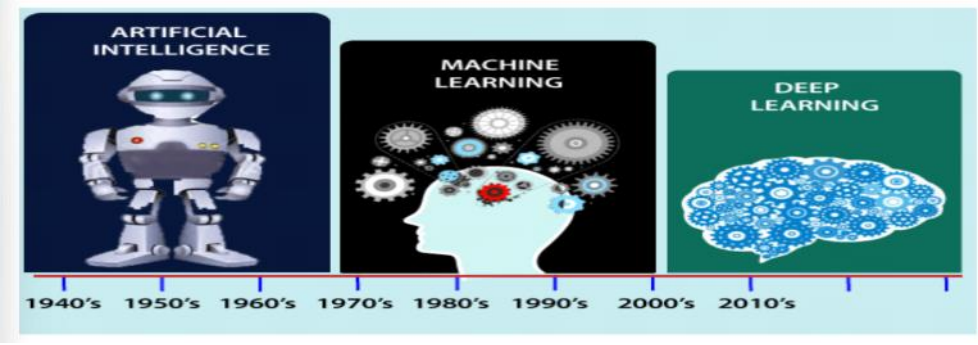
# UNIT I : MACHINE LEARNING

## Syllabus For Unit 1

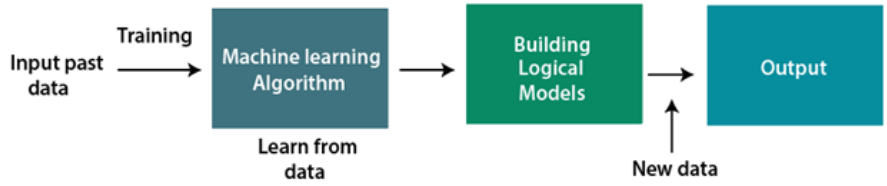
**Introduction to Machine Learning, Types of Machines Learning, Linear Regression, Classification and Logistic Regression, Decision Tree and Random Forest, Naïve Bayes and Support Vector Machine. Applications of machine learning.**




## ❖ Introduction to Machine Learning

→ **Machine Learning & It's Working.**

- "Machine learning is a subfield of Artificial Intelligence (AI) that involves the development of algorithms and statistical models, which allow computers to learn from data, improve performance from experiences, and predict things without being explicitly programmed."
- Machine learning uses various algorithms for building mathematical models and making predictions using historical data or information.
- The accuracy of predicted output depends upon the amount of data, as the huge amount of data helps to build a better model which predicts the output more accurately.
- Currently, it is being used for various tasks such as image recognition, speech recognition, email filtering, Facebook auto-tagging, recommender system, and many more.
- "Machine learning enables a machine to automatically learn from data, improve performance from experiences, and predict things without being explicitly programmed."
- Suppose we have a complex problem, where we need to perform some predictions, so instead of writing a code for it, we just need to feed the data to generic algorithms, and with the help of these algorithms, machine builds the logic as per the data and predict the output.
-




→ **Features Of Machine Learning**

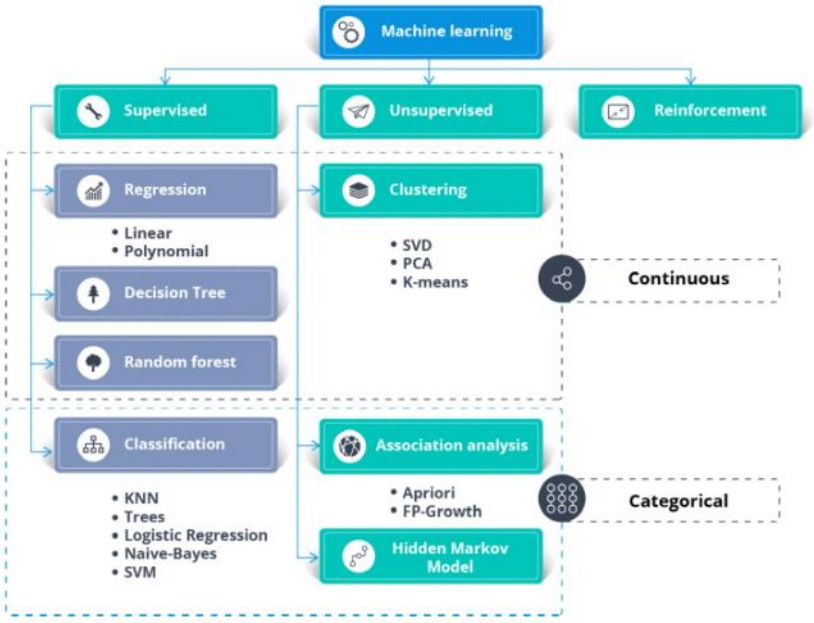
- Predictive modeling: Data is used by machine learning algorithms to create models that forecast future events. Example : These models can be used to determine the risk of a loan default.
- Automation: Machine learning algorithms automate the process of finding patterns in data, requiring less human involvement and enabling more precise and effective analysis.
- Scalability: Machine learning techniques are well suited for processing big data because they are made to handle massive amounts of data.
- Generalization: Algorithms for machine learning are capable of discovering broad patterns in data that can be used to analyze fresh, unexplored data, they are useful for forecasting future events.
- Adaptiveness: As new data becomes available, machine learning algorithms are built to learn and adapt continuously. As a result, they can enhance their performance over time, becoming more precise and efficient as more data is made available to them.

→ **Need For Machine Learning**

- The reason behind the need for machine learning is that it is capable of doing tasks that are too complex for a person to implement directly.
- As a human, we have some limitations as we cannot access the huge amount of data manually, so for this, we need some computer systems and here comes the machine learning to make things easy for us.
- We can train machine learning algorithms by providing them the huge amount of data and let them explore the data, construct the models, and predict the required output automatically.
- The performance of the machine learning algorithm depends on the amount of data, and it can be determined by the cost function. With the help of machine learning, we can save both time and money.
- Currently, machine learning is used in self-driving cars, cyber fraud detection, face recognition, and friend suggestion by Facebook, etc.
- Various top companies such as Netflix and Amazon have built machine learning models that are using a vast amount of data to analyze the user interest and recommend product accordingly.

## ➢ Types Of Machine Learning

Examples of Machine Learning Algorithms:

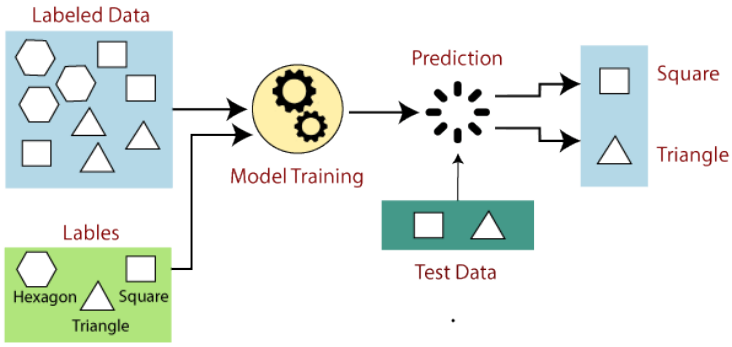



### ➢ Supervised Learning

- Supervised learning is the types of machine learning in which machines are trained using well "labelled" training data, and on basis of that data, machines predict the output. The labelled data means some input data is already tagged with the correct output.
- In supervised learning, the training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly.
- The aim of a supervised learning algorithm is to find a mapping function to map the input variable(x) with the output variable(y). Y = f(X)
- In the real-world, supervised learning can be used for Risk Assessment, Image classification, Fraud Detection, spam filtering, etc.

### → Working Of Supervised Learning

- In supervised learning, models are trained using labelled dataset, where the model learns about each type of data. Once the training process is completed, the model is tested on the basis of test data (a subset of the training set), and then it predicts the output.
- Working of Supervised Learning example and diagram:




- Suppose we have a dataset of different types of shapes: square, rectangle, triangle, and polygon. Now the first step is that we need to train the model for each shape.
  - If the given shape has four sides, and all the sides are equal, then it will be labelled as a Square.
  - If the given shape has three sides, then it will be labelled as a triangle.
  - If the given shape has six equal sides then it will be labelled as hexagon.
- Now, after training, we test our model using the test set, and the task of the model is to identify the shape.
- The machine is already trained on all types of shapes, and when it finds a new shape, it classifies the shape on the bases of a number of sides, and predicts the output.

→ **Types Of Supervised Machine Learning Algorithms**

(a) Regression :

- Regression algorithms are used if there is a relationship between the input variable and the output variable.
- It is used for the prediction of continuous variables, such as Weather forecasting, Market Trends, etc.
- Below are some popular Regression algorithms which come under supervised learning.
  - Linear Regression
  - Regression Trees
  - Non-Linear Regression
  - Bayesian Linear Regression
  - Polynomial Regression

(b) Classification :

- Classification algorithms are used when the output variable is categorical, which means there are two classes such as Yes-No, Male-Female, True-false, etc.
- Below are some popular Classification algorithms which come under supervised learning.
  - Random Forest
  - Decision Trees
  - Logistic Regression
  - Support Vector Machines

→ **Advantages Of Supervised Learning**

- With the help of supervised learning, the model can predict the output on the basis of prior experiences.
- In supervised learning, we can have an exact idea about the classes of objects.
- Supervised learning model helps us to solve various real-world problems such as fraud detection, spam filtering, etc.

→ **Disadvantages of supervised learning:**

- Supervised learning models are not suitable for handling the complex tasks.
- Supervised learning cannot predict the correct output if the test data is different from the training dataset.
- Training required lots of computation times.
- In supervised learning, we need enough knowledge about the classes of object.

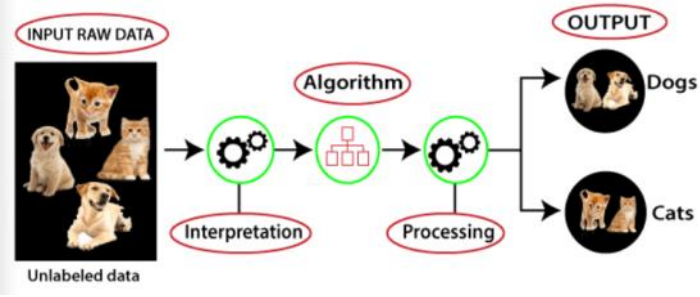
➢ **Unsupervised Learning**

- Unsupervised learning is a type of machine learning in which models are trained using unlabeled dataset and are allowed to act on that data without any supervision.
- Models itself find the hidden patterns and insights from the given data.

- The goal of unsupervised learning is to find the underlying structure of dataset, group that data according to similarities, and represent that dataset in a compressed format.
- In real-world, we do not always have input data with the corresponding output so to solve such cases, we need unsupervised learning.
- Unsupervised Learning Algorithms : K-means Clustering, KNN (k-nearest neighbors), Hierarchal Clustering, Anomaly Detection, Neural Networks, Principle Component Analysis, Independent Component Analysis, Apriori Algorithm, Singular Value Decomposition

→ **Working Of Unsupervised Learning**

- Unsupervised learning works on unlabeled and uncategorized data which make unsupervised learning more important.
-




- Here, we have taken an unlabeled input data, which means it is not categorized and corresponding outputs are also not given.
- Now, this unlabeled input data is fed to the machine learning model in order to train it.
- Firstly, it will interpret the raw data to find the hidden patterns from the data and then will apply suitable algorithms such as k-means clustering, Decision tree, etc.
- Once it applies the suitable algorithm, the algorithm divides the data objects into groups according to the similarities and difference between the objects.

→ **Types Of Unsupervised Learning**

(a) Clustering :

- Clustering is a method of grouping the objects into clusters such that objects with most similarities remains into a group and has less or no similarities with the objects of another group.
- Cluster analysis finds the commonalities between the data objects and categorizes them as per the presence and absence of those commonalities.

(b) Association

- An association rule is an unsupervised learning method which is used for finding the relationships between variables in the large database. It determines the set of items that occurs together in the dataset.
- Association rule makes marketing strategy more effective. Such as people who buy X item (suppose a bread) are also tend to purchase Y (Butter/Jam) item. A typical example of Association rule is Market Basket Analysis.

→ **Advantages of Unsupervised Learning**

- Unsupervised learning is used for more complex tasks as compared to supervised learning because, in unsupervised learning, we don't have labeled input data.
- Unsupervised learning is preferable as it is easy to get unlabeled data in comparison to labeled data.

→ **Disadvantages of Unsupervised Learning**

- Unsupervised learning is intrinsically more difficult than supervised learning as it does not have corresponding output.
- The result of the unsupervised learning algorithm might be less accurate as input data is not labeled, and algorithms do not know the exact output in advance.

➢ **Reinforcement Learning**

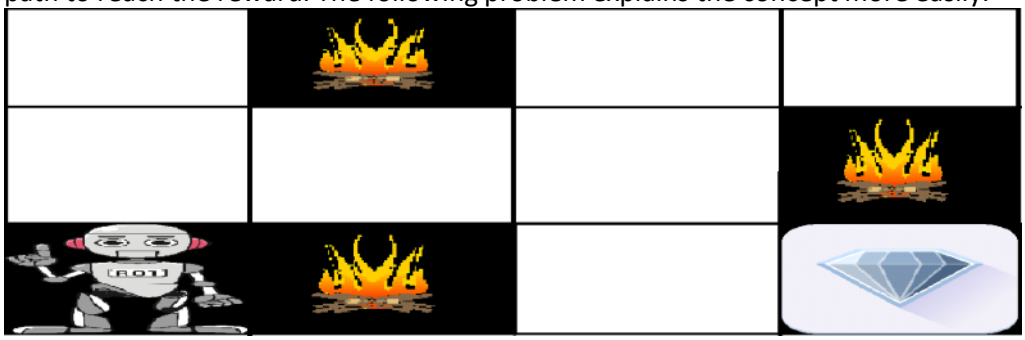
- Reinforcement learning is a feedback-based learning method, in which a learning agent gets a reward for each right action and gets a penalty for each wrong action.
- The agent learns automatically with these feedbacks without any labelled data and improves its performance.
- In reinforcement learning, the agent interacts with the environment and explores it. The goal of an agent is to get the most reward points, and hence, it improves its performance.
- It is a good technique to use for automated systems that have to make a lot of small decisions without human

guidance.

– 


→ **Working Of Reinforcement Learning**

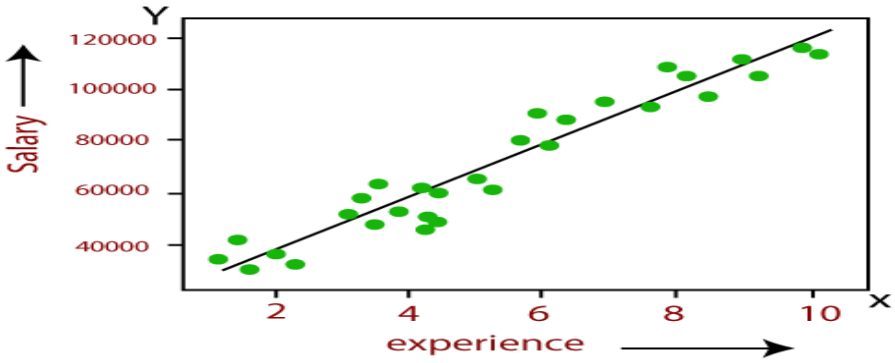
– To understand working process of the RL, we need to consider two main things :
  - Environment : It can be anything such as a room, maze, football ground, etc.
  - Agent : An intelligent agent such as AI robot.
– Example :
– We have an agent and a reward, with many hurdles in between. The agent is supposed to find the best possible path to reach the reward. The following problem explains the concept more easily.

– 

– The above image shows the robot, diamond, and fire. The goal of the robot is to get the reward that is the diamond and avoid the hurdles that are fired.
– The robot learns by trying all the possible paths and then choosing the path which gives him the reward with the least hurdles.
– Each right step will give the robot a reward (+1) and each wrong step will subtract the reward (-1) of the robot. The total reward will be calculated when it reaches the final reward that is the diamond.

## ❖ Linear Regression

– Linear regression is a statistical regression method which is used for predictive analysis.
– It shows the relationship between the continuous variables
– Linear regression algorithm shows the linear relationship between the independent variable (X-axis) and the dependent variable (Y-axis), hence called linear regression.
– Since, it shows the linear relationship between the continuous variables, it finds how the value of the dependent variable is changing according to the value of the independent variable.
– If there is only one input variable (x), then such linear regression is called simple linear regression and if there is more than one input variable, then such linear regression is called multiple linear regression.
– Linear regression makes predictions for continuous/real or numeric variables such as sales, salary, age, product price, etc.
– The relationship between variables in the linear regression model can be explained using the below image. Here we are predicting the salary of an employee on the basis of the year of experience.



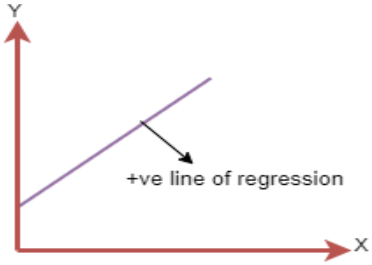

- Mathematical equation for Linear Regression :
  - **Y = a0+ a1X + ε**

  Y= Dependent Variable (Target Variable)
  X= Independent Variable (predictor Variable)
  a0= intercept of the line (Gives an additional degree of freedom)
  a1 = Linear regression coefficient (scale factor to each input value).
  ε = random error
- The values for x and y variables are training datasets for Linear Regression model representation.

→ **Linear Regression Line**

- A linear line showing the relationship between the dependent and independent variables is called a regression line.
- A regression line can show two types of relationship:
  - **Positive Linear Relationship:**

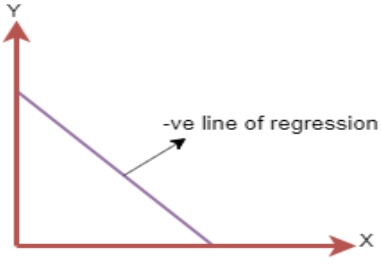
    If the dependent variable increases on the Y-axis and independent variable increases on X-axis, then such a relationship is termed as a Positive linear relationship.




The line equation will be: $Y= a_0+a_1x$

  - **Negative Linear Relationship:**

    If the dependent variable decreases on the Y-axis and independent variable increases on the X-axis, then such a relationship is called a negative linear relationship.



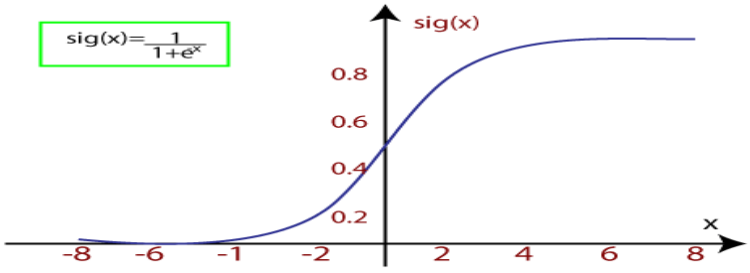

The line of equation will be: $Y= -a_0+a_1x$

## ❖ Logistic Regression

- Logistic regression is another supervised learning algorithm which is used to solve the classification problems. In classification problems, we have dependent variables in a binary or discrete format such as 0 or 1.
- Logistic regression algorithm works with the categorical variable such as 0 or 1, Yes or No, True or False, Spam or

not spam, etc. But instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1.

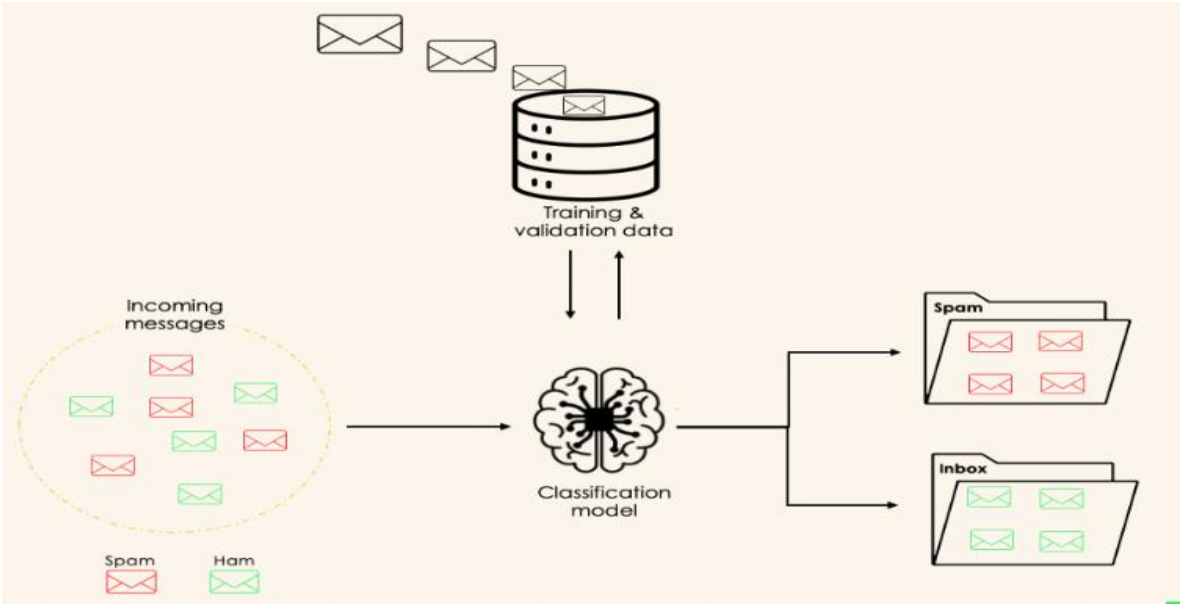
- It is a predictive analysis algorithm which works on the concept of probability.
- In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1).
- Logistic Regression is a significant machine learning algorithm because it has the ability to provide probabilities and classify new data using continuous and discrete datasets.
- Logistic regression uses sigmoid function or logistic function which is a complex cost function. This sigmoid function is used to model the data in logistic regression. The function can be represented as:
- $f(x)= \frac{1}{1+e^{-x}}$
  - f(x)= Output between the 0 and 1 value
  - x= input to the function
  - e= base of natural logarithm
- When we provide the input values (data) to the function, it gives the S-curve as follows:

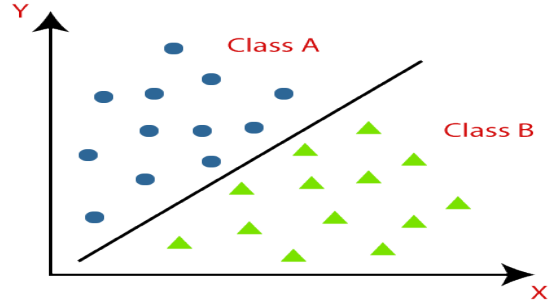



- It uses the concept of threshold levels, values above the threshold level are rounded up to 1, and values below the threshold level are rounded up to 0.
- Assumptions for logistic regression : The dependent variable must be categorical in nature and the independent variable should not have multi-collinearity.
- There are three types of logistic regression:
  - **Binary(0/1, pass/fail) :** In binomial logistic regression, there can be only two possible types of the dependent variables.
  - **Multi(cats, dogs, lions) :** In multinomial logistic regression, there can be 3 or more possible unordered types of the dependent variable.
  - **Ordinal(low, medium, high) :** In ordinal logistic regression, there can be 3 or more possible ordered types of dependent variables.

## ❖ Classification Algorithm In Machine Learning

- The Classification algorithm is a Supervised Learning technique that is used to identify the category of new observations on the basis of training data.
- In Classification, a program learns from the given dataset or observations and then classifies new observation into a number of classes or groups. Such as, Yes or No, 0 or 1, Spam or Not Spam, cat or dog, etc. Classes can be called as targets/labels or categories.
- Unlike regression, the output variable of Classification is a category, not a value, such as "Green or Blue", "fruit or animal", etc.
- Since the Classification algorithm is a Supervised learning technique, hence it takes labeled input data, which means it contains input with the corresponding output.
- In classification algorithm, a discrete output function(y) is mapped to input variable(x).
- **Y = f(X)** where Y = categorical output.
- The best example of an ML classification algorithm is Email Spam Detector.

- 

- The main goal of the Classification algorithm is to identify the category of a given dataset, and these algorithms are mainly used to predict the output for the categorical data.
- Classification algorithms can be better understood using the below diagram. In the below diagram, there are two classes, class A and Class B. These classes have features that are similar to each other and dissimilar to other classes.
- 

- The algorithm which implements the classification on a dataset is known as a classifier. There are two types of Classifications :
  - **Binary Classifier :** If the classification problem has only two possible outcomes, then it is called as binary classifier. Eg., yes or no, male or female, spam or not spam, cat or dog, etc.
  - **Multi-class Classifier :** If a classification problem has more than two outcomes, then it is called as multi-class classifier. Eg., Classifications of types of crops, Classification of types of music.

→ **Learners In The Classification Problems**

- **Lazy Learners:** Lazy Learner firstly stores the training dataset and wait until it receives the test dataset. In Lazy learner case, classification is done on the basis of the most related data stored in the training dataset. It takes less time in training but more time for predictions.
  Example: K-NN algorithm, Case-based reasoning
- **Eager Learners:** Eager Learners develop a classification model based on a training dataset before receiving a test dataset. Opposite to Lazy learners, Eager Learner takes more time in learning, and less time in prediction. Example: Decision Trees, Naïve Bayes, ANN.

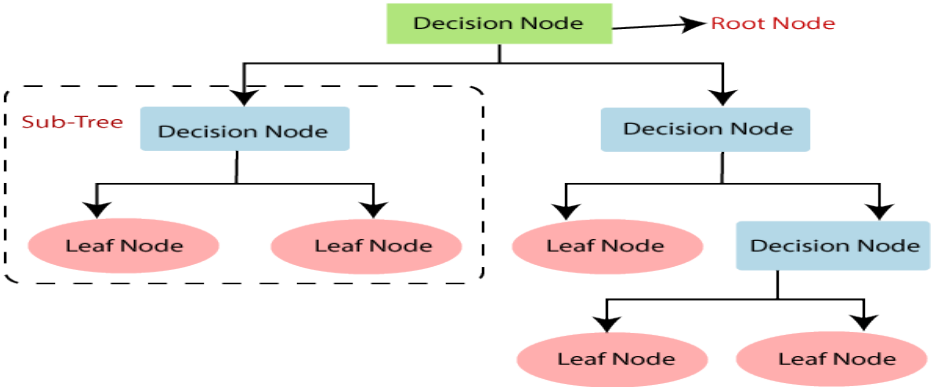
→ **Types Of ML Classification Algorithms :**

- Linear Models : Logistic Regression, Support Vector Machines.
- Non-Linear Models : K-Nearest Neighbours, Kernel SVM, Naïve Bayes, Decision Tree Classification, Random Forest Classification

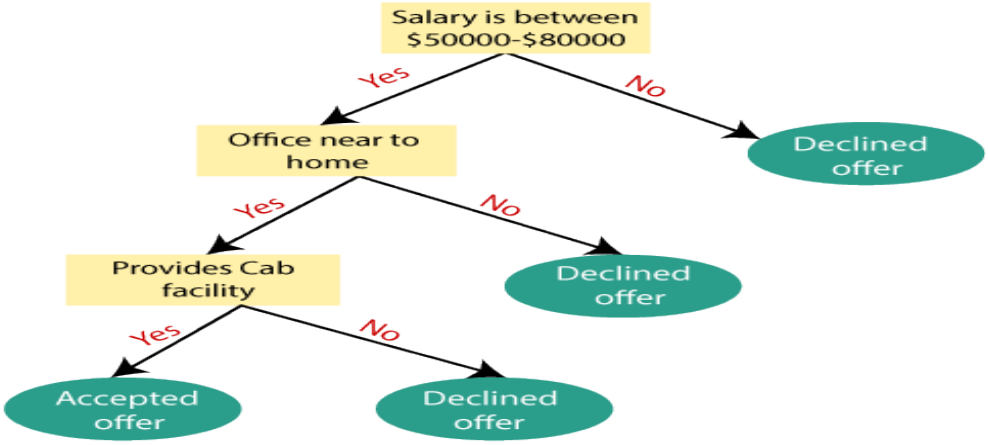
## ❖ Decision Tree

- Decision Tree is a supervised learning algorithm which can be used for solving both classification and regression problems. But preferably used for solving classification problems.

- It can solve problems for both categorical and numerical data.
- In a Decision tree, there are two nodes, which are the Decision Node and Leaf Node. Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches.
- It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions.
- In order to build a tree, we use the CART algorithm, which stands for Classification and Regression Tree algorithm.
- Consider the below image:




- Example : Suppose there is a candidate who has a job offer and wants to decide whether he should accept the offer or Not. So, to solve this problem, the decision tree starts with the root node (Salary attribute by ASM). The root node splits further into the next decision node (distance from the office) and one leaf node. The next decision node further gets split into one decision node (Cab facility) and one leaf node. Finally, the decision node splits into two leaf nodes (Accepted offers and Declined offer). Consider the below diagram:




→ **Advantages & Disadvantages :**

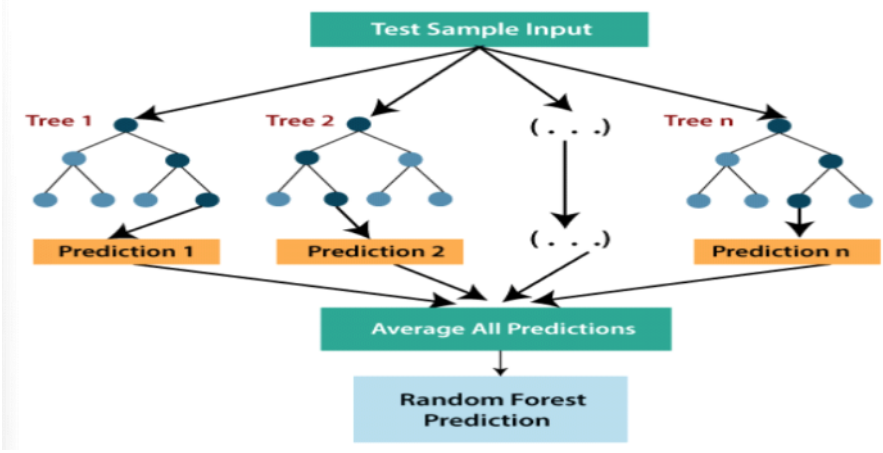
- Decision trees are interpretable, easy to understand.
- Can handle both numerical and categorical data.
- There is less requirement of data cleaning compared to other algorithm.
- It may have an overfitting issue, which can be resolved using the Random Forest Algorithm.
- For more class labels, the computational complexity of the decision tree may increase.

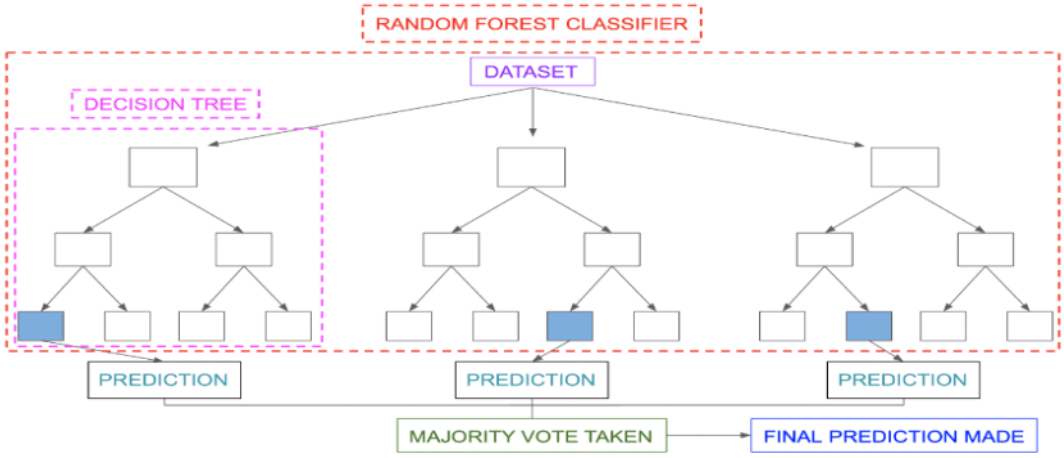
## ❖ Random Forest

- Random forest is one of the most powerful supervised learning algorithms which is capable of performing regression as well as classification tasks.
- The Random Forest regression is an ensemble learning method which combines multiple decision trees and predicts the final output based on the average of each tree output.
- Random Forest regression is an is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.
- Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.
- The combined decision trees are called as base models, and it can be represented more formally as :

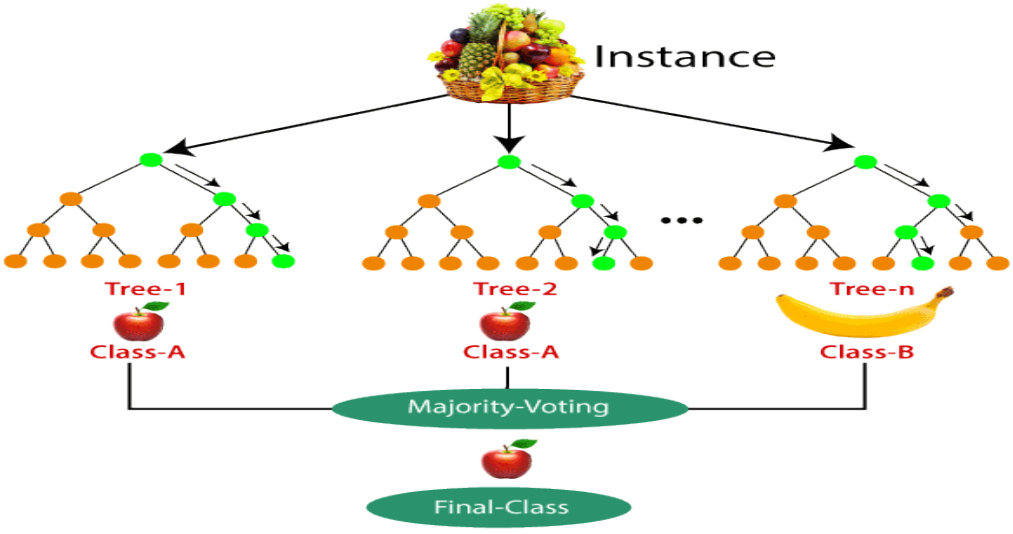
**g(x) = f0(x) + f1(x) + f2(x) + ….**

- Random forest uses Bagging or Bootstrap Aggregation technique of ensemble learning in which aggregated decision tree runs in parallel and do not interact with each other.
- With the help of Random Forest regression, we can prevent Overfitting in the model by creating random subsets of the dataset.



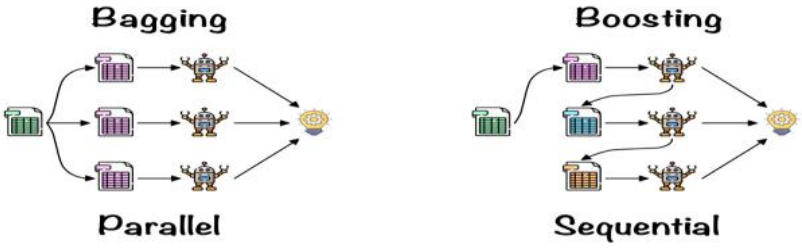




- Example : Suppose there is a dataset that contains multiple fruit images. So, this dataset is given to the Random forest classifier. The dataset is divided into subsets and given to each decision tree. During the training phase, each decision tree produces a prediction result, and when a new data point occurs, then based on the majority of results, the Random Forest classifier predicts the final decision. Consider the below image:




→ **Ensemble uses two types of methods:**

- Bagging– It creates a different training subset from sample training data with replacement & the final output is based on majority voting. For example, Random Forest.
- Boosting– It combines weak learners into strong learners by creating sequential models such that the final model has the highest accuracy. For example, ADA BOOST, XG BOOST.




→ **Advantages & Disadvantages**

- Random Forest is capable of performing both Classification and Regression tasks.
- It is capable of handling large datasets with high dimensionality.
- It enhances the accuracy of the model and prevents the overfitting issue.
- Although it can perform both regression and classification tasks, it is not suitable for regression task and more resources are required for computation.
- It consumes more time compared to a decision tree algorithm.

→ **Applications of Random Forest**

- Banking: Banking sector mostly uses this algorithm for the identification of loan risk.
- Medicine: With the help of this algorithm, disease trends and risks of the disease can be identified.
- Land Use: We can identify the areas of similar land use by this algorithm.
- Marketing: Marketing trends can be identified using this algorithm.

## ❖ Naïve Bayes

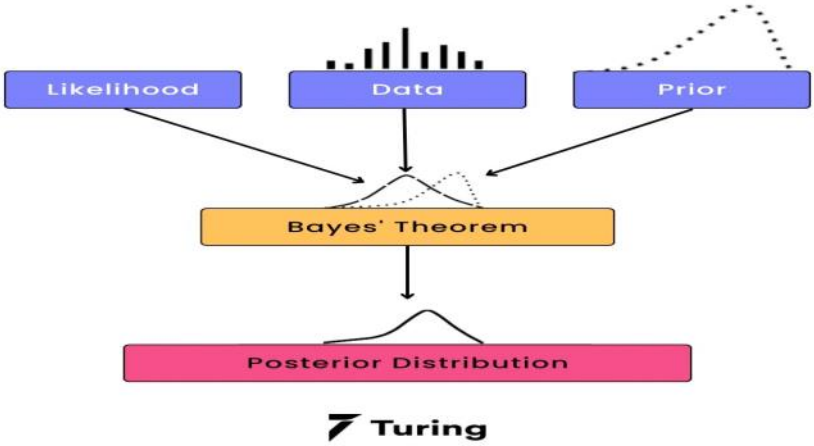
- Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems.
- It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.
- It assumes the features are independent of each other. E.g., if given a banana, the classifier will see that the fruit is of yellow colour, oblong shaped and long and tapered. All of these features will contribute independently to the probability of it being a banana and are not independent on each other.
- Algorithm is widely used for various classification task, especially in NLP and document classification that includes a high-dimensional training dataset.
- Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles.

→ **Bayes Theorem**

- Bayes Theory works on coming to a hypothesis (H) from a given set of evidence (E). It relates to two things: The probability of the hypothesis before the evidence P(H) and the probability after the evidence P(H|E).
- The Bayes Theory is explained by the following equation:
  - $\mathbf{P(H|E)} = \frac{\mathbf{P(E|H)} * \mathbf{P(H)}}{\mathbf{P(E)}}$

where,

P(H|E) Posterior Probability - denotes how event H happens when event E takes place.
P(E|H) Likelihood Probability - represents how often event E happens when event H takes place first.
P(H) Prior Probability - represents the probability of event X happening on its own.
P(E) Marginal Probability - represents the probability of event Y happening on its own.




→ **Types of the Naive Bayes Model**


– There are four types of the Naive Bayes Model, which are explained below:

**(a) Gaussian Naive Bayes**

– The Gaussian model assumes that features follow a normal distribution. This means if predictors take continuous values instead of discrete, then the model assumes that these values are sampled from the Gaussian Distribution.

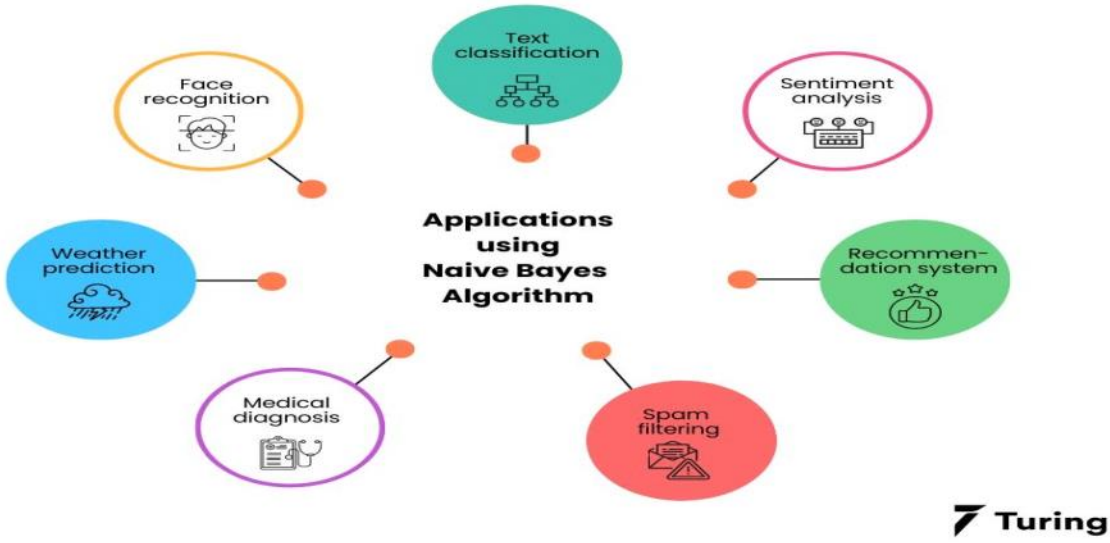
**(b) Bernoulli Naive Bayes**

– Bernoulli Naive Bayes is an algorithm that is useful for data that has binary or boolean attributes. The attributes will have a value of yes or no, useful or not, granted or rejected, etc. This model is also famous for document classification task.

**(c) Multinominal Naive Bayes**

– Multinominal Naive Bayes is used on documentation classification problems, it means a particular document belongs to which category such as sports, politics, education, etc. The classifier uses the frequency of words for the predictors.

→ **Applications that use Naive Bayes**

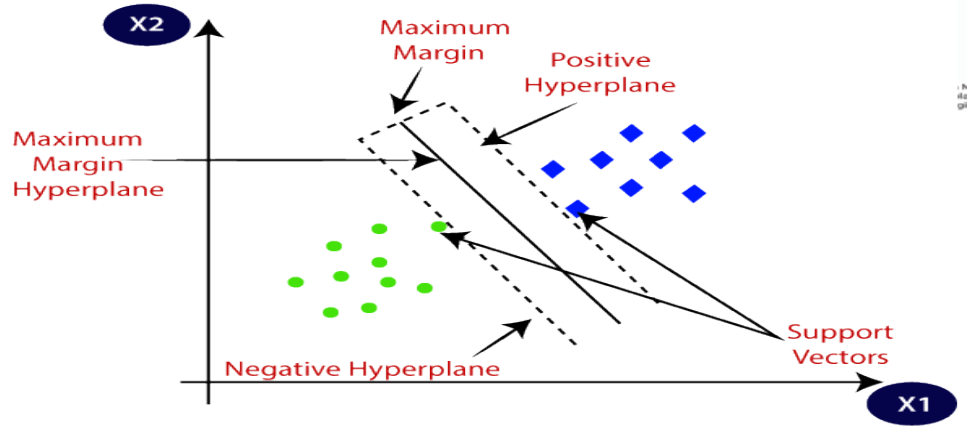



→ **Advantages & of Naïve Bayes Classifier**

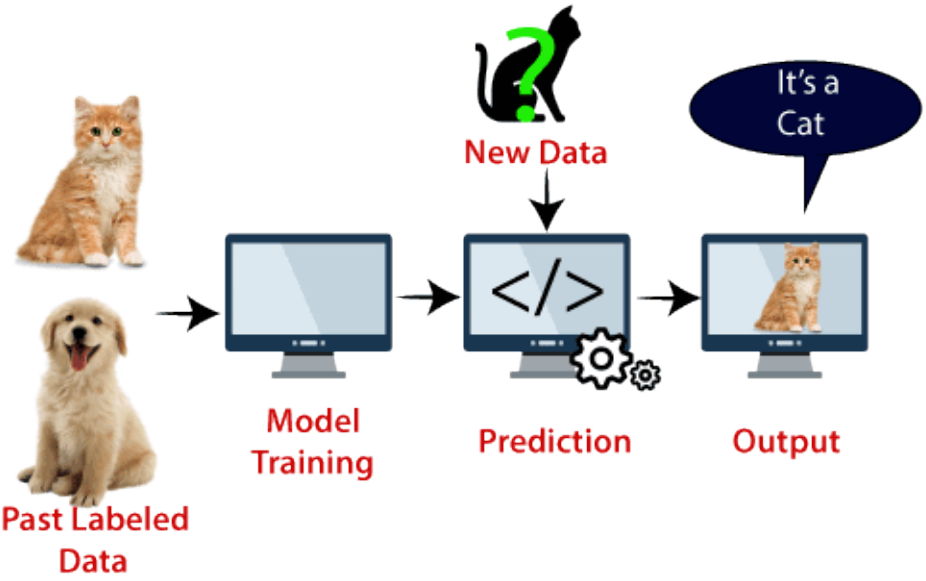
- Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.
- It can be used for Binary as well as Multi-class Classifications.
- It performs well in Multi-class predictions as compared to the other Algorithms.
- It is the most popular choice for text classification problems.
- Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between features.

## ❖ Support Vector Machine (SVM)

- Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.
- The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.
- SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine.
- Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane:




- **Example:** Suppose we see a strange cat that also has some features of dogs, so if we want a model that can accurately identify whether it is a cat or dog, so such a model can be created by using the SVM algorithm. We will first train our model with lots of images of cats and dogs so that it can learn about different features of cats and dogs, and then we test it with this strange creature. So as support vector creates a decision boundary between these two data (cat and dog) and choose extreme cases (support vectors), it will see the extreme case of cat and dog. On the basis of the support vectors, it will classify it as a cat. Consider the below diagram:

– 


– SVM algorithm can be used for **Face detection, image classification, text categorization,** etc.

→ **Types of SVM**

– **Linear SVM:** Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier employed is known as Linear SVM Classifier.

– **Non-linear SVM:** Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.

→ **Hyperplane and Support Vectors in the SVM algorithm**

**(c) Hyperplane:**

– There can be multiple lines/decision boundaries to segregate the classes in n-dimensional space, but we need to find out the best decision boundary that helps to classify the data points. This best boundary is known as the hyperplane of SVM.

– The dimensions of the hyperplane depend on the features present in the dataset, which means if there are 2 features (as shown in image), then hyperplane will be a straight line. And if there are 3 features, then hyperplane will be a 2-dimension plane.

– We always create a hyperplane that has a maximum margin, which means the maximum distance between the data points.

**(d) Support Vectors:**

– The data points or vectors that are the closest to the hyperplane and which affect the position of the hyperplane are termed as Support Vector.

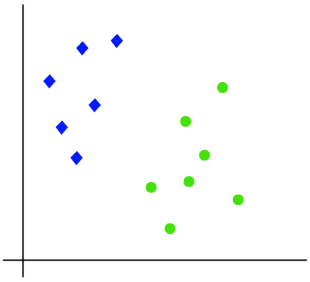
– Since these vectors support the hyperplane, hence called a Support vector.

→ **Working Of SVM**

**(a) Linear SVM:**

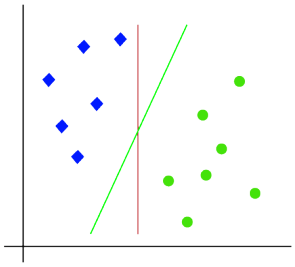
– The working of the SVM algorithm can be understood by using an example. Suppose we have a dataset that has two tags (green and blue), and the dataset has two features x1 and x2.

– We want a classifier that can classify the pair(x1, x2) of coordinates in either green or blue. Consider the below image:

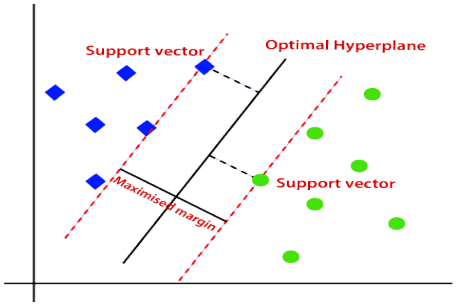


– So as it is 2-d space so by just using a straight line, we can easily separate these two classes. But there can be

multiple lines that can separate these classes. Consider the below image:

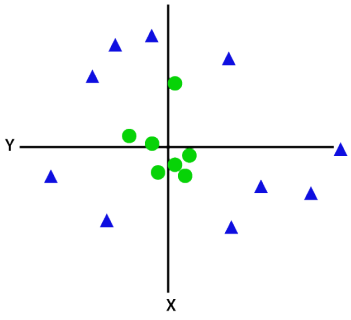


- Hence, the SVM algorithm helps to find the best line or decision boundary; this best boundary or region is called as a hyperplane.
- SVM algorithm finds the closest point of the lines from both the classes. These points are called support vectors.
- The distance between the vectors and the hyperplane is called as margin and the goal of SVM is to maximize this margin. The hyperplane with maximum margin is called the optimal hyperplane.

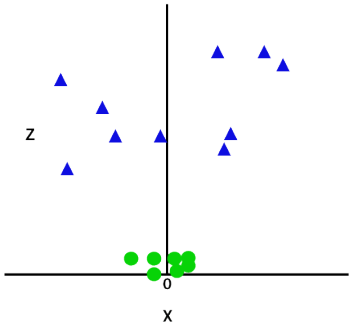



**(b) Non-Linear SVM:**

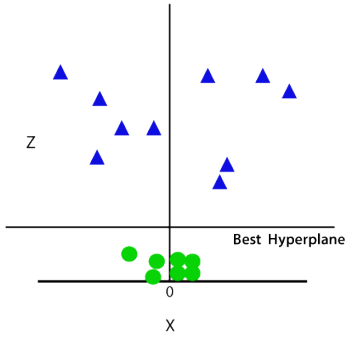
- If data is linearly arranged, then we can separate it by using a straight line, but for non-linear data, we cannot draw a single straight line. Consider the below image:



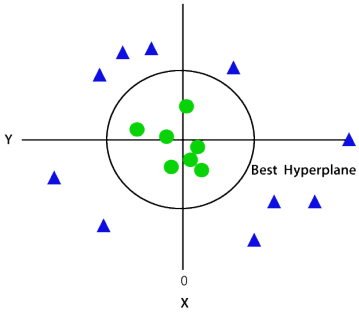

- So to separate these data points, we need to add one more dimension. For linear data, we have used two dimensions x and y, so for non-linear data, we will add a third dimension z.
- It can be calculated as:
  - **z = x2 + y2**
- By adding the third dimension, the sample space will become as below image:




- So now, SVM will divide the datasets into classes in the following way. Consider the below image:

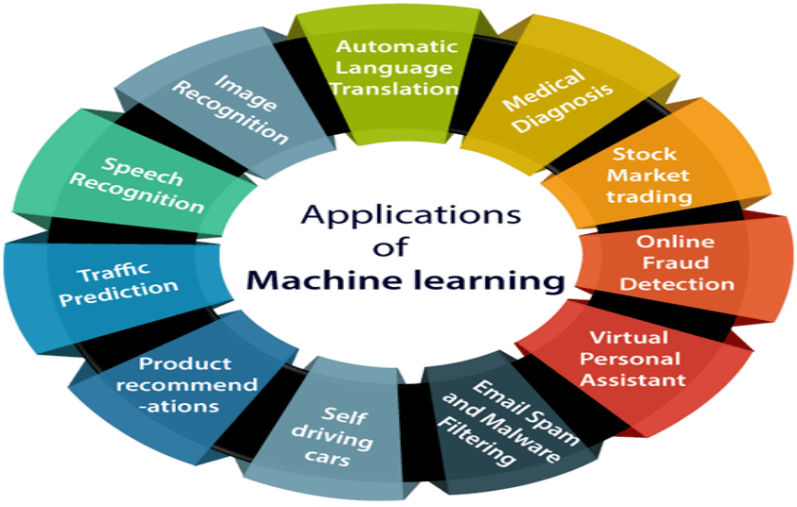



– Since we are in 3-d Space, hence it is looking like a plane parallel to the x-axis. If we convert it in 2d space with z= 1, then it will become as:




– Hence we get a circumference of radius 1 in case of non-linear data.

## ❖ Applications Of Machine Learning




**(a) Image Recognition:**

– Image recognition is one of the most common applications of machine learning. It is used to identify objects, persons, places, digital images, etc. The popular use case of image recognition and face detection is, Automatic friend tagging suggestion:

– Facebook provides us a feature of auto friend tagging suggestion. Whenever we upload a photo with our Facebook friends, then we automatically get a tagging suggestion with name, and the technology behind this is machine learning's face detection and recognition algorithm.

– It is based on the Facebook project named "Deep Face," which is responsible for face recognition and person identification in the picture.

**(b) Speech Recognition**

– While using Google, we get an option of "Search by voice," it comes under speech recognition, and it's a popular application of machine learning.

– Speech recognition is a process of converting voice instructions into text, and it is also known as "Speech to text", or "Computer speech recognition."

– At present, machine learning algorithms are widely used by various applications of speech recognition. Google assistant, Siri, Cortana, and Alexa are using speech recognition technology to follow the voice instructions.

**(c) Traffic prediction:**

- If we want to visit a new place, we take help of Google Maps, which shows us the correct path with the shortest route and predicts the traffic conditions.
- It predicts the traffic conditions such as whether traffic is cleared, slow-moving, or heavily congested with the help of two ways:
    - Real Time location of the vehicle form Google Map app and sensors
    - Average time has taken on past days at the same time.
- Everyone who is using Google Map is helping this app to make it better. It takes information from the user and sends back to its database to improve the performance.

**(d) Email Spam and Malware Filtering:**

- Whenever we receive a new email, it is filtered automatically as important, normal, and spam. We always receive an important mail in our inbox with the important symbol and spam emails in our spam box, and the technology behind this is Machine learning.
- Below are some spam filters used by Gmail:
    - Content Filter
    - Header filter
    - General blacklists filter
    - Rules-based filters
    - Permission filters
- Some machine learning algorithms such as Multi-Layer Perceptron, Decision tree, and Naïve Bayes classifier are used for email spam filtering and malware detection.

**(e) Virtual Personal Assistant:**

- We have various virtual personal assistants such as Google assistant, Alexa, Cortana, Siri. As the name suggests, they help us in finding the information using our voice instruction. These assistants can help us in various ways just by our voice instructions such as Play music, call someone, Open an email, Scheduling an appointment, etc.
- These virtual assistants use machine learning algorithms as an important part.
- These assistant record our voice instructions, send it over the server on a cloud, and decode it using ML algorithms and act accordingly.