COURSE NAME: ARTIFICIAL INTELLIGENCE
COURSE CODE: CIS 412

SYED TANGIM PASHA

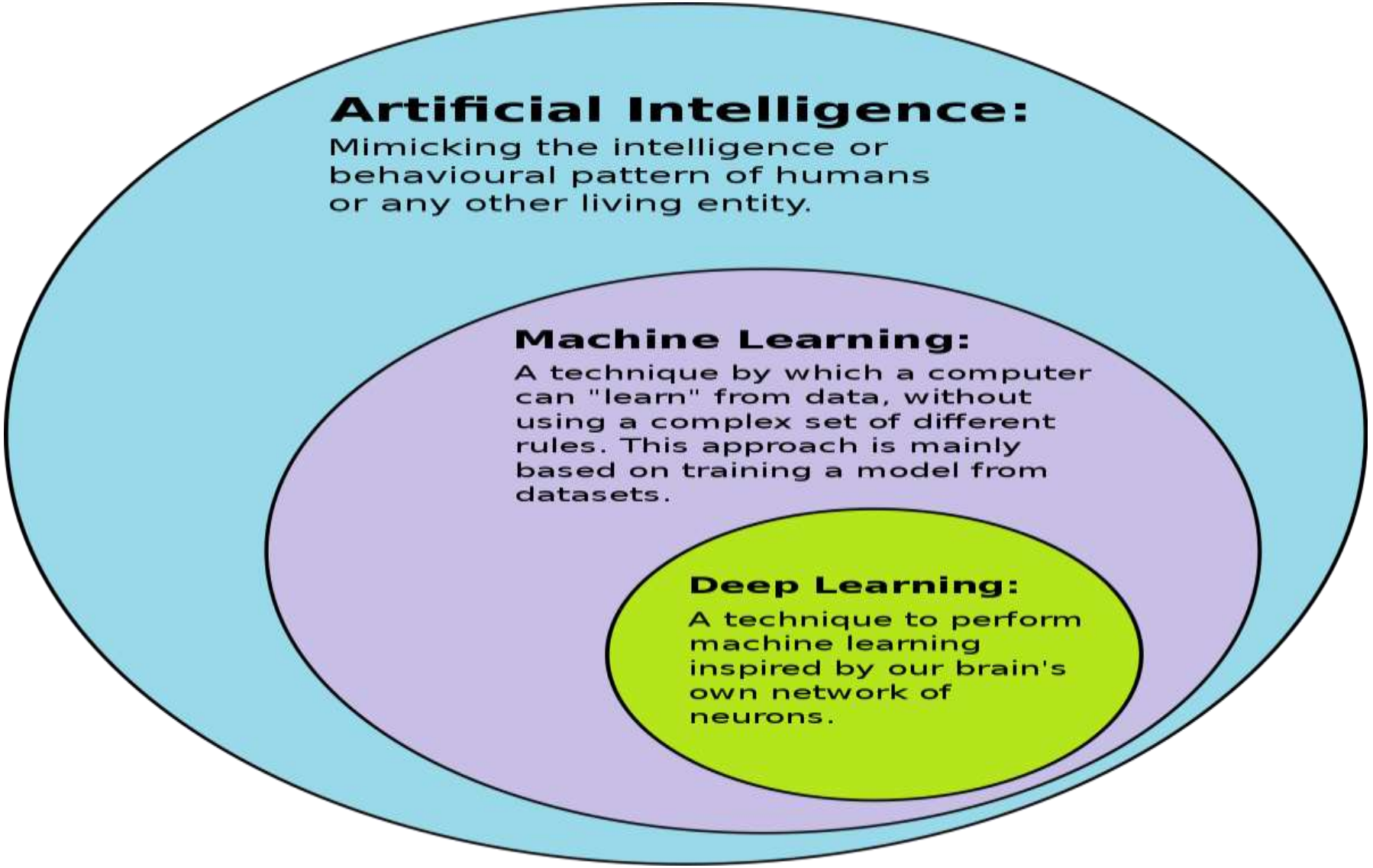
LECTURER,

DEPARTMENT OF COMPUTING AND INFORMATION SYSTEM (CIS)

DAFFODIL INTERNATIONAL UNIVERSITY (DIU)

DHAKA,BANGLADESH

MACHINE LEARNING



MACHINE LEARNING

- **Machine Learning:**
- Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.
- Machine Learning (ML) is the study of computer algorithms that can improve automatically through experience and by the use of data.
- **Arthur Samuel**, a pioneer in the field of Artificial Intelligence and computer gaming, coined the term “Machine Learning”.

Traditional Program vs ML

- **Traditional Programming:** We feed in data(Input) + program(logic), run it on machine and get output.

- **Machine Learning:** We feed in data(Input) + output, run it on machine during training and the machine creates its own program(logic), which can be evaluated while testing.

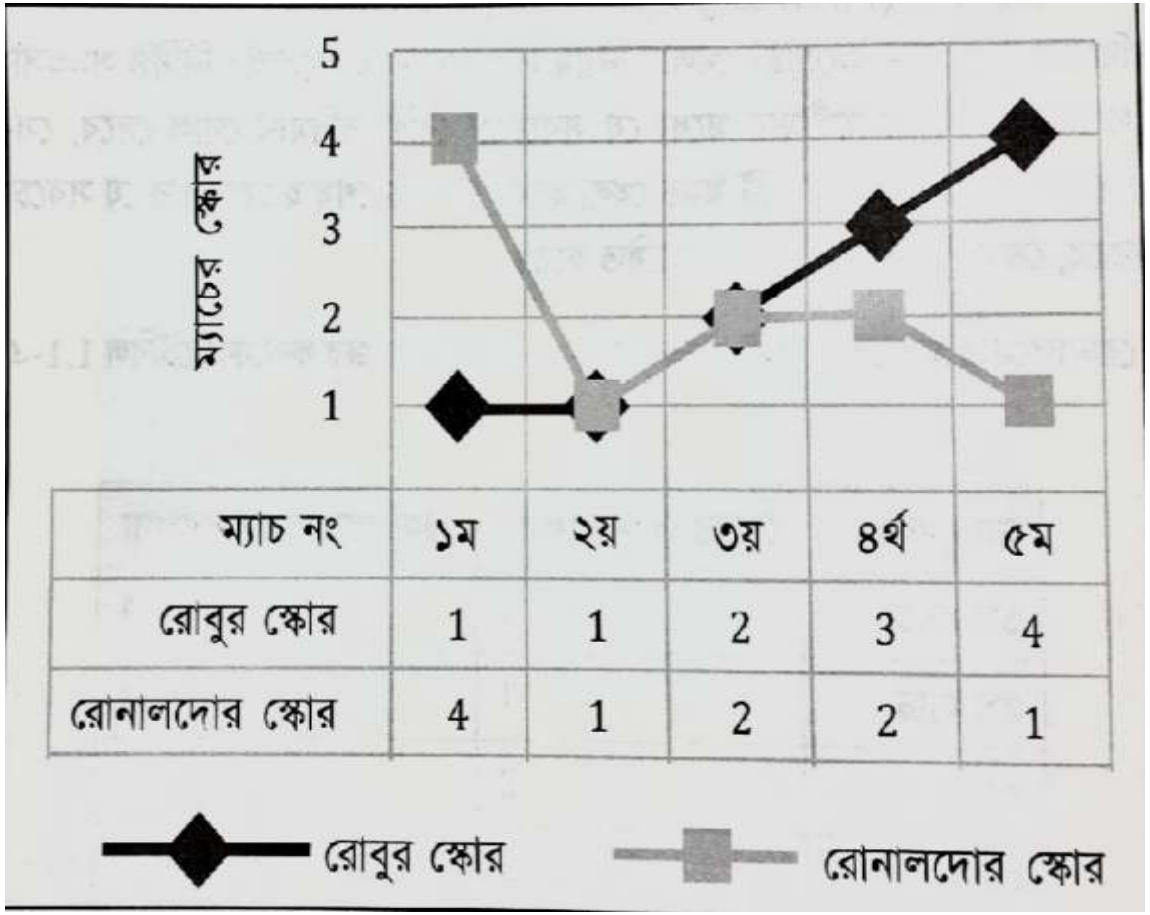
Learning in ML

- A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P**, if its performance at tasks in **T**, as measured by **P**, improves with experience **E.**

- **Example:** playing checkers.
  E = the experience of playing many games of checkers
  T = the task of playing checkers.
  P = the probability that the program will win the next game

MACHINE LEARNING

ম্যাচ নম্বর	রোবুর গোলসংখ্যা	রোনালদোর গোলসংখ্যা
১ম ম্যাচ	1	4
২য় ম্যাচ	1	1
৩য় ম্যাচ	2	2
৪র্থ ম্যাচ	3	2
৫ম ম্যাচ	4	1

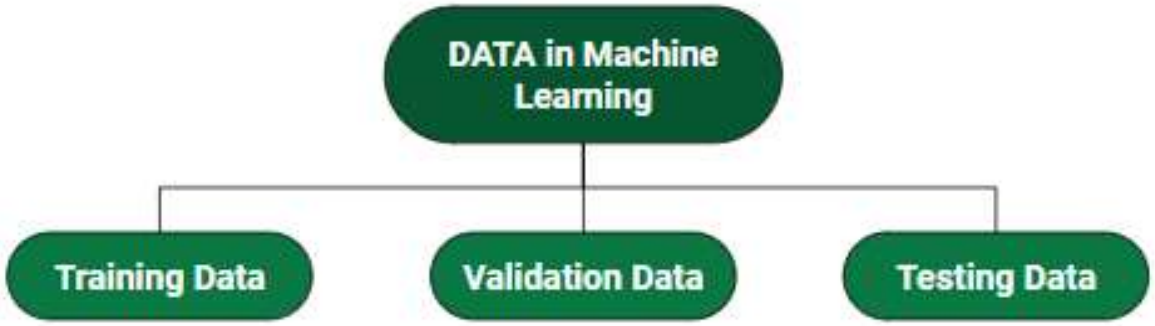


DATA

- **Data:** In computing, data is information that has been translated into a form that is efficient for movement or processing. It can be unprocessed fact, value, text, sound, or picture that is not being interpreted and analyzed.
- **Information:** Information is processed, organized and structured data. It provides context for data and enables decision making process.
- Example: 100, 212, 0, 32-**Data**; the freezing and boiling points of water in Fahrenheit and Celsius-**Information**.
- **Knowledge:** Combination of information, experiences, learning and insights.

DATA

- **Training Data:** The part of data we use to train our model. This is the data that your model actually sees(both input and output) and learns from.
- **Validation Data:** The part of data that is used to do a frequent evaluation of the model, fit on the training dataset along with improving involved hyper parameters (initially set parameters before the model begins learning). This data plays its part when the model is actually training.
- **Testing Data:** Once our model is completely trained, testing data provides an unbiased evaluation. When we feed in the inputs of testing data, our model will predict some values(without seeing actual output). After prediction, we evaluate our model by comparing it with the actual output present in the testing data.



FEATURE

- **Feature:** A Feature is a measurable property of the object you're trying to analyze. In datasets, features appear as columns.
- Each feature or column, represents a measureable piece of data that can be used for analysis: Name, age, sex, fare and so on. Features are also sometimes referred to as "variables" or attributes.

	A	B	C	D	E	F	G	H	I	J	K	L
1	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
2	1	0	3	Braund, Mr. (	male	22	1	0	A/5 21171	7.25		S
3	2	1	1	Cumings, Mr	female	38	1	0	PC 17599	71.2833	C85	C
4	3	1	3	Heikkinen, M	female	26	0	0	STON/O2. 31	7.925		S
5	4	1	1	Futrelle, Mrs	female	35	1	0	113803	53.1	C123	S
6	5	0	3	Allen, Mr. Wi	male	35	0	0	373450	8.05		S
7	6	0	3	Moran, Mr. J	male		0	0	330877	8.4583		Q

FEATURE

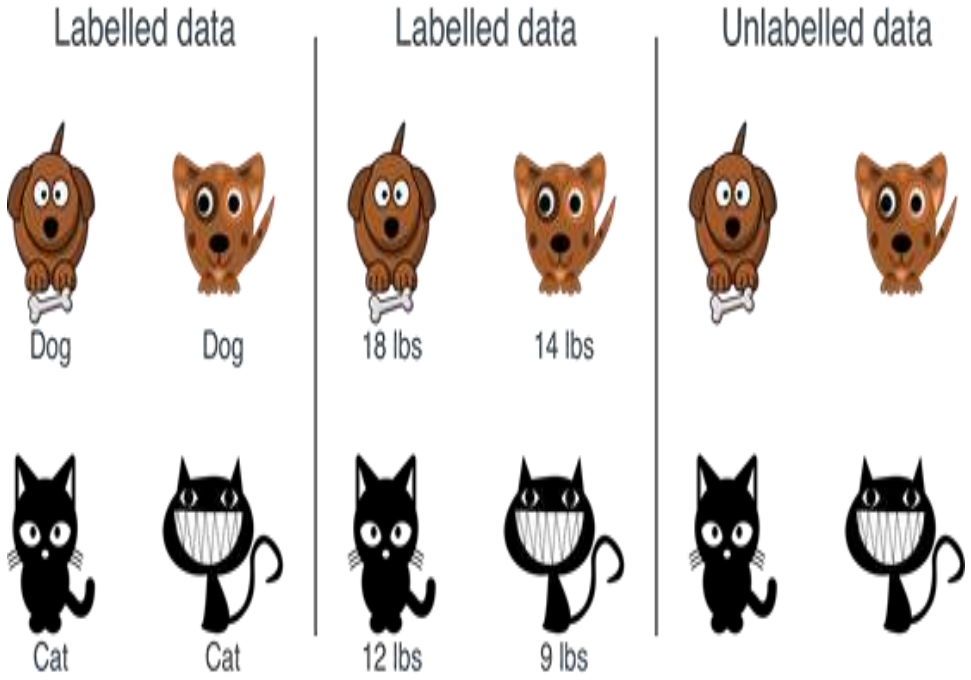
ছবি নম্বর	লেজের দৈর্ঘ্য (সে.মি.)	গলার দৈর্ঘ্য (সে.মি.)	শিঙের মতো বস্তু কি আছে?	প্রাণীটি কি জিরাফ?
1	5	8	হ্যাঁ	হ্যাঁ
2	2	3	না	না
3	1	2	না	না
4	0	2	না	না
5	5.5	7.5	হ্যাঁ	হ্যাঁ

LABEL & UNLABEL DATA

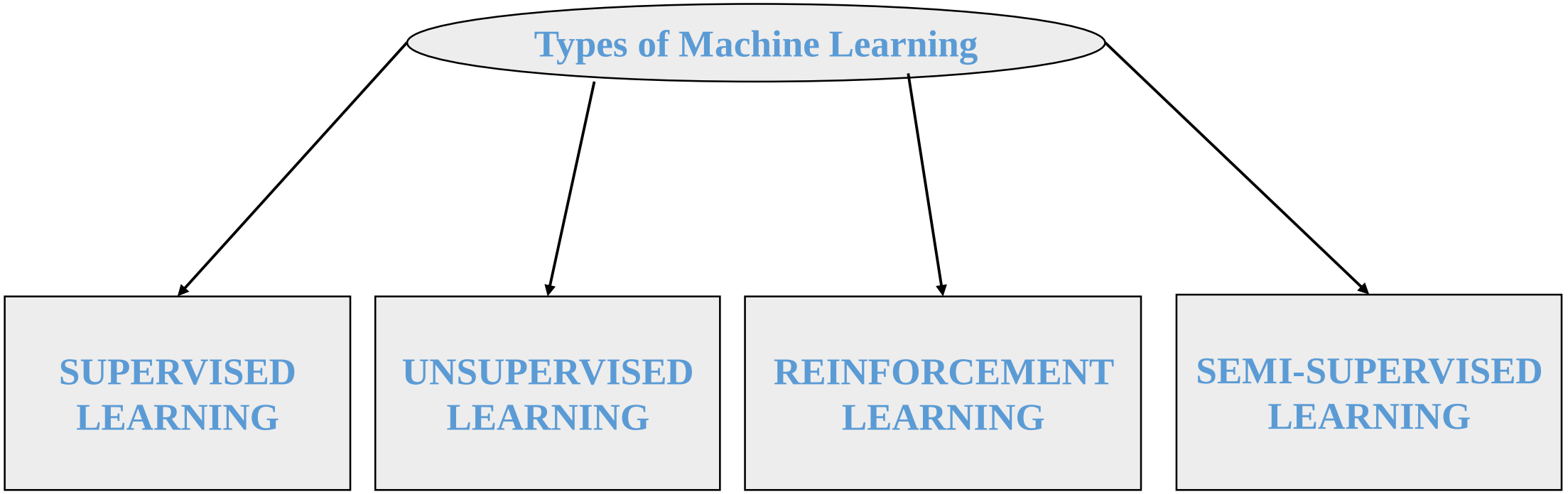
- **Labeled data:** Data that comes with a label.

- **Unlabeled data:** Data that comes without a label.

- The set of algorithms in which we use a labeled dataset is called Supervised Learning. The set of algorithms in which we use an unlabeled dataset, is called Unsupervised Learning.



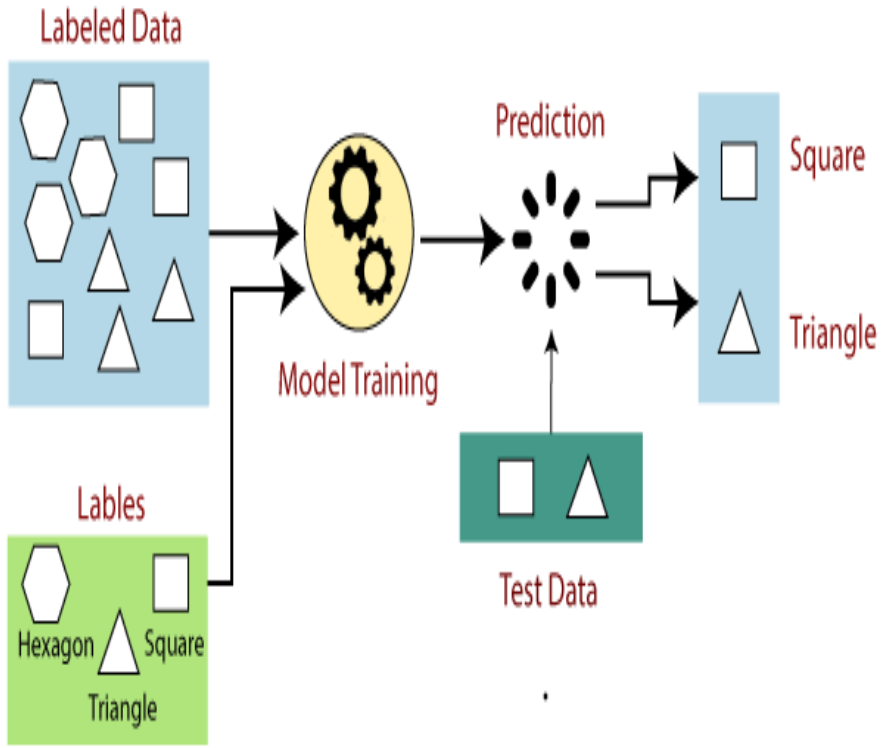
MACHINE LEARNING



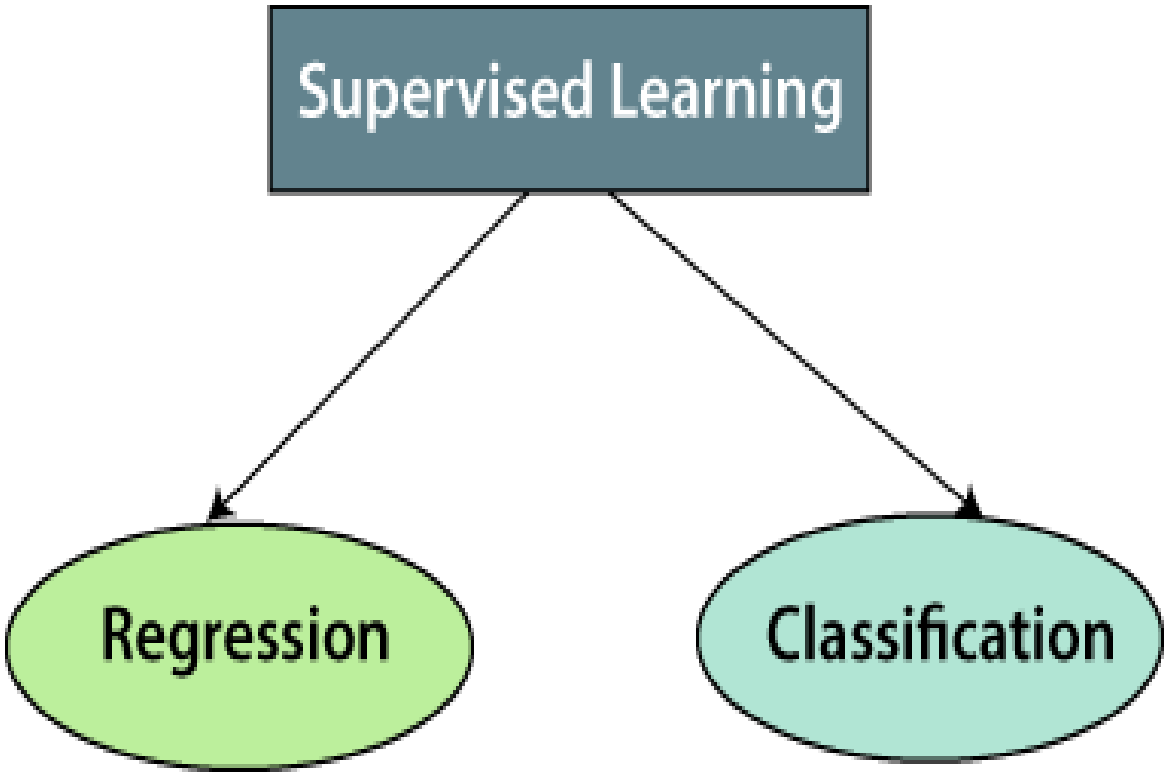
SUPERVISED LEARNING

- **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.

- Supervised Learning is a process of providing input data as well as correct output data to the machine learning model. The aim of a supervised learning algorithm is to find a mapping function to map the input variable(x) with the output variable(y).

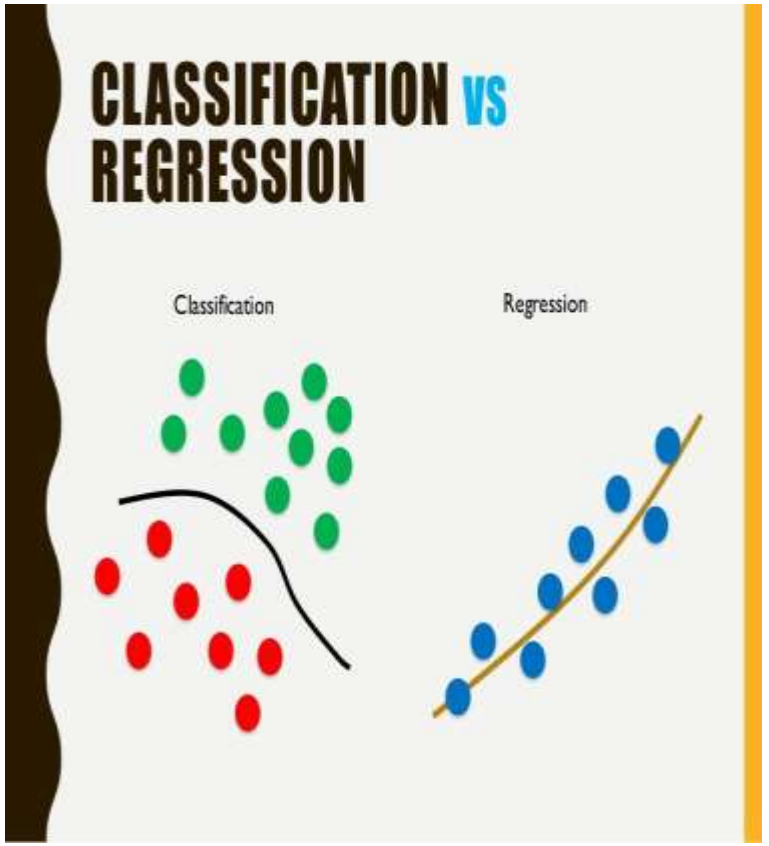


SUPERVISED LEARNING

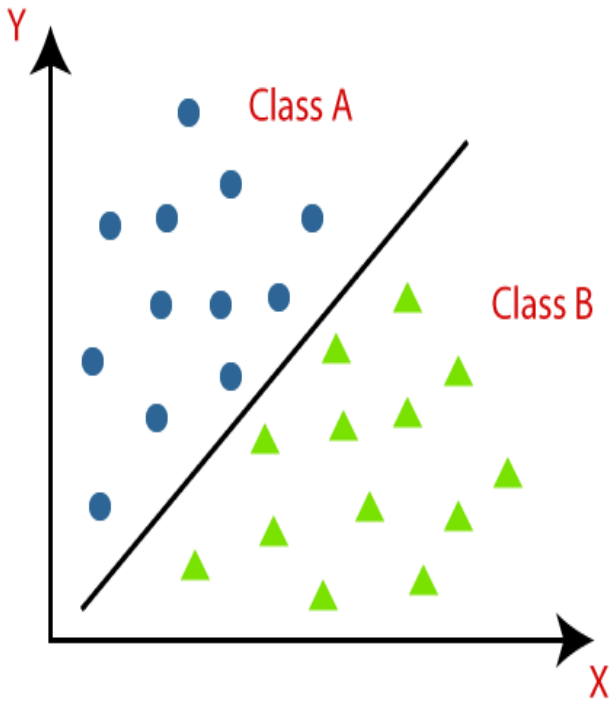


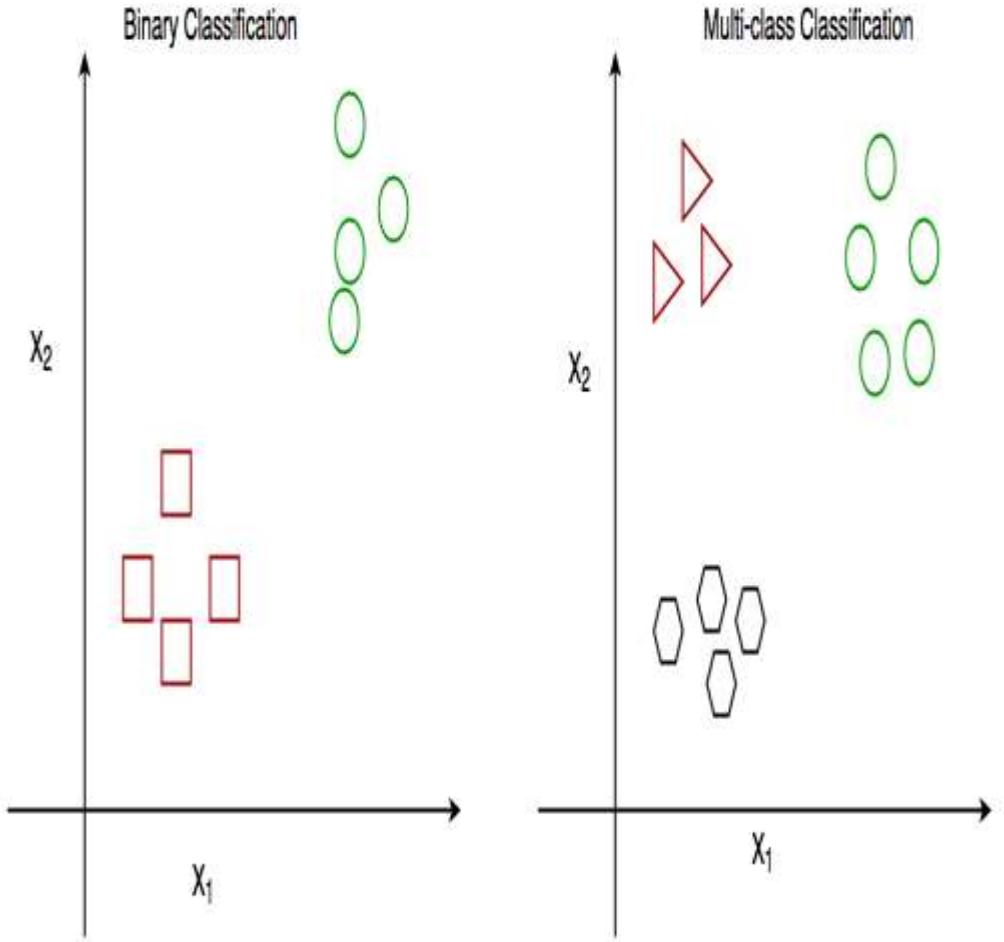
SUPERVISED LEARNING

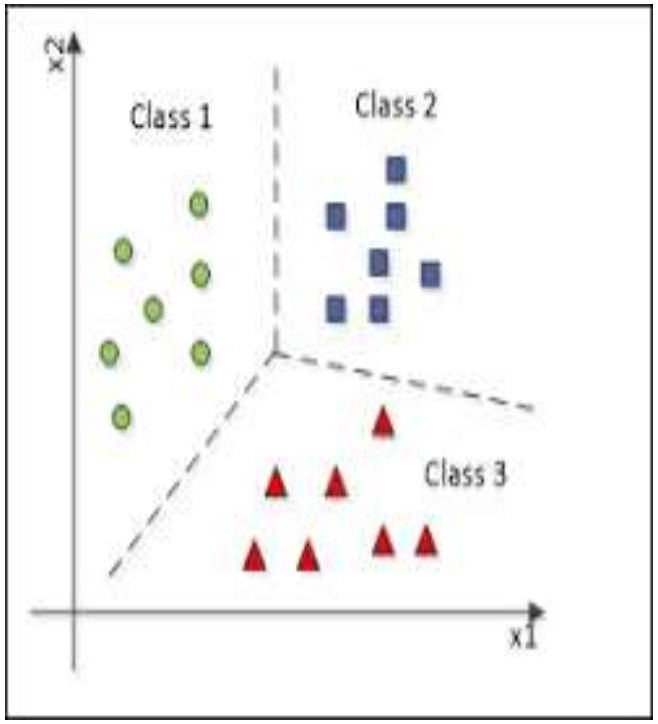
- **Classification:** Classification algorithms are used when the output variable is categorical, which means there are two classes such as Yes-No, Male-Female, True-False etc.
- **Regression:** Regression algorithms are used if there is a relationship between the input variable and the output variable. It is used for the prediction of continuous variables, such as weather forecasting, astronomical analysis, social network analysis etc.



CLASSIFICATION

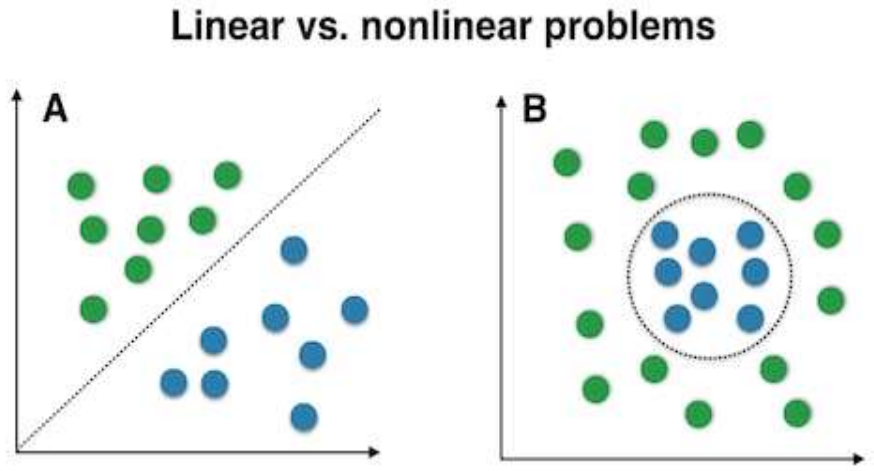




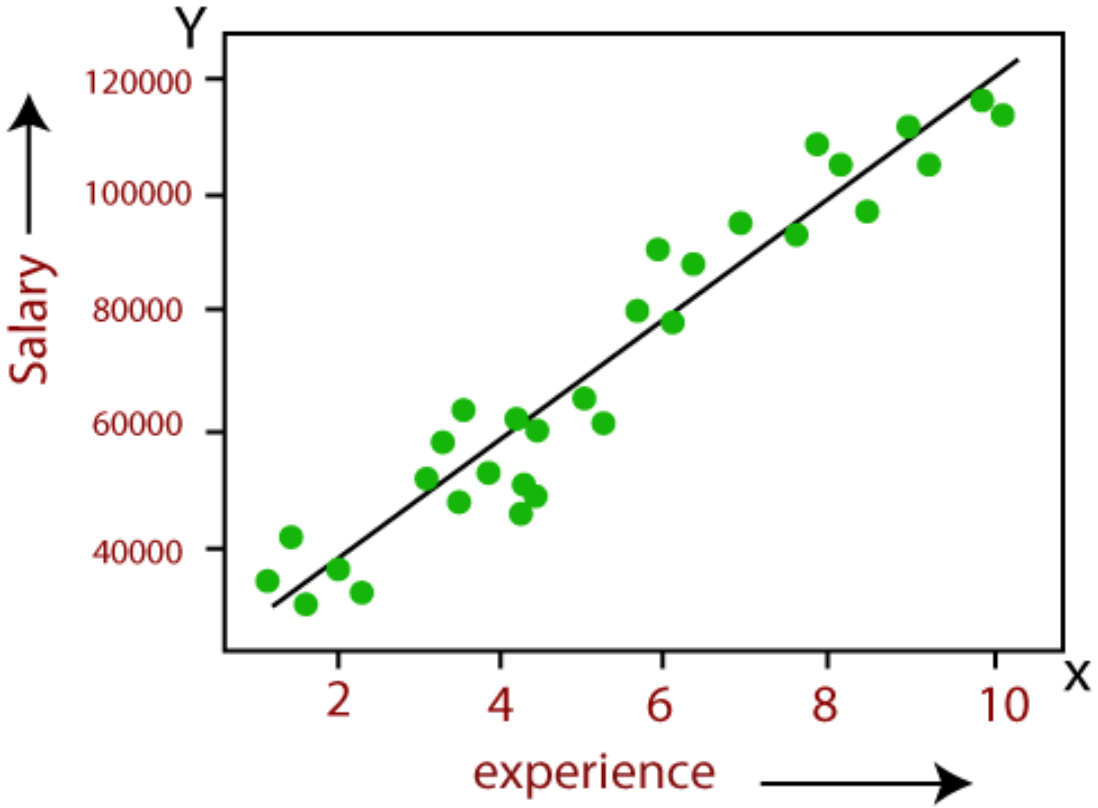


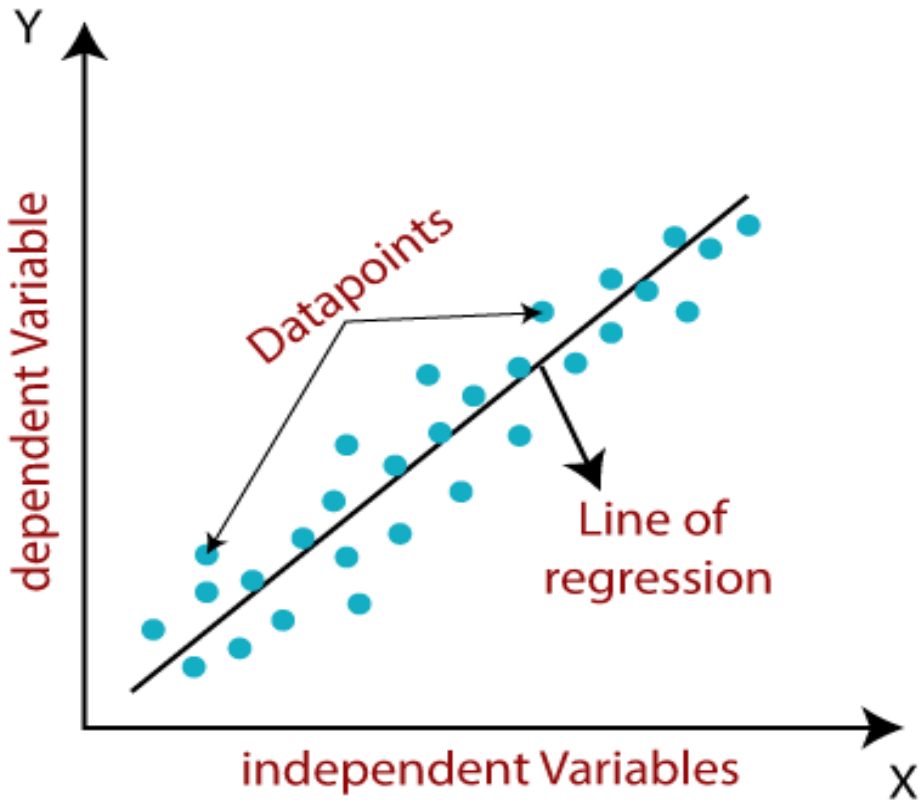
CLASSIFICATION

- Logistic Regression (**Linear**)
- SVM (Support Vector Machine) (**Linear**)
- Naïve Bayes (**Non-Linear**)
- KNN-(K-Nearest Neighbors) (**Non-Linear**)
- Decision Tree (**Non-Linear**)
- Random Forest (**Non-Linear**)
- Kernel SVM (Support Vector Machine) (**Non-Linear**)
- NN (Neural Network) (**Non-Linear**)



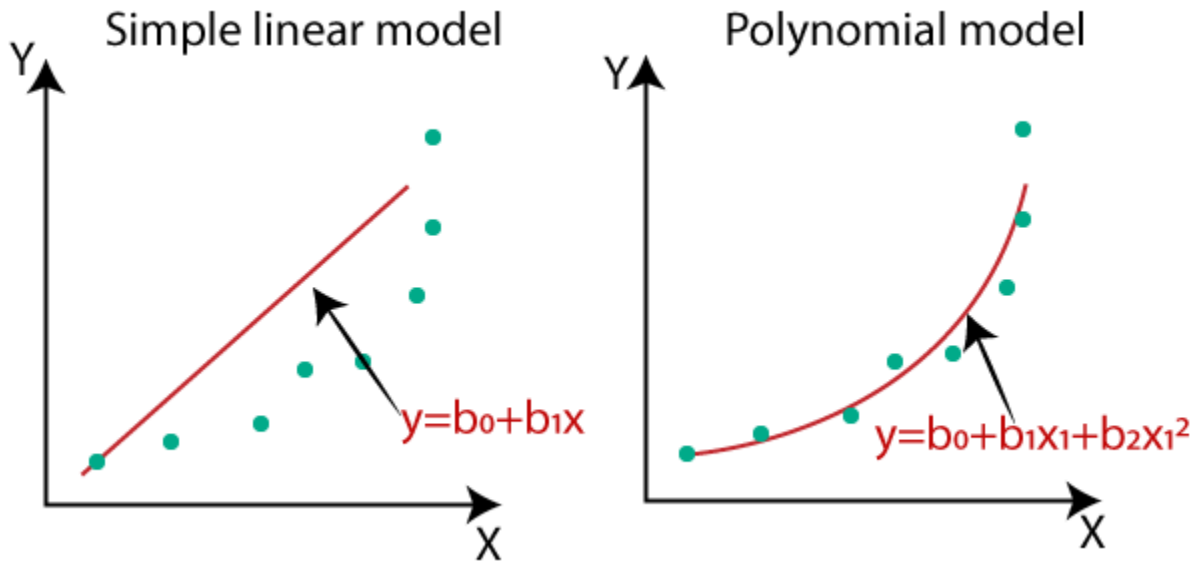
REGRESSION


Y
120000
100000
80000
60000
40000
Salary
2
4
6
8
10
X
experience


Y
dependent Variable
Datapoints
Line of
regression
independent Variables
X

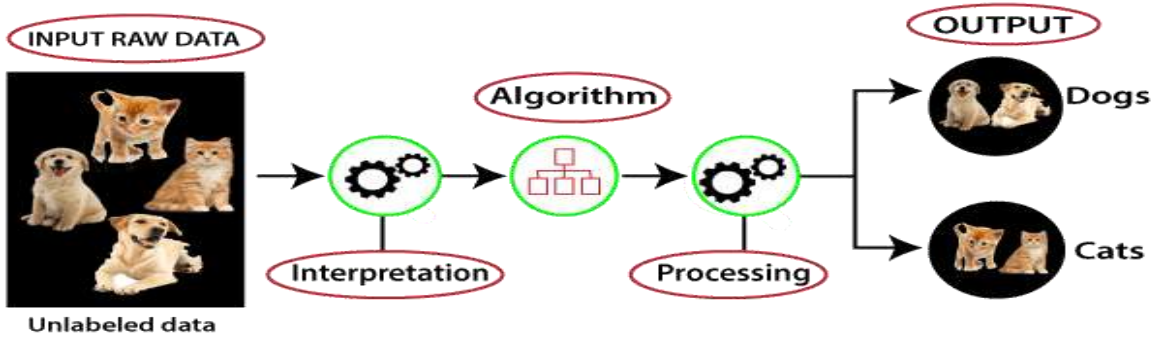
REGRESSION

- Linear Regression
- Ridge Regression
- Lasso Regression
- Polynomial Regression
- Elastic Net Regression

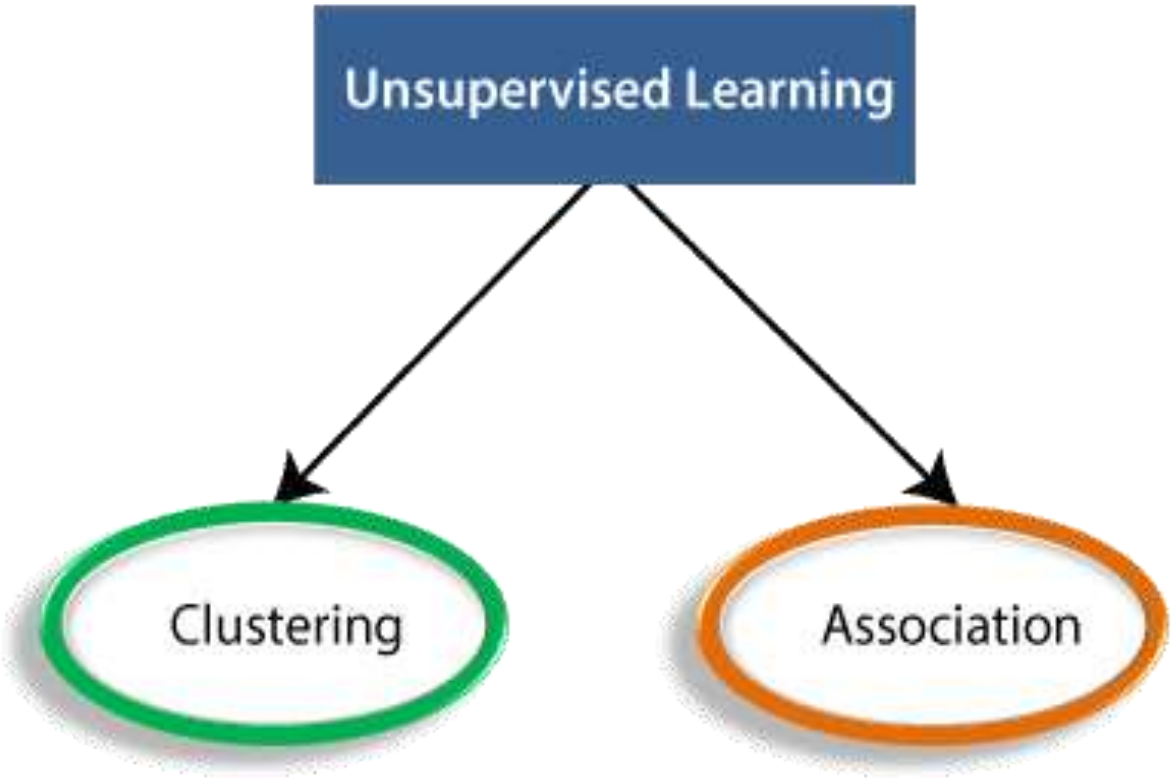


UNSUPERVISED LEARNING

- **Unsupervised Learning:** As the name suggests, unsupervised learning is a machine learning technique in which models are not supervised using training dataset. Instead, models itself find the hidden patterns and insights from the given data. It can be compared to learning which takes place in the human brain while learning new things. Unsupervised learning cannot be directly applied to a regression or classification problem because unlike supervised learning, we have the input data but no corresponding output 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.

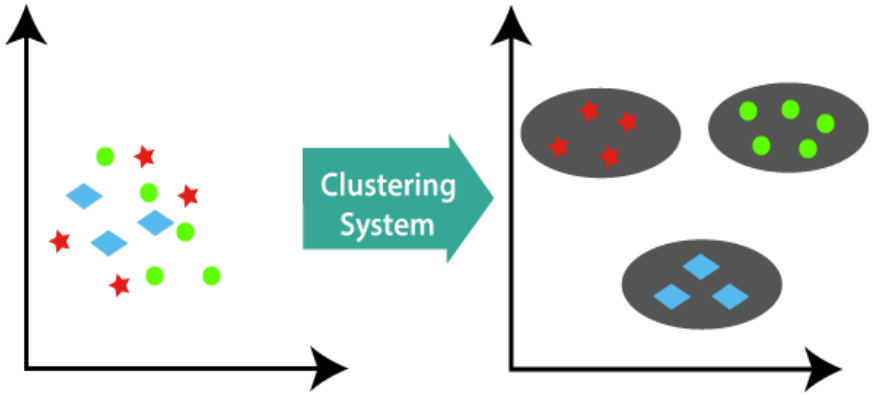


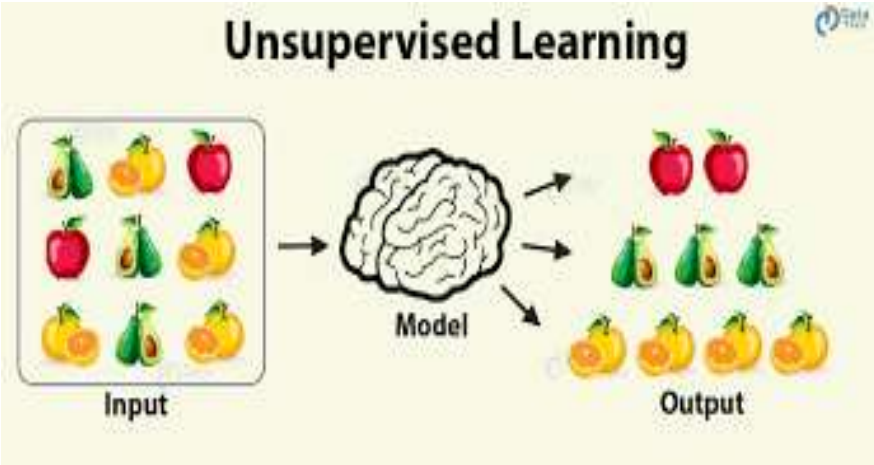
UNSUPERVISED LEARNING



CLUSTERING

- **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.





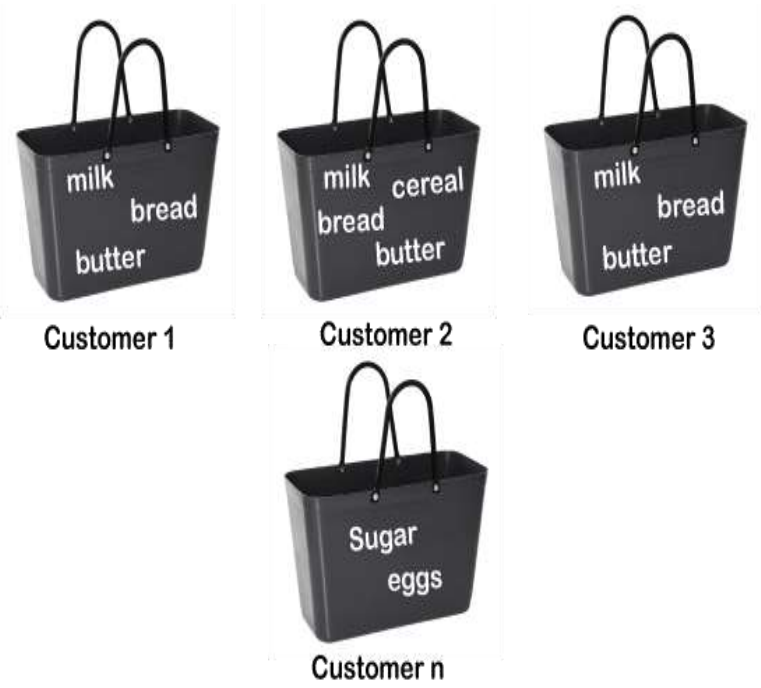
CLUSTERING

- K-means clustering
- DBSCAN(Density Based Spatial Clustering Applications with Noise)

ASSOCIATION RULE LEARNING

- **Association Rule learning:** Association rule learning is a type of unsupervised learning technique that checks for the dependency of one data item on another data item and maps accordingly so that it can be more profitable. Example: Market Basket analysis.

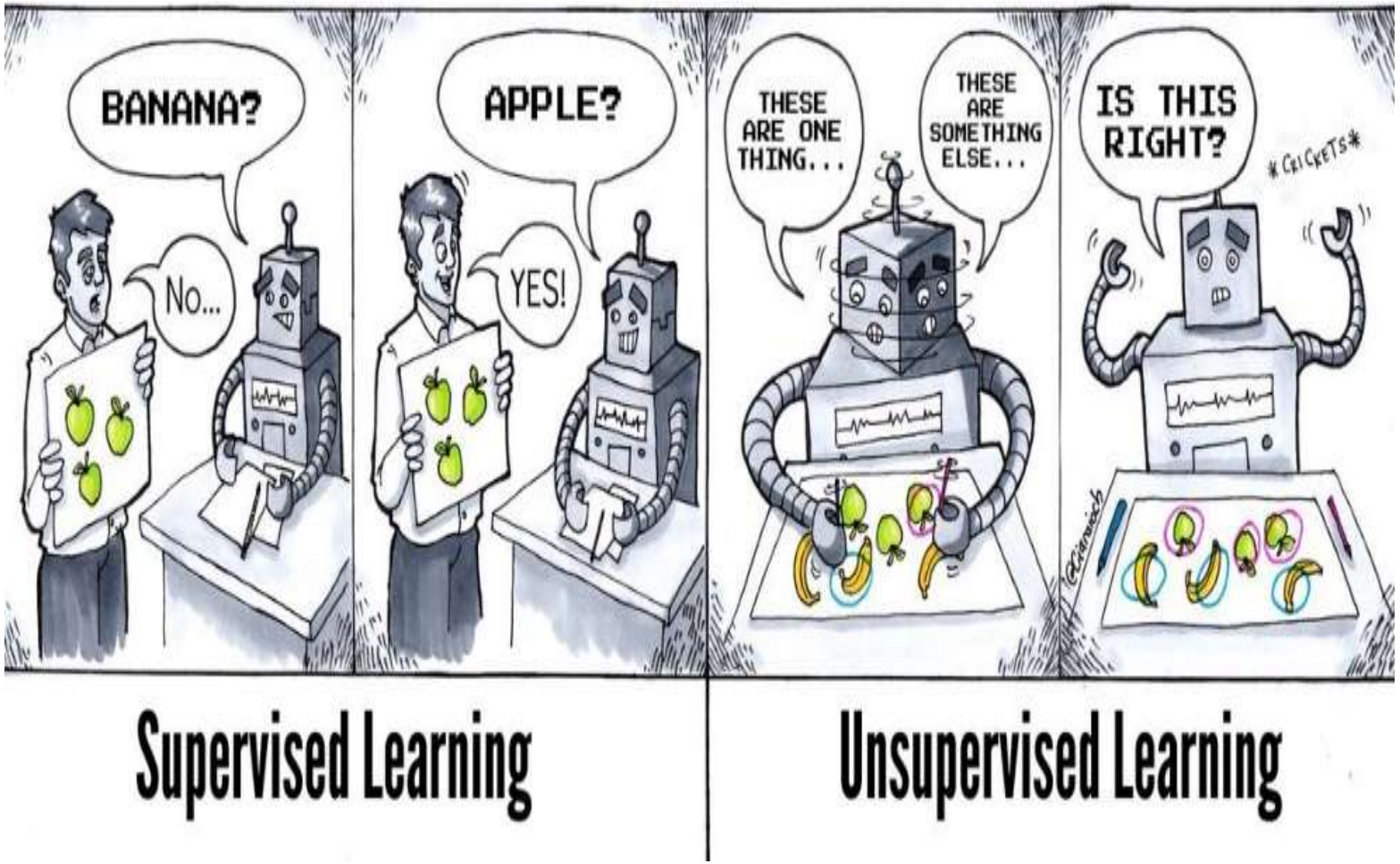
- Here, market basket analysis is a technique used by the various big retailer to discover the association between items. We can understand it by taking an example of a supermarket, as in supermarket, all products that are purchased together are put together.
- For example, if a customer buys bread, he most likely can also buy butter, eggs, or milk, so these products are stored within a shelf or mostly nearby.



ASSOCIATION RULE LEARNING

- Apriori
- F-P Growth algorithm

EXTRAS

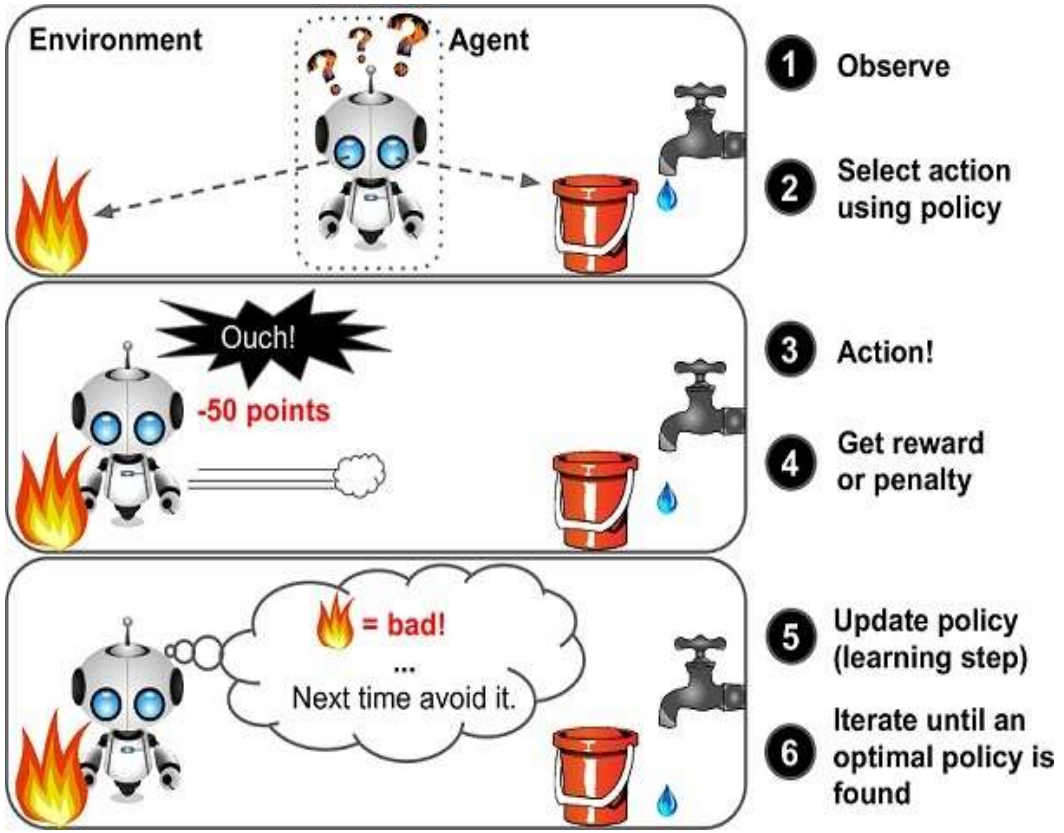

BANANA?
No...
APPLE?
YES!
THESE ARE ONE THING...
THESE ARE SOMETHING ELSE...
IS THIS RIGHT?
CRICKETS
Supervised Learning
Unsupervised Learning

REINFORCEMENT LEARNING

- **Reinforcement Learning:** Reinforcement Learning is a type of machine learning method where an intelligent agent (computer program) interacts with the environment and learns to act within that.
- It is a core part of Artificial Intelligence, and all AI agent works on the concept of reinforcement learning. Here we do not need to pre-program the agent, as it learns from its own experience without any human intervention.
- In reinforcement learning, the agent learns automatically using feedbacks without any labeled data, unlike supervised learning.
- Since there is no labeled data, so the agent is bound to learn by its experience only.
- RL solves a specific type of problem where decision making is sequential, and the goal is long term such as game-playing, robotics etc.
- The agent interacts with the environment and explores it by itself. The primary goal of an agent in RL is to improve the performance by getting the maximum positive rewards.

REINFORCEMENT LEARNING

- Four main elements of RL:
- Policy
- Reward Signal
- Value Function
- Model of the environment

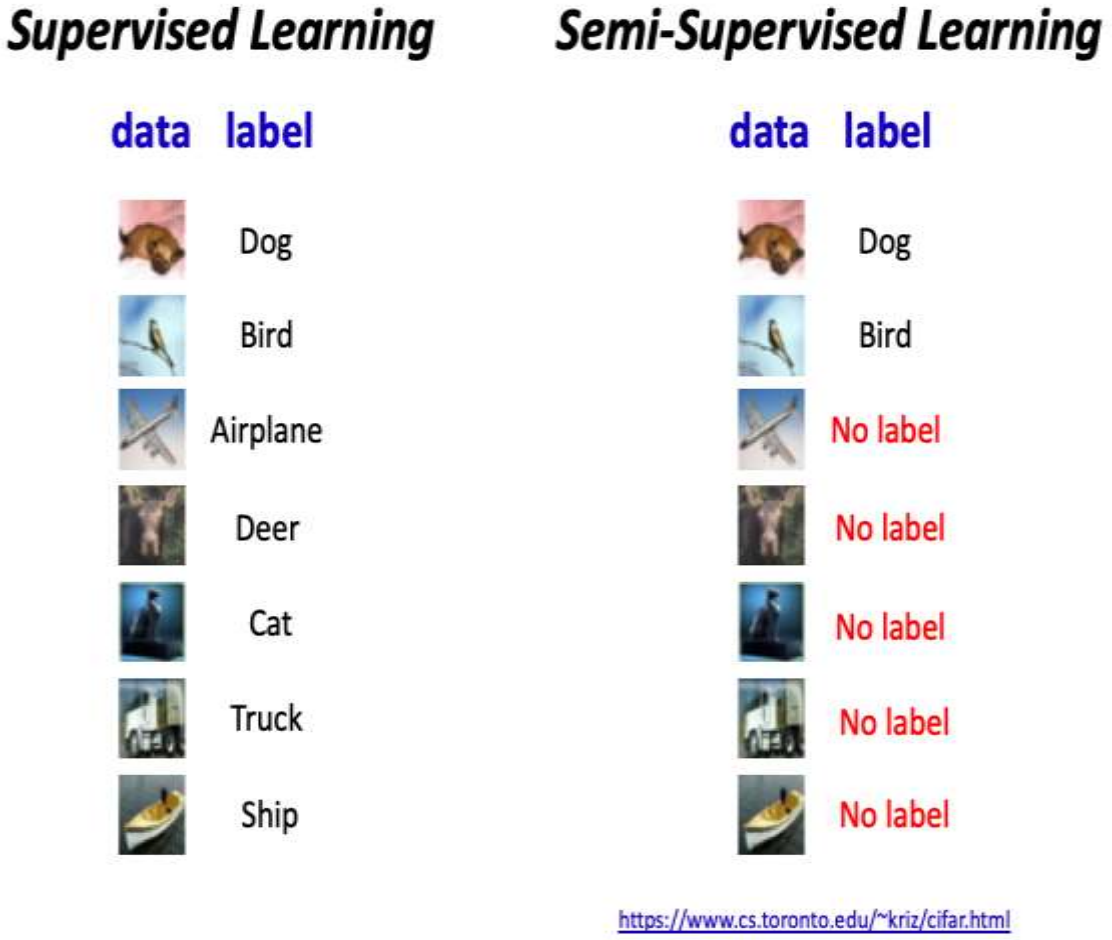


REINFORCEMENT LEARNING

- **Positive RL**
- **Negative RL**

SEMI SUPERVISED LEARNING

- **Semi Supervised Learning:** Semi-supervised learning is the type of machine learning that uses a combination of a small amount of labeled data and a large amount of unlabeled data to train models.



SEMI SUPERVISED LEARNING

- **Speech Analysis:** It is the most classic example of semi-supervised learning applications. Since, labelling the audio data is the most impassable task that requires many human resources, this problem can be naturally overcome with the help of applying SSL in a semi-supervised learning model.
- **Web content classification:** However, this is very critical and impossible to label each page on the internet because it needs mode human intervention. Still this problem can be reduced through semi-supervised learning algorithms. Further, Google also uses semi-supervised learning algorithms to rank a webpage for a given query.
- **Protein sequence classification:** DNA strands are larger, they require active human intervention. So, the rise of the semi-supervised model has been proximate in this field.
- **Text document classifier:** As we know, it would be very unfeasible to find a large amount of labeled text data, so semi-supervised learning is an ideal model to overcome this.

EXTRAS

PROTEIN SEQUENCE CLASSIFICATION