



Introduction to Machine Learning

(Supervised vs Unsupervised Learning)

Dr. Virendra Singh Kushwah

Assistant Professor Grade-II

School of Computing Science and Engineering

Virendra.Kushwah@vitbhopal.ac.in

7415869616

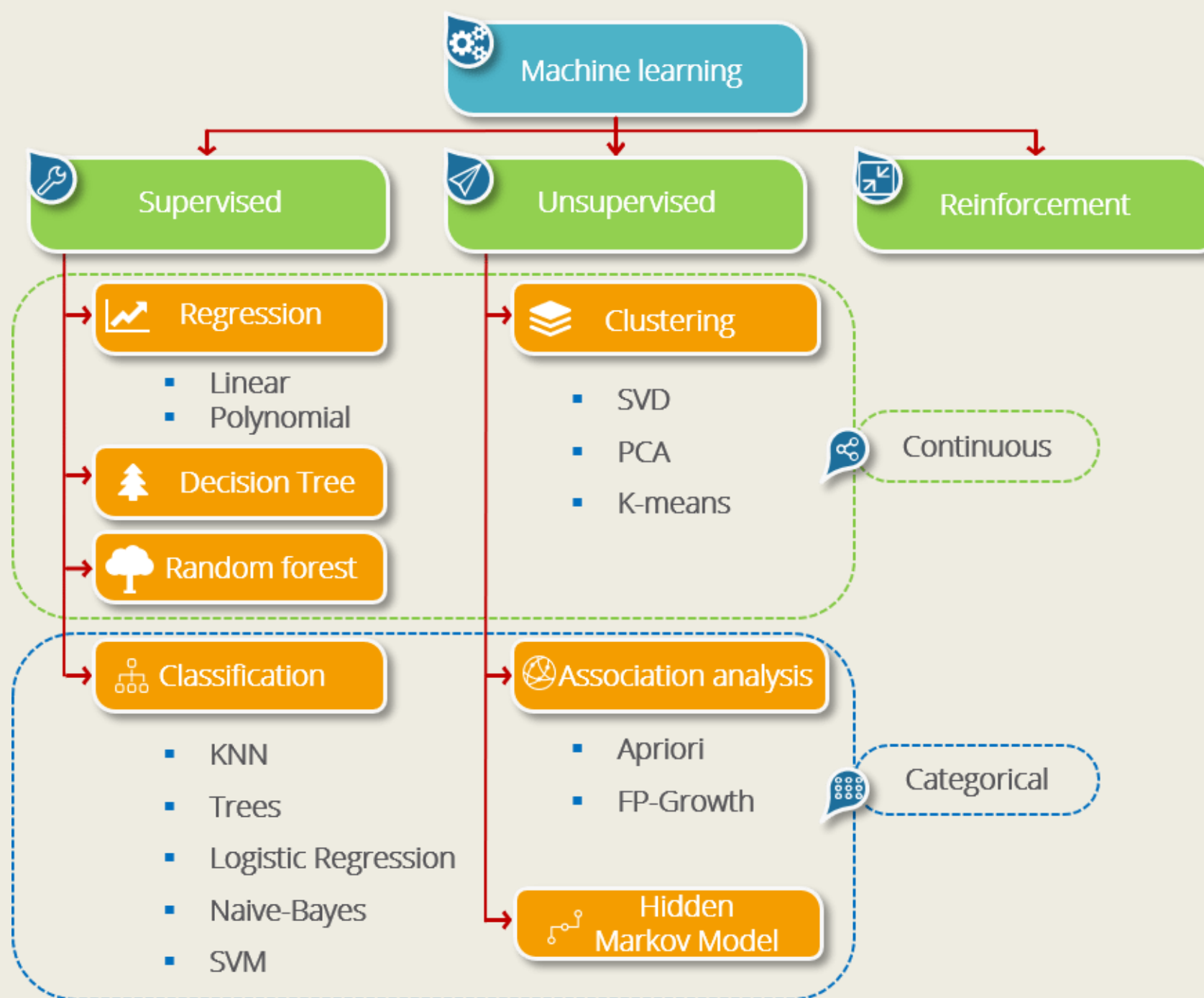


Lecture-2

- Need For Machine Learning
- Machine Learning Definitions
- Machine Learning Process
- How does Machine Learning Work?

Machine Learning Types

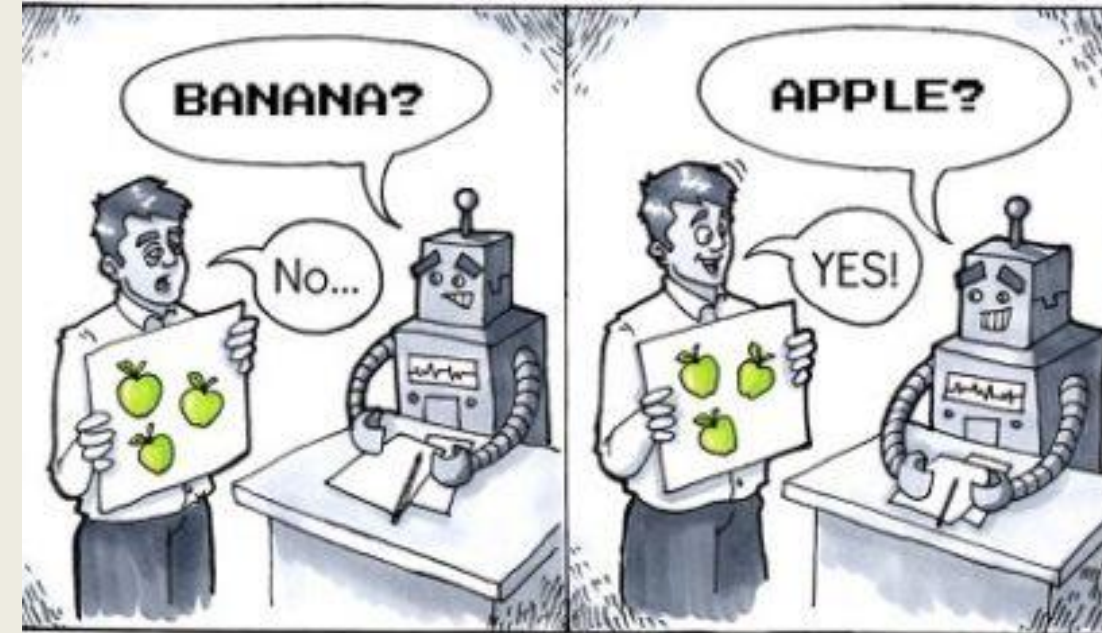
- A machine can learn to solve a problem by following any one of the following three approaches. These are the ways in which a machine can learn:
 1. Supervised Learning
 2. Unsupervised Learning
 3. Reinforcement Learning



What is Supervised Learning?

- Supervised Learning is the process of making an algorithm to learn to map an input to a particular output. This is achieved using the labelled datasets that you have collected. If the mapping is correct, the algorithm has successfully learned. Else, you make the necessary changes to the algorithm so that it can learn correctly. Supervised Learning algorithms can help make predictions for new unseen data that we obtain later in the future.

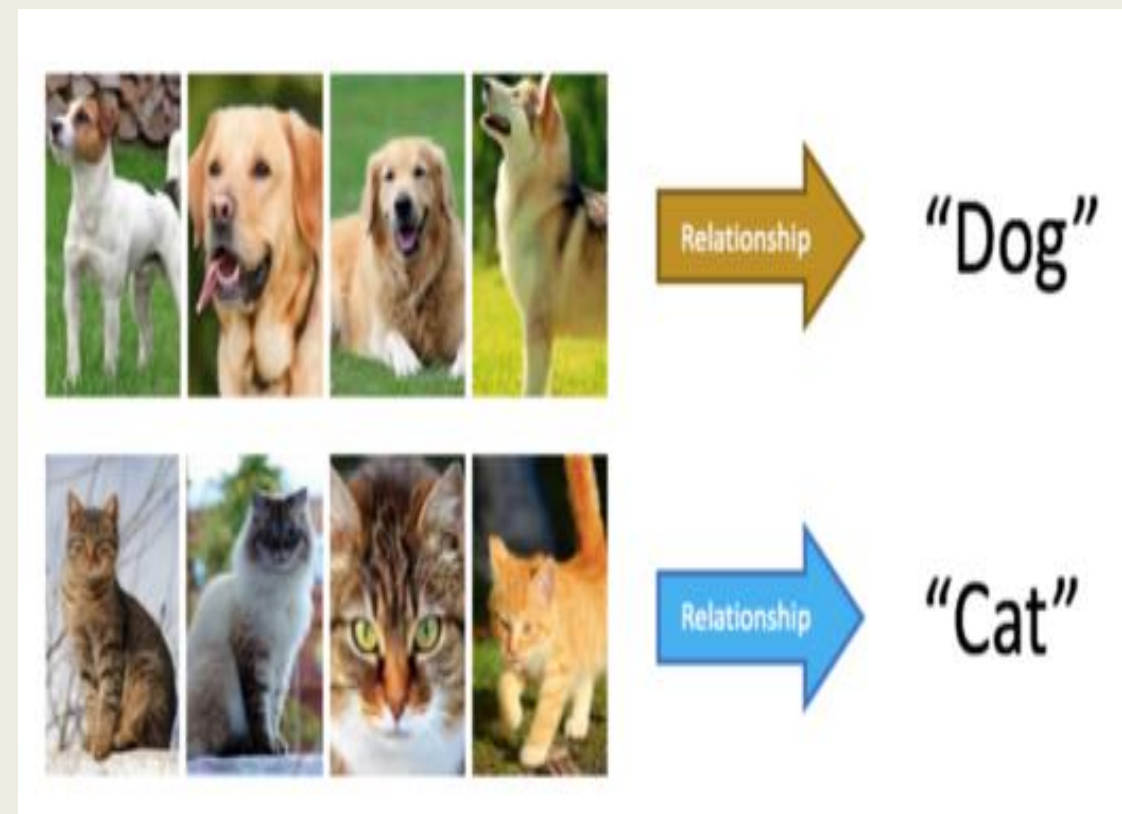
- This is similar to a teacher-student scenario. There is a teacher who guides the student to learn from books and other materials. The student is then tested and if correct, the student passes. Else, the teacher tunes the student and makes the student learn from the mistakes that he or she had made in the past. That is the basic principle of Supervised Learning.



Supervised Learning

Example of Supervised Learning

- Suppose you have a niece who has just turned 2 years old and is learning to speak. She knows the words, Papa and Mummy, as her parents have taught her how she needs to call them. You want to teach her what a dog and a cat is. So what do you do? You either show her videos of dogs and cats or you bring a dog and a cat and show them to her in real-life so that she can understand how they are different.



Observations

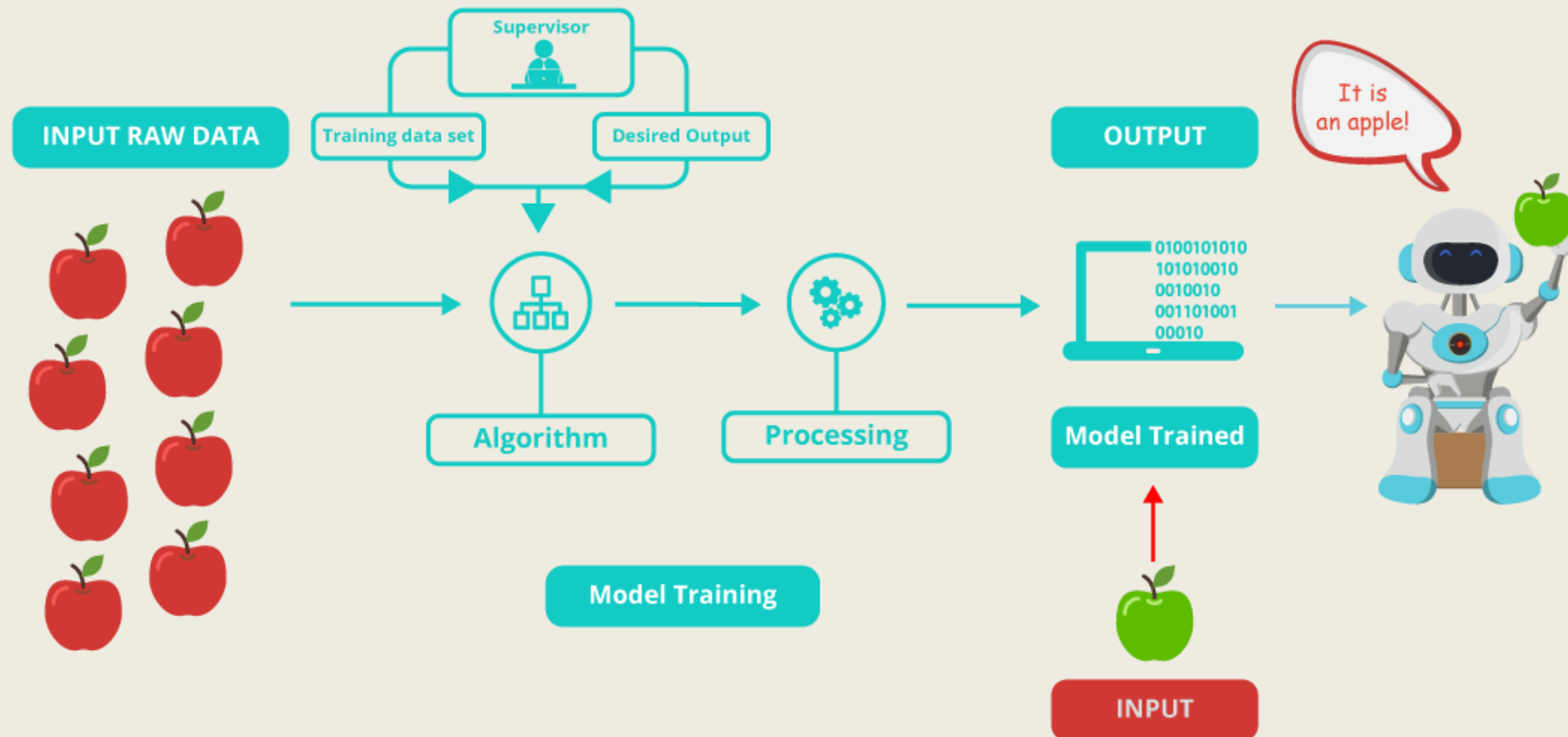
- Now there are certain things you tell her so that she understands the differences between the 2 animals.
- Dogs and cats both have 4 legs and a tail.
- Dogs come in small to large sizes. Cats, on the other hand, are always small.
- Dogs have a long mouth while cats have smaller mouths.
- Dogs bark while cats meow.
- Different dogs have different ears while cats have almost the same kind of ears.

- Now you take your niece back home and show her pictures of different dogs and cats. If she is able to differentiate between the dog and cat, you have successfully taught her.
- So what happened here? You were there to guide her to the goal of differentiating between a dog and a cat. You taught her every difference there is between a dog and a cat. You then tested her if she was able to learn. If she was able to learn, she called the dog as a dog and a cat as a cat. If not, you taught her more and were able to teach her. You acted as the supervisor and your niece acted as the algorithm that had to learn. You even knew what was a dog and what was a cat. Making sure that she was learning the correct thing. That is the principle that Supervised Learning follows.

Why is it Important?

- Learning gives the algorithm experience which can be used to output the predictions for new unseen data
- Experience also helps in optimizing the performance of the algorithm
- Real-world computations can also be taken care of by the Supervised Learning algorithms

Another Examples

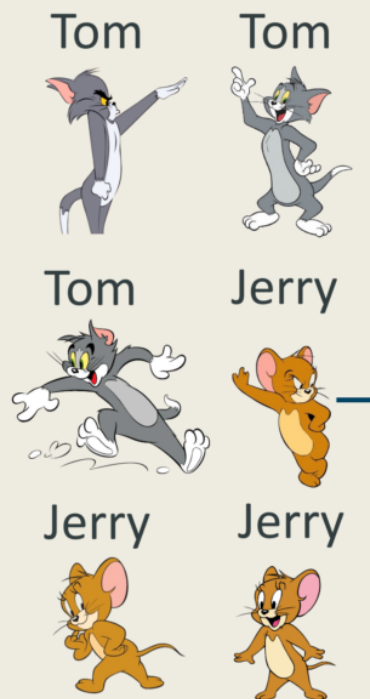


Supervised Learning

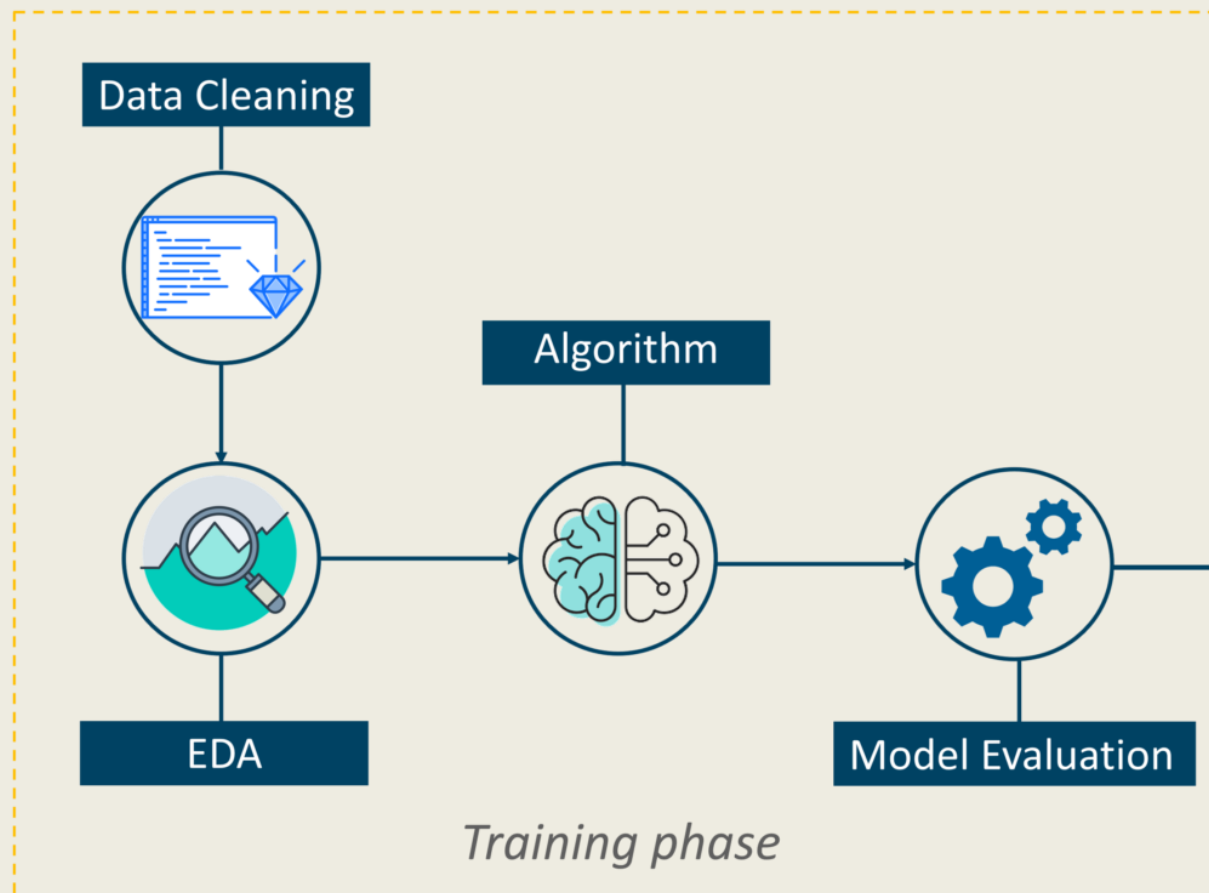
- *“Supervised learning is a technique in which we teach or train the machine using data which is well labeled.”*
- Supervised Learning is the one, where you can consider the learning is guided by a teacher. We have a dataset which acts as a teacher and its role is to train the model or the machine. Once the model gets trained it can start making a prediction or decision when new data is given to it.

- To understand Supervised Learning let's consider an analogy. As kids we all needed guidance to solve math problems. Our teachers helped us understand what addition is and how it is done. Similarly, you can think of supervised learning as a type of Machine Learning that involves a guide. The labeled data set is the teacher that will train you to understand patterns in the data. The labeled data set is nothing but the training data set.

Labelled Data



Known output



Labelled Output



Class 'Jerry'



Class 'Tom'

Types of Supervised Learning

- Supervised Learning has been broadly classified into 2 types.
 - Regression
 - Classification
- Regression is the kind of Supervised Learning that learns from the Labelled Datasets and is then able to predict a continuous-valued output for the new data given to the algorithm. It is used whenever the output required is a number such as money or height etc.

What is Unsupervised Learning?

- Unsupervised Learning can be thought of as self-learning where the algorithm can find previously unknown patterns in datasets that do not have any sort of labels.
- It helps in modelling probability density functions, finding anomalies in the data, and much more.
- To give you a simple example, think of a student who has textbooks and all the required material to study but has no teacher to guide. Ultimately, the student will have to learn by himself or herself to pass the exams. This sort of self-learning is what we have scaled into Unsupervised Learning for machines.

Example of Unsupervised Learning

- Suppose you have never watched a cricket match in your entire life and you have been invited by your friends to hang out at their house for a match between India and England. You have no idea about what cricket is but just for your friends, you say yes and head over with them. The match starts and you just sit there, blank. Your friends are enjoying the way Virat Kohli plays and want to join in the fun. Here is when you start learning about the game. You analyze the screen and come up with certain conclusions that you can use to understand the game better.



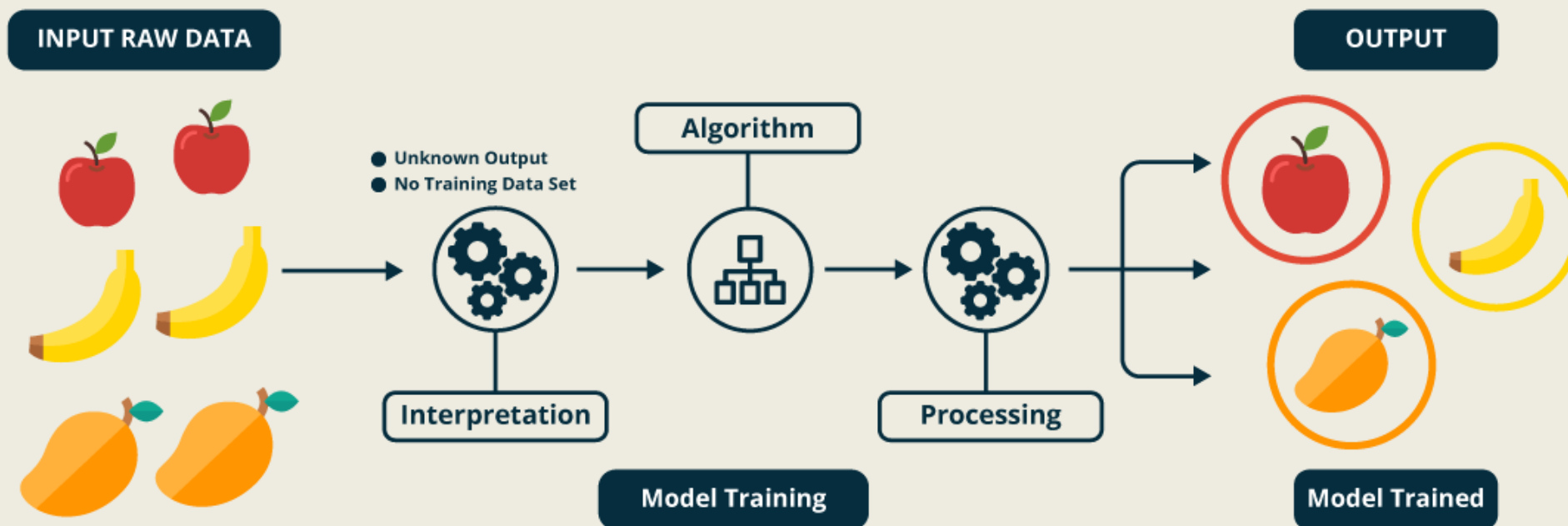
Observations

- There are 2 teams with jerseys of color Blue and Dark Blue. Since Virat Kohli belongs to India and you see the score of India on the screen, you conclude that India has the jersey of Blue which makes England have Dark Blue Jersey.
- There are different types of players on the field. 2 which belong to India have bats in their hand meaning that they are batting. There is someone who runs up and bowls the ball, making him a bowler. There are around 9 players around the field who try to stop the ball from reaching the boundary of the stadium. There is someone behind the wickets and 2 umpires to manage the match.
- If the ball hits the wickets or if the ball is caught by the fielders, the batsman is out and has to walk back.
- Virat Kohli has the number 18 and his name on the back of his jersey and if this player scores a 4 or a 6, you need to cheer.

Why is it important?

- Unsupervised Learning algorithms work on datasets that are unlabeled and find patterns which would previously not be known to us.
- These patterns obtained are helpful if we need to categorize the elements or find an association between them.
- They can also help detect anomalies and defects in the data which can be taken care of by us.

Another Example



Unsupervised Learning

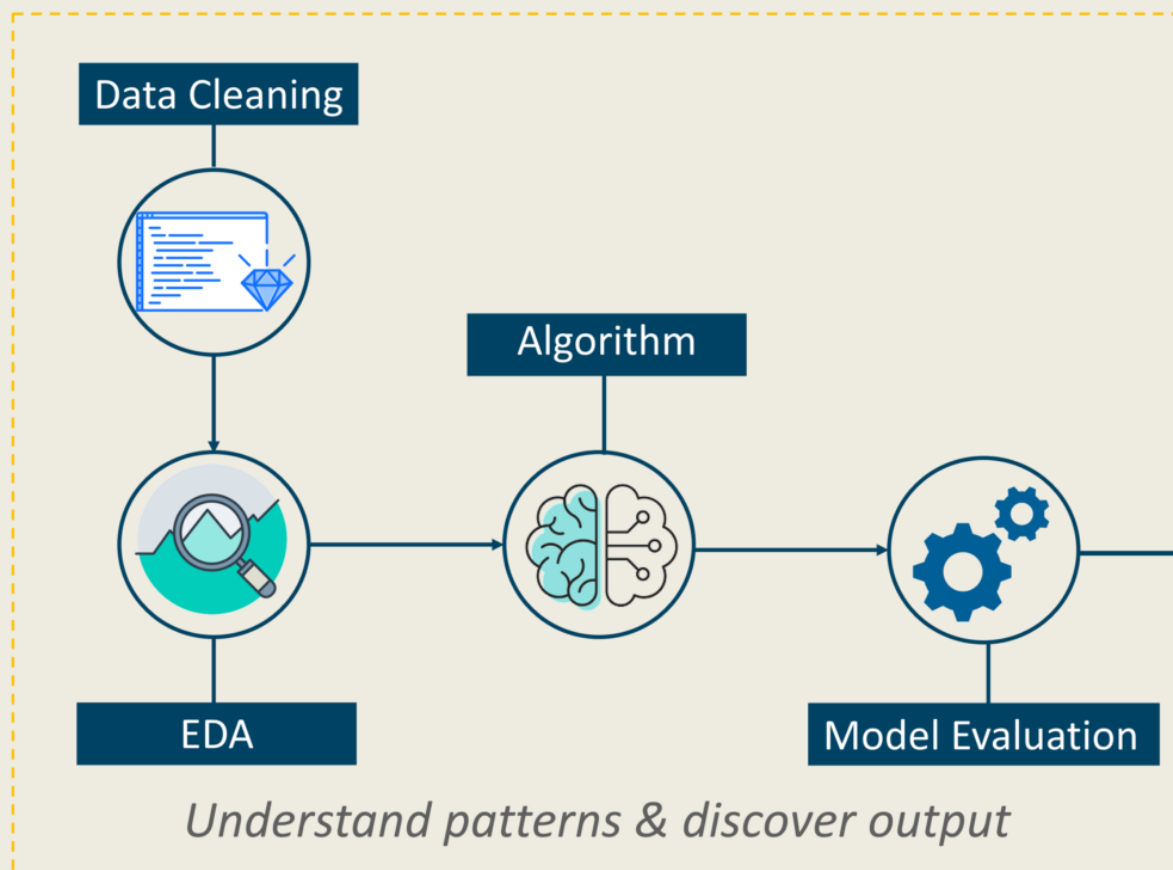
- *“Unsupervised learning involves training by using unlabeled data and allowing the model to act on that information without guidance.”*
- The model learns through observation and finds structures in the data. Once the model is given a dataset, it automatically finds patterns and relationships in the dataset by creating clusters in it. What it cannot do is add labels to the cluster, like it cannot say this a group of apples or mangoes, but it will separate all the apples from mangoes.
- Suppose we presented images of apples, bananas and mangoes to the model, so what it does, based on some patterns and relationships it creates clusters and divides the dataset into those clusters. Now if a new data is fed to the model, it adds it to one of the created clusters.

- Think of unsupervised learning as a smart kid that learns without any guidance. In this type of Machine Learning, the model is not fed with labeled data, as in the model has no clue that ‘this image is Tom and this is Jerry’, it figures out patterns and the differences between Tom and Jerry on its own by taking in tons of data.
- For example, it identifies prominent features of Tom such as pointy ears, bigger size, etc, to understand that this image is of type 1. Similarly, it finds such features in Jerry and knows that this image is of type 2. Therefore, it classifies the images into two different classes without knowing who Tom is or Jerry is.

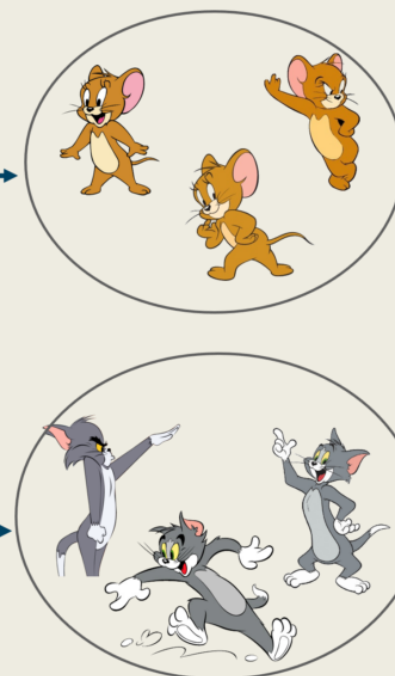
Unlabelled Data



Unknown output



Unlabelled Output



*Clusters formed based on
feature similarity*



Types of Unsupervised Learning

- Unsupervised Learning has been split up majorly into 2 types:
 - Clustering
 - Association

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.”*
- It is the ability of an agent to interact with the environment and find out what is the best outcome. It follows the concept of hit and trial method. The agent is rewarded or penalized with a point for a correct or a wrong answer, and on the basis of the positive reward points gained the model trains itself. And again once trained it gets ready to predict the new data presented to it.
- Reinforcement Learning is mainly used in advanced Machine Learning areas such as self-driving cars, AlphaGo, etc.



VIT[®]
BHOPAL
www.vitbhopal.ac.in