computer perform tasks on their own by previous experiences.
- Therefore we can say in machine language artificial intelligence is generated on the basis of experience.

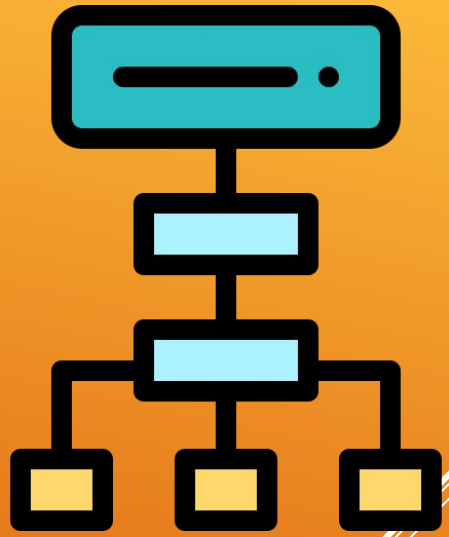
Types of Machine Learning

There are three types of machine learning

- Supervised learning
- Unsupervised learning
- Reinforcement learning

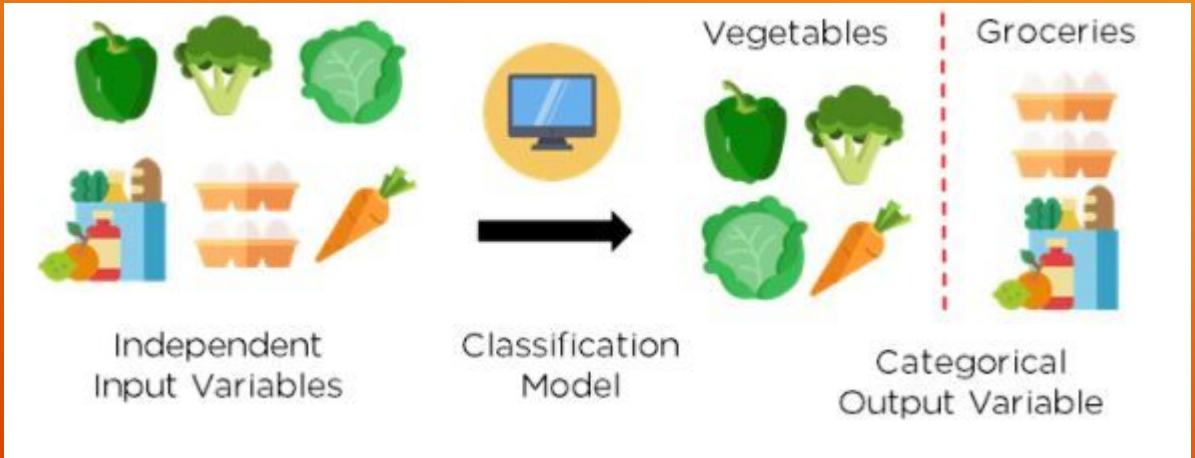


Classification Algorithms



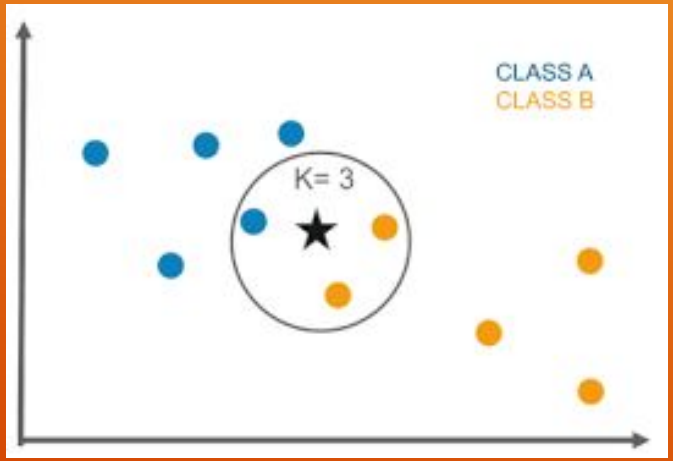
Classification

Classification algorithms used in machine learning utilize input training data for the purpose of predicting the likelihood or probability that the data that follows will fall into one of the predetermined categories. One of the most common applications of classification is for filtering emails into “spam” or “non-spam”, as used by today’s top email service providers.



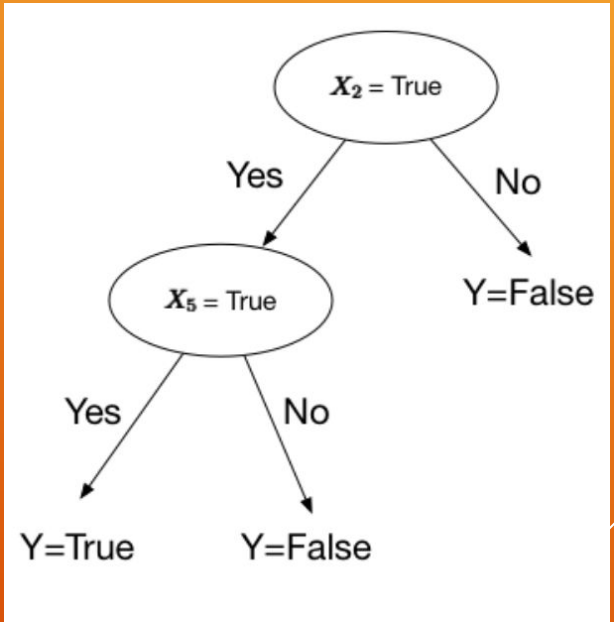
K-Nearest Neighbor(KNN)

It is a supervised machine learning algorithm. The algorithm can be used to solve both classification and regression problem statements. The number of nearest neighbours to a new unknown variable that has to be predicted or classified is denoted by the symbol 'K'.



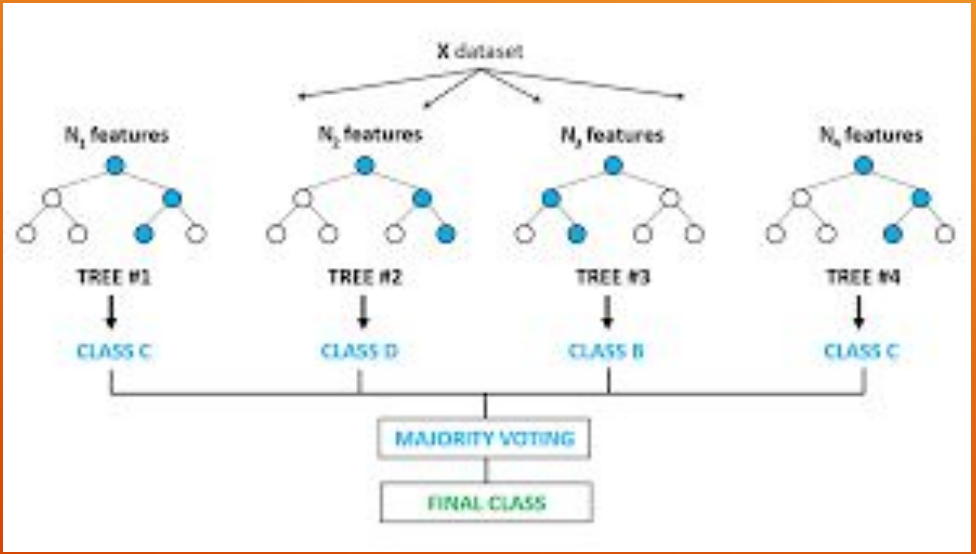
Decision Tree

A decision tree is a very specific type of probability tree that enables you to make a decision about some kind of process. For example, you might want to choose between manufacturing item A or item B, or investing in choice 1, choice 2, or choice 3.

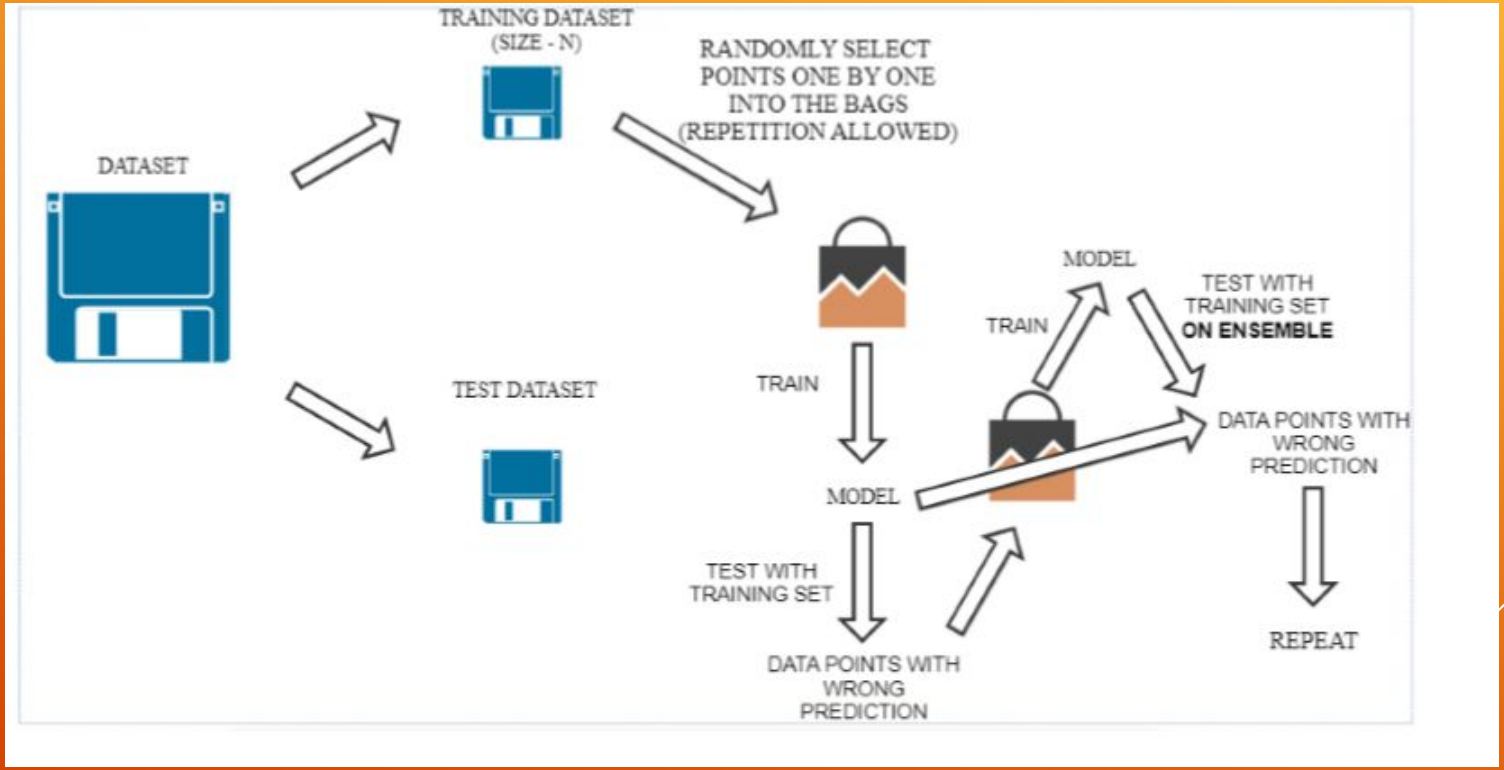


Random Forest

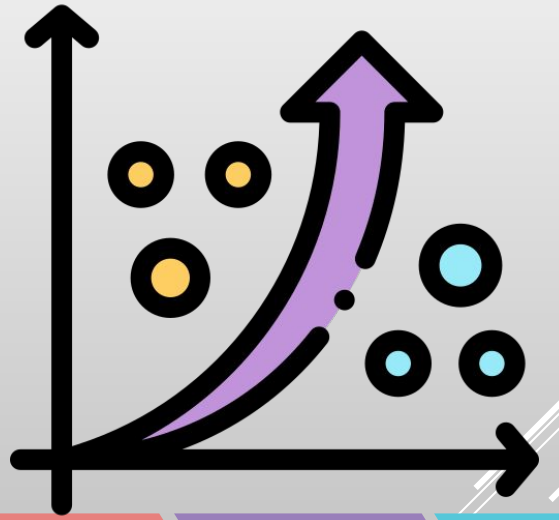
Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees.



Boosting

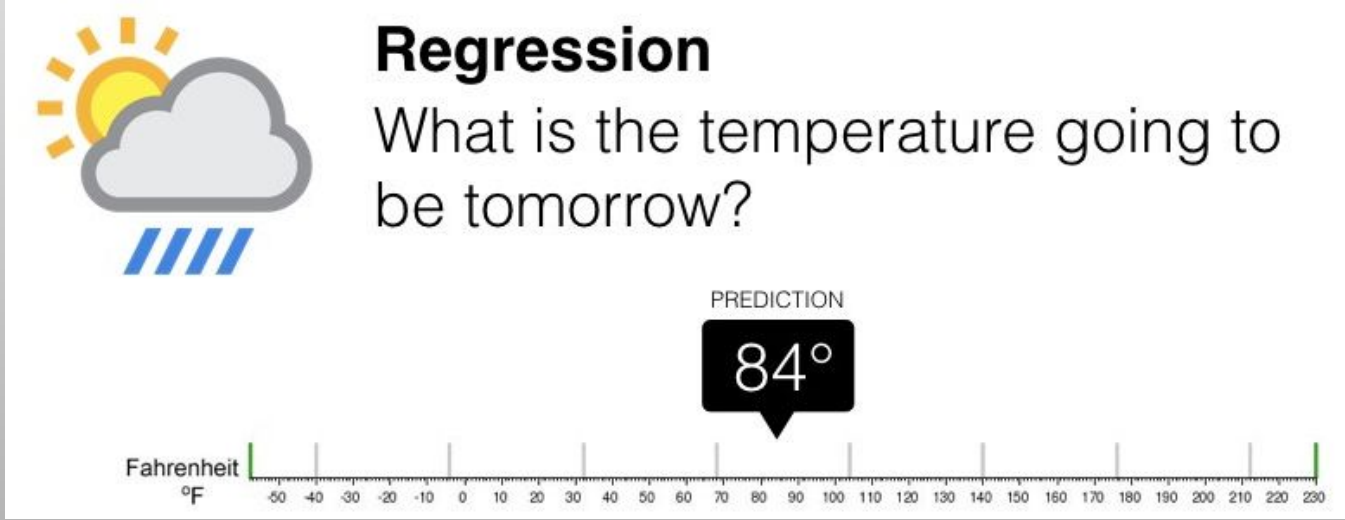

TRAINING DATASET
(SIZE - N)
RANDOMLY SELECT
POINTS ONE BY ONE
INTO THE BAGS
(REPETITION ALLOWED)
DATASET
MODEL
TEST WITH
TRAINING SET
ON ENSEMBLE
TRAIN
TEST DATASET
TRAIN
DATA POINTS WITH
WRONG
PREDICTION
MODEL
TEST WITH
TRAINING SET
REPEAT
DATA POINTS WITH
WRONG
PREDICTION

Regression Algorithms



Regression

Regression analysis is a statistical method to model the relationship between a dependent (target) and independent (predictor) variables with one or more independent variables. More specifically, Regression analysis helps us to understand how the value of the dependent variable is changing corresponding to an independent variable when other independent variables are held fixed. It predicts continuous/real values such as **temperature, age, salary, price,** etc



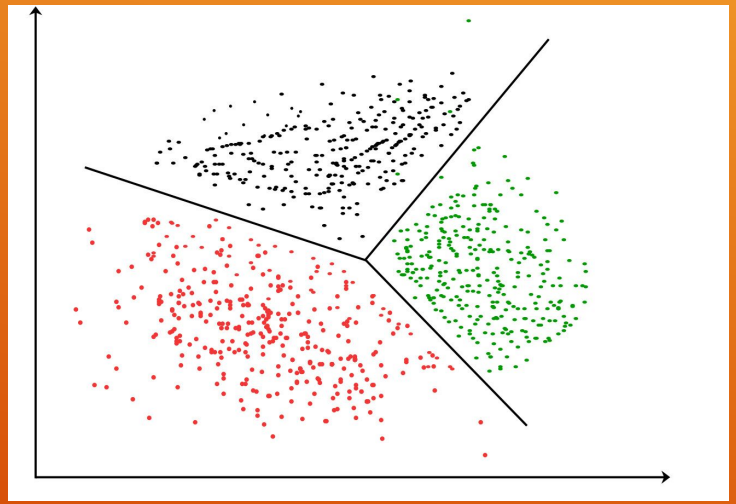
Clustering Algorithms

Clustering

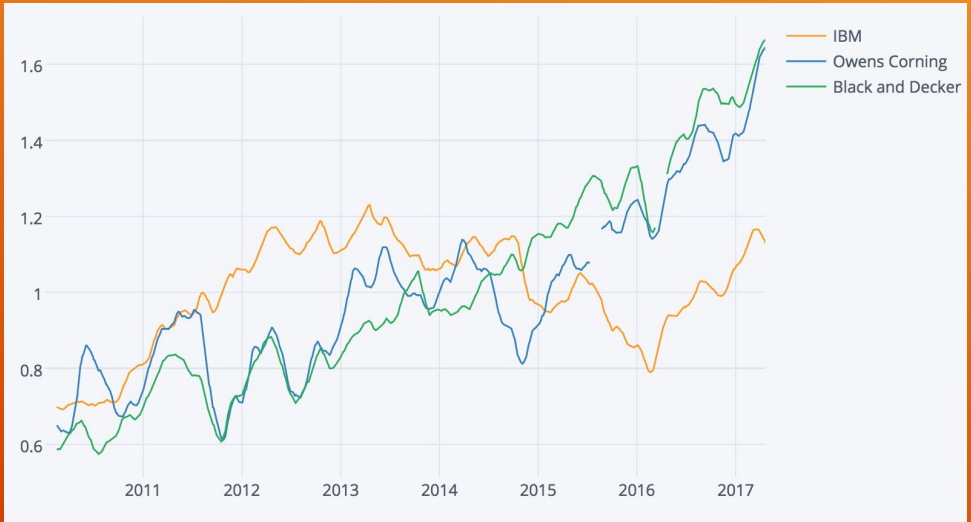
Clustering or cluster analysis is a machine learning technique, which groups the unlabelled dataset. It can be defined as "A way of grouping the data points into different clusters, consisting of similar data points. The objects with the possible similarities remain in a group that has less or no similarities with another group."

The clustering technique can be widely used in various tasks. Some most common uses of this technique are:

- *Market Segmentation*
- *Statistical data analysis*
- *Social network analysis*
- *Image segmentation*
- *Anomaly detection, etc.*



Time Series Analysis

Time Series Analysis

A **time series** is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time. Thus it is a sequence of discrete-time data. Examples of time series are heights of ocean tides, counts of sunspots, and the daily closing value of the Dow Jones Industrial Average.