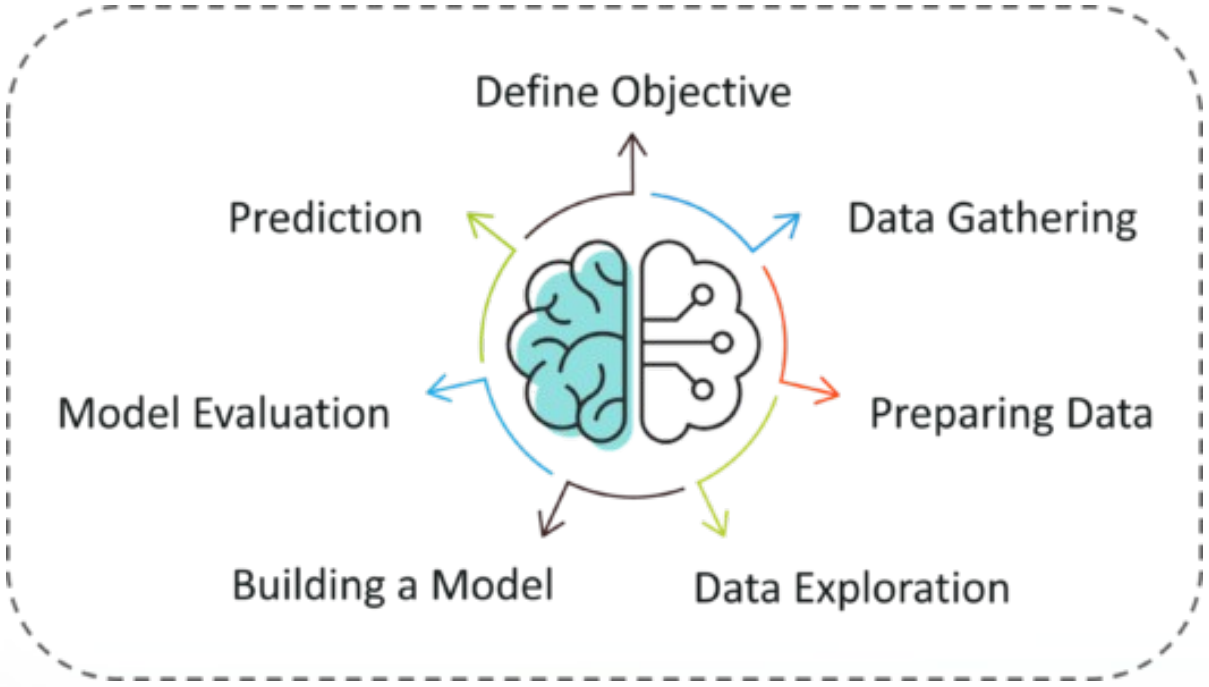
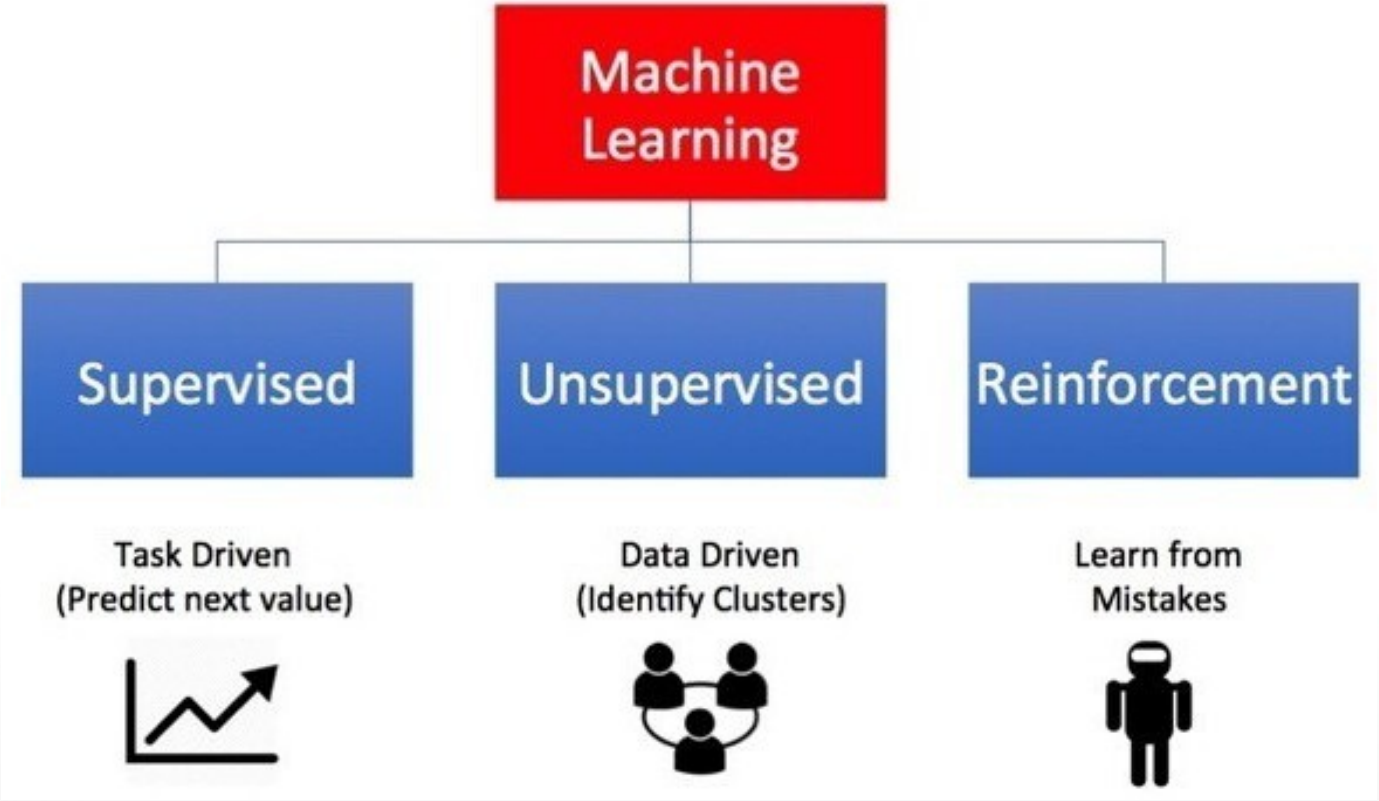
Machine Learning

Machine Learning Process



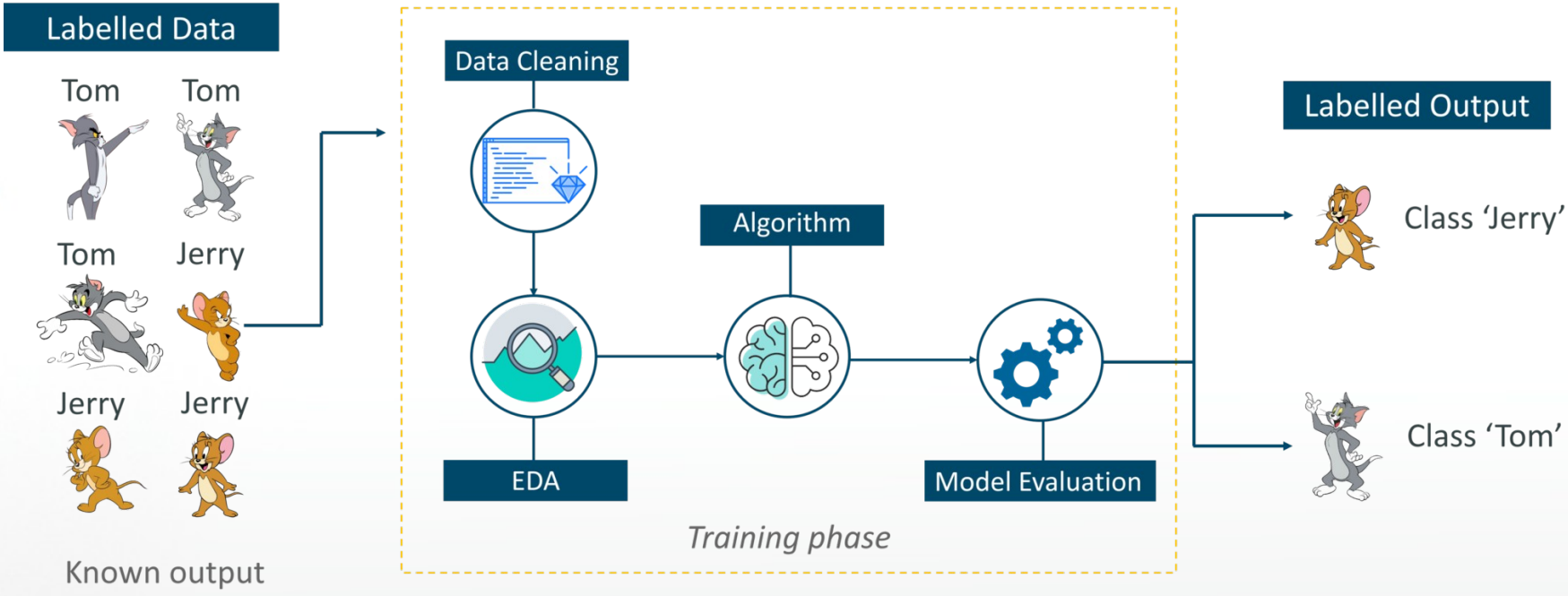
- In above steps, most time consuming steps are preparing data and data explorartion where we have to clean data and we have to put it in structured form as well as we have to do feature engineering.

Types of Machine Learning



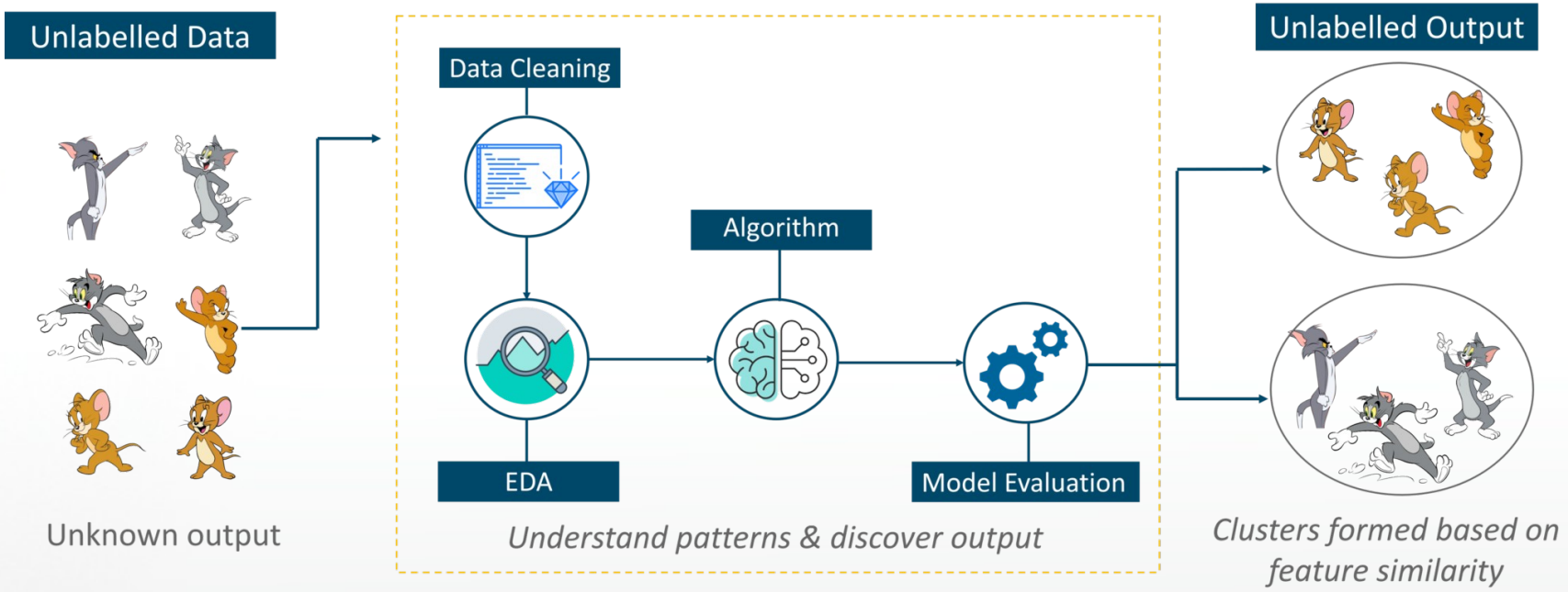
Supervised Learning

- Supervised learning is a technique in which we teach or train the machine using data which is well labeled.
- This labelled data set is nothing but tarining data set.



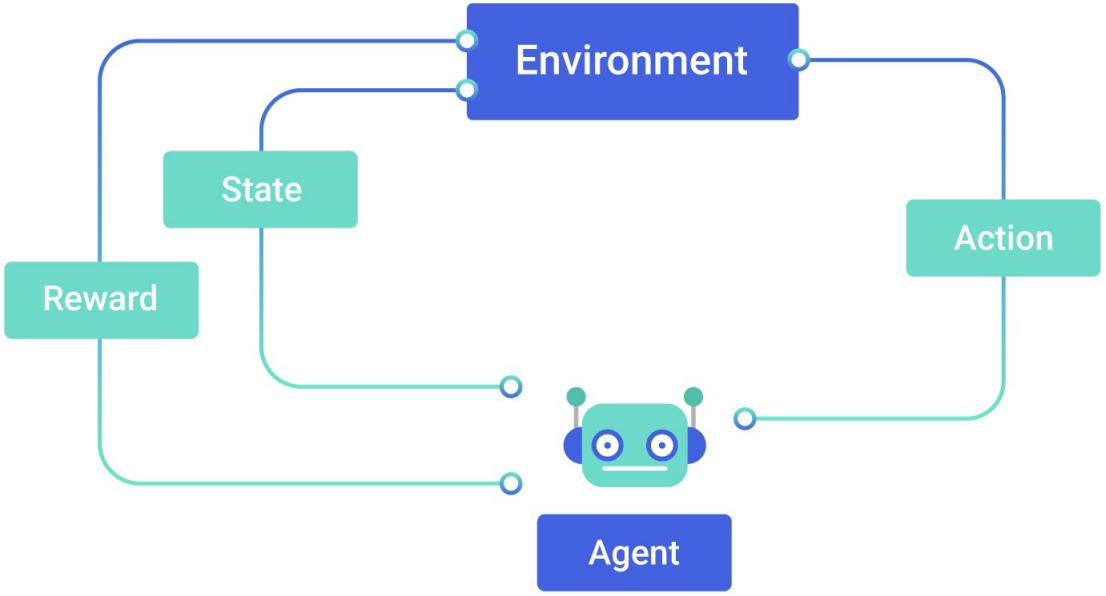
Unsupervised Learning

- Unsupervised learning involves training by using unlabeled data and allowing the model to act on that information without guidance.
- Ex: Clustering and Association Analysis

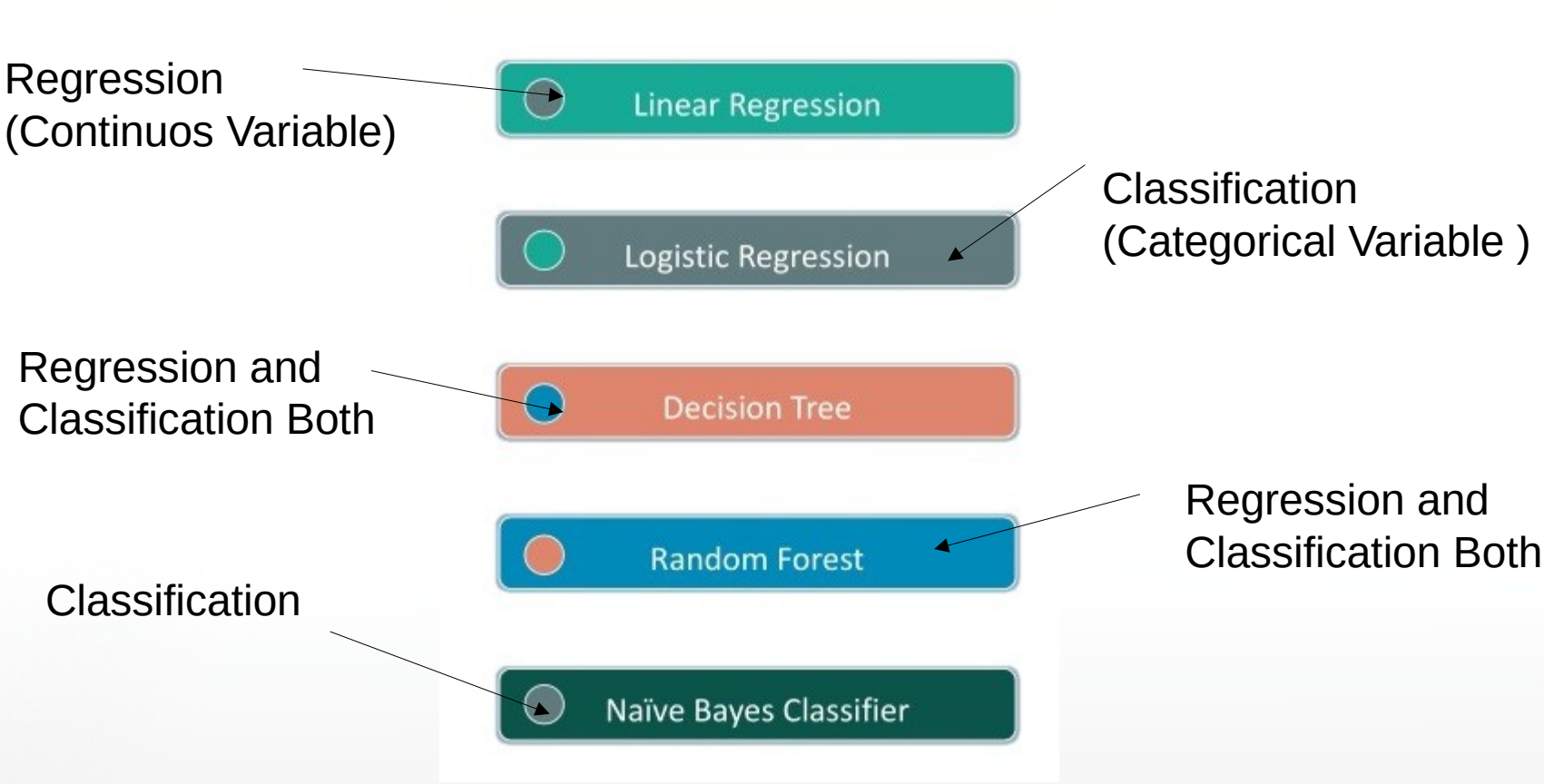


Reinforcement Learning

- Reinforcement Learning is a part of Machine learning where an agent is put in an environment and he learns to behave in this environment by performing certain actions and observing the rewards which it gets from those actions.

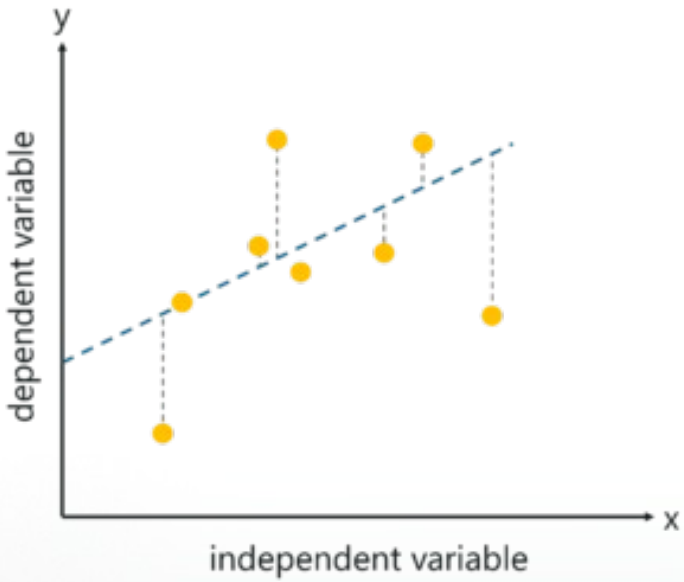


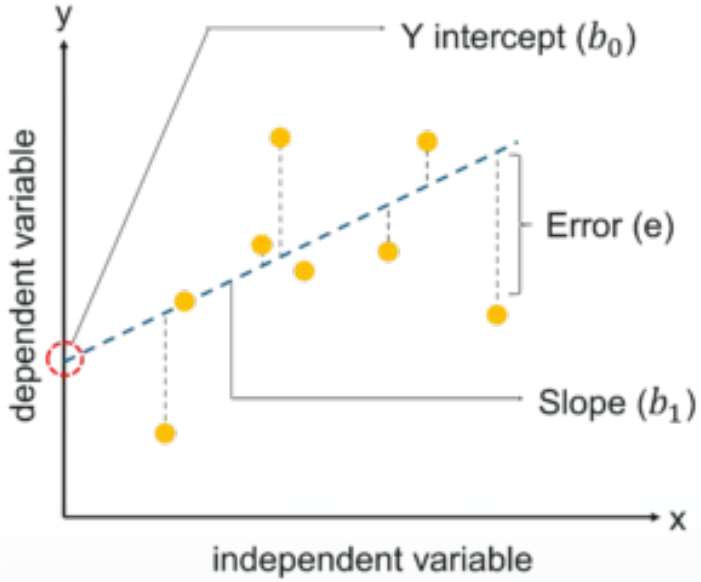
Types of Supervised Learning



Linear Regression

- It is a predictive modelling technique used to predict a continuous dependent variable, given one or more independent variables.

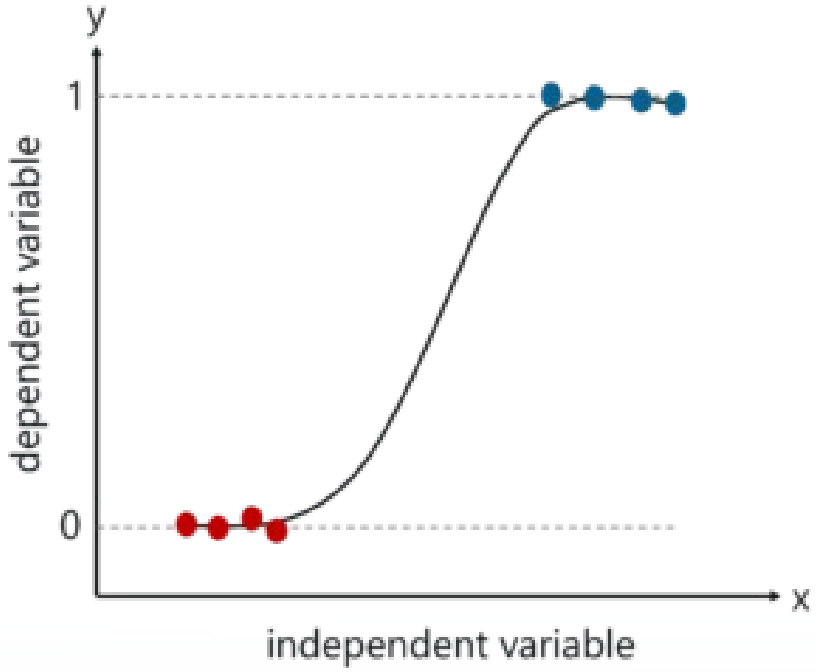




Linear Regression Equation : $\mathbf{Y = b_0 + b_1 x + e}$

Logistic Regression

- Logistic Regression is a method used to predict a dependent variable (Y), given an independent variable (X), such that the dependent variable is categorical.
- predict the outcome in a binary variable, meaning it will have only two outcomes.
- We use the Sigmoid function/curve to predict the categorical value. The threshold value decides the outcome



Logistic Regression Equation :

$$P(X) = \frac{e^{(\beta_0+\beta_1 x)}}{e^{(\beta_0+\beta x)} + 1} \longrightarrow \ln\left[\frac{p}{(1-p)}\right] = (\beta_0 + \beta_1 x)$$

Naive Bayes

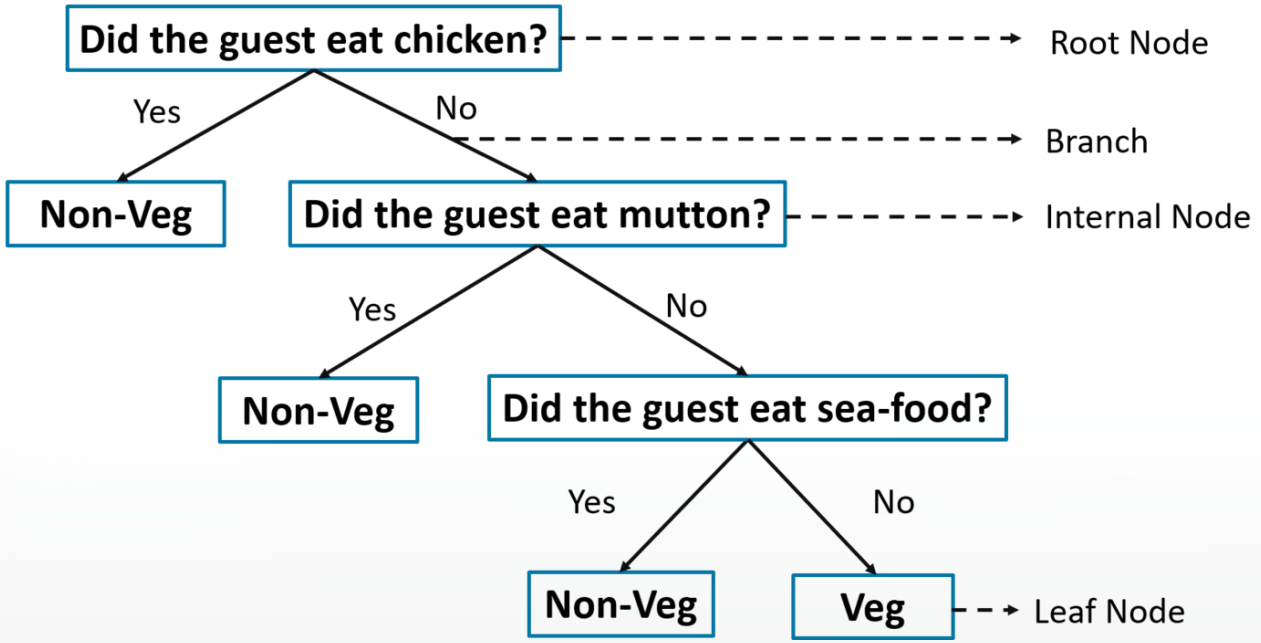
- The principle behind Naive Bayes is the Bayes theorem also known as the Bayes Rule.

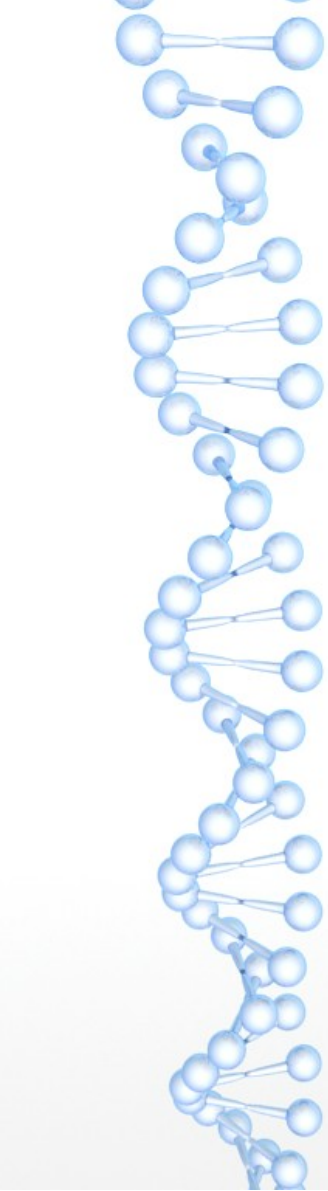
Bayes theorem : $$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- P(A|B): Conditional probability of event A occurring, given the event B
- P(A): Probability of event A occurring
- P(B): Probability of event B occurring
- P(B|A): Conditional probability of event B occurring, given the event A

Decision Tree

- A Decision Tree is a Supervised Machine Learning algorithm which looks like an inverted tree, wherein each node represents a predictor variable (feature), the link between the nodes represents a Decision and each leaf node represents an outcome (response variable).



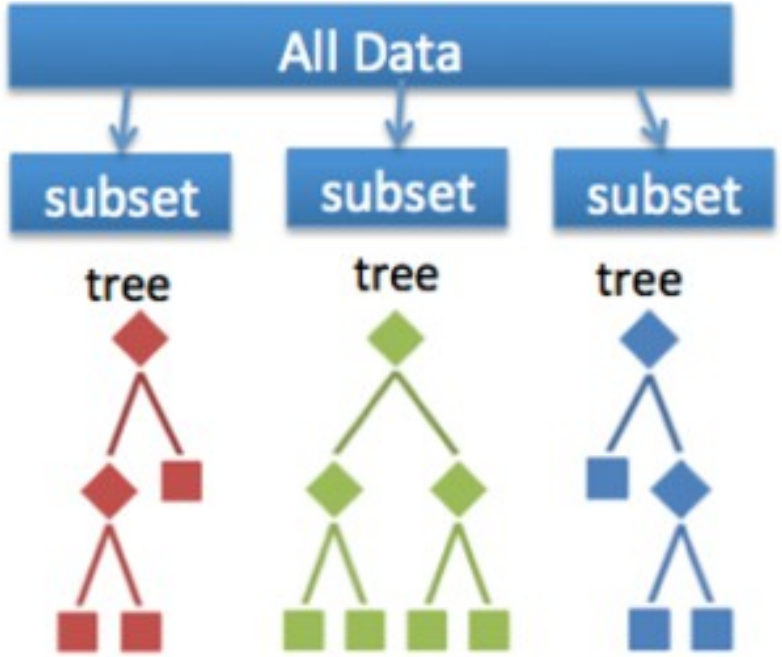


Decision Tree Using ID3 Algorithm

- The ID3 algorithm follows the below workflow in order to build a Decision Tree:
  - Select Best Attribute (A)
  - Assign A as a decision variable for the root node.
  - For each value of A, build a descendant of the node.
  - Assign classification labels to the leaf node.
  - If data is correctly classified: Stop.
  - Else: Iterate over the tree.
- Two measures are used to decide the best attribute:
  - Information Gain
  - Entropy

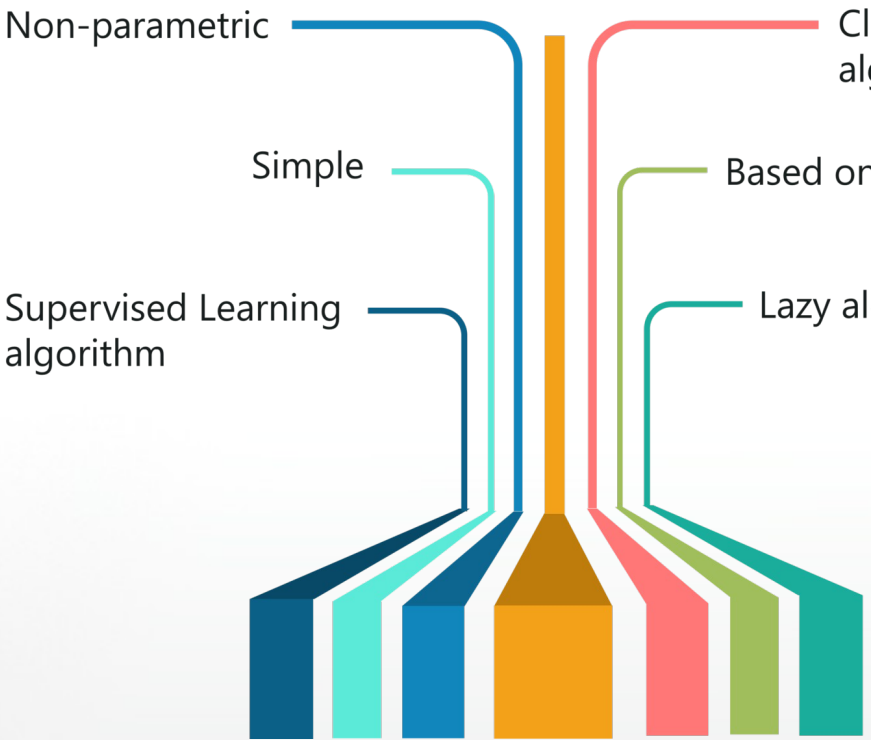
Random Forest

- Random forest builds multiple decision trees (called the forest) and glues them together to get a more accurate and stable prediction.
- Why Rnadom Forest ?
  - More Accuaracy
  - Avoid Overfitting
  - Bagging
- Bootstrpping is used for creating subset
- Out of bag dataset is for testing model
- This is used for classification as well as regression exaple



K Nearest Neighbour

- KNN which stand for K Nearest Neighbor is a Supervised Machine Learning algorithm
- It is used classifies a new data point into the target class, depending on the features of its neighboring data points.

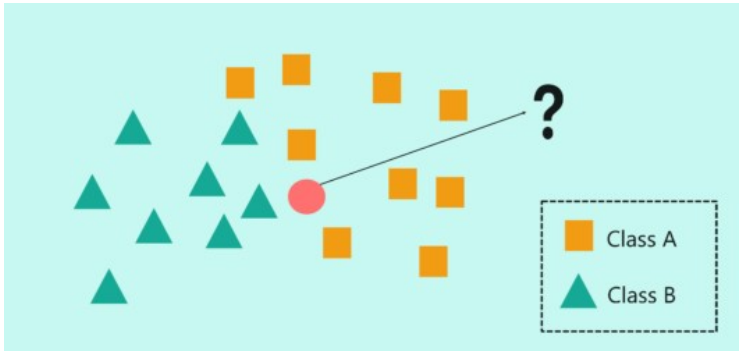


Point P1 = (1,4)

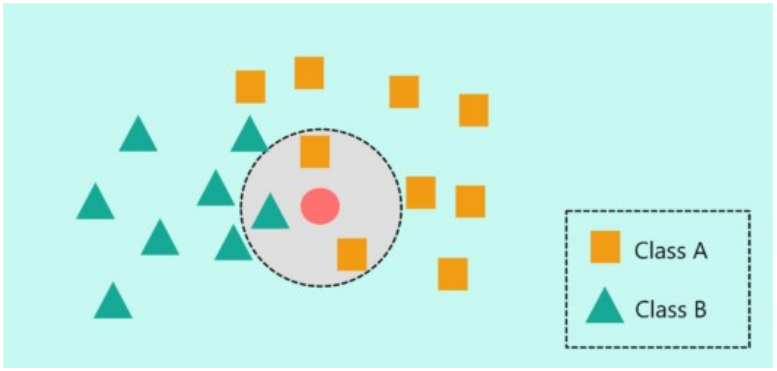
Point P2 = (5,1)

Euclidian distance = $\sqrt{(5-1)^2 + (4-1)^2} = 5$

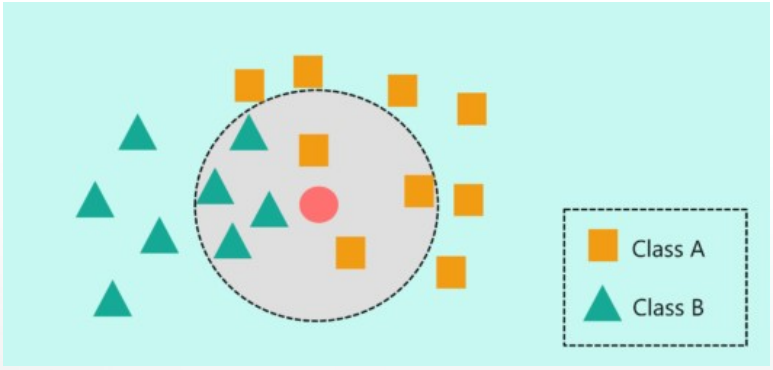
Example

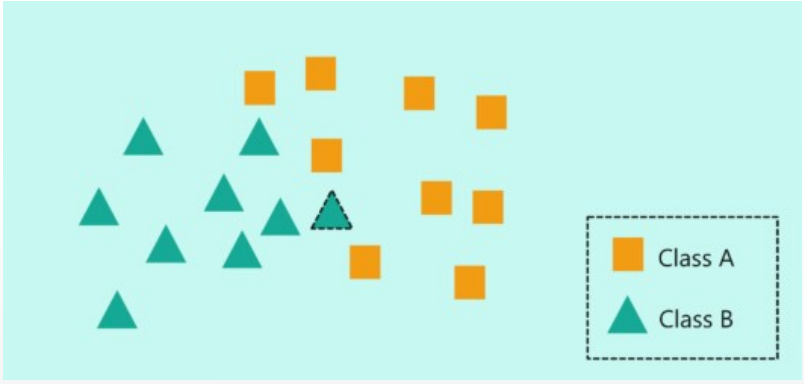


k=3



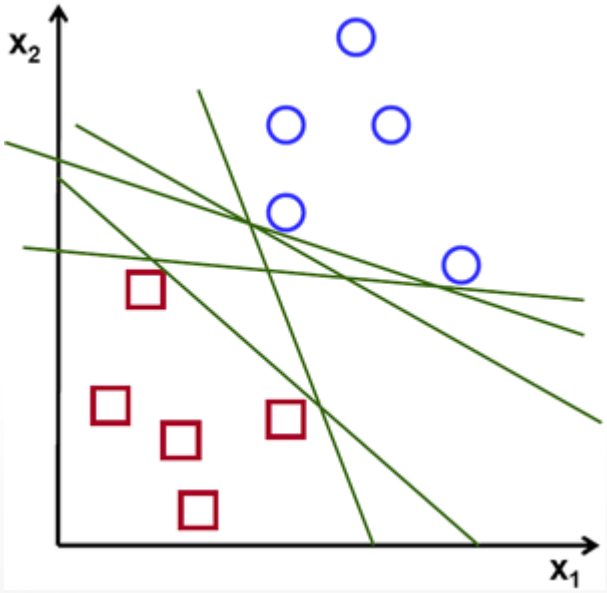
k=7

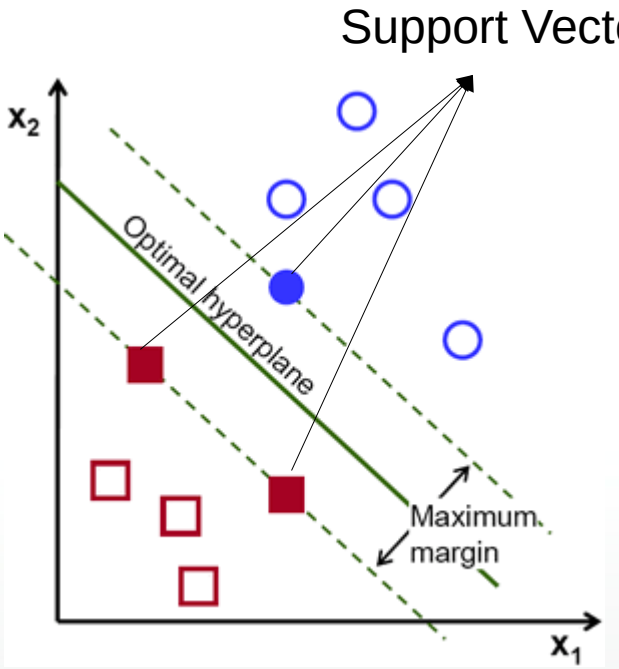




Support Vector Machine

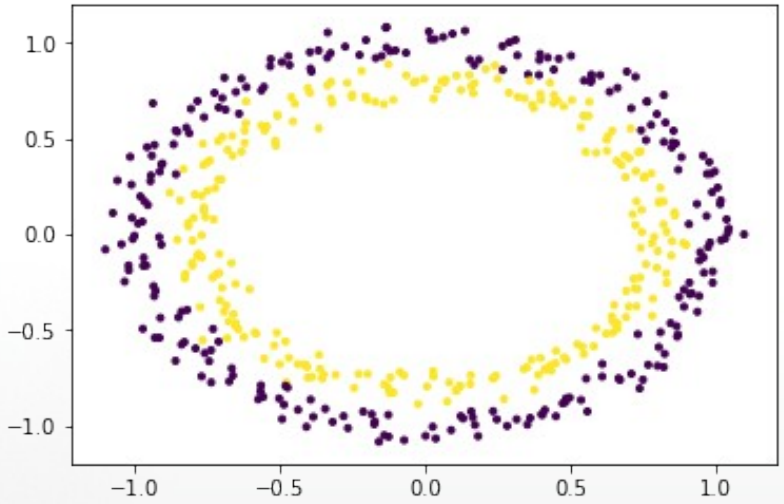
- The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space(N — the number of features) that distinctly classifies the data points.

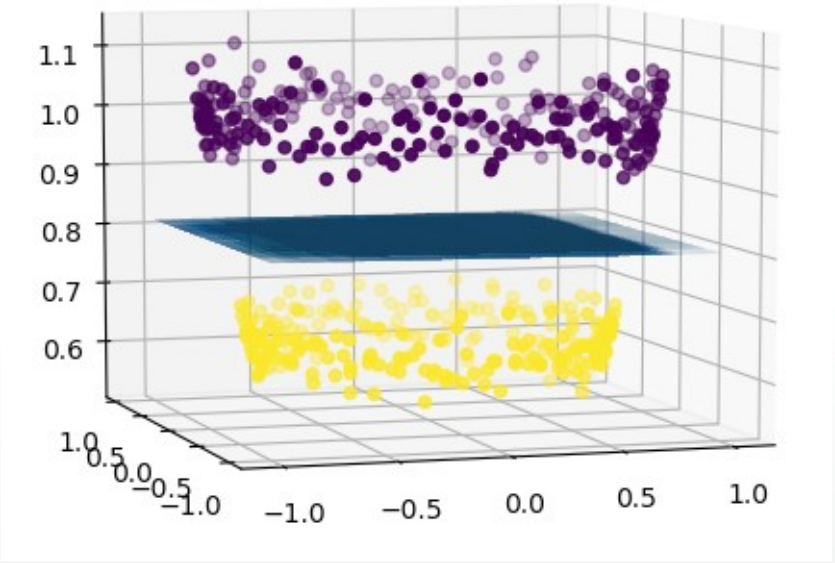




Non-Linear SVM

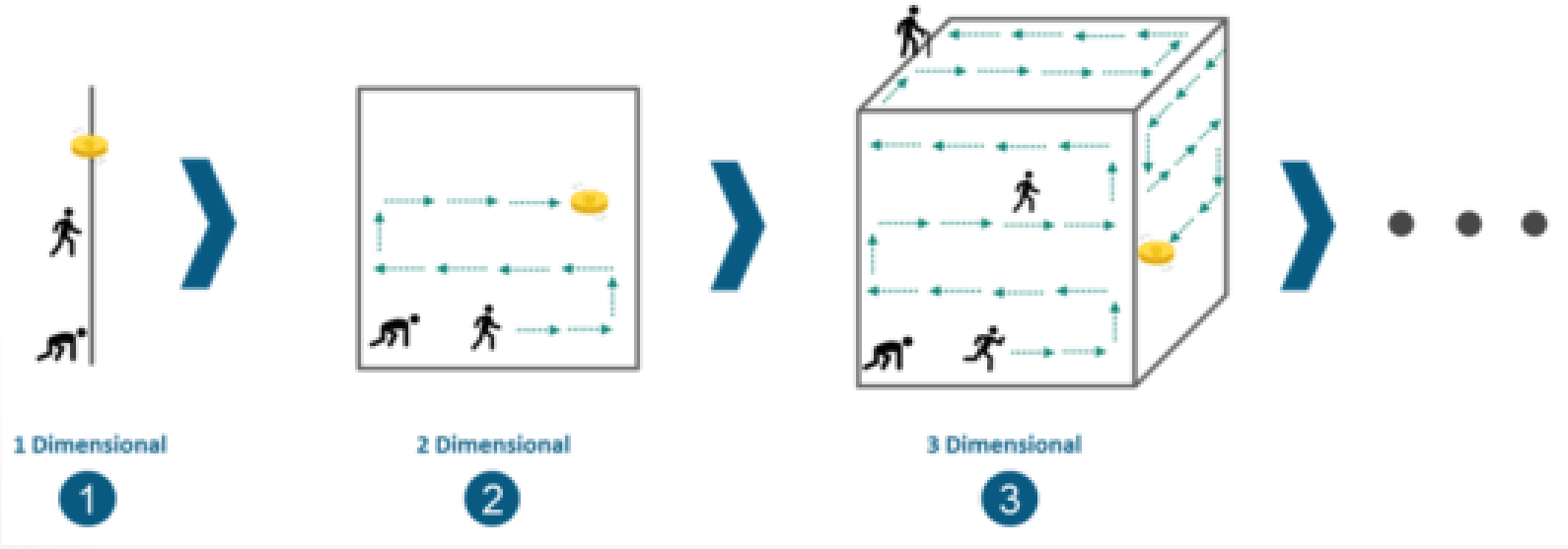
- In machine learning, a trick known as “kernel trick” is used to learn a linear classifier to classify a non-linear dataset. It transforms the linearly inseparable data into a linearly separable one by projecting it into a higher dimension.





Limitations of Machine Learning

- Machine Learning is not capable of handling high dimensional data that is where input & output is quite large.
- Handling and processing such type of data becomes very complex and resource exhaustive. This is termed as Curse of Dimensionality



Limitations of Machine Learning

- One of the biggest challange in machine learning model process is called feature extaraction.
- For complex problems like image recongnization, handwritting recognization its huge problem.