



Unit 1 : Introduction to Machine Learning

Outline

- Overview of Human Learning and Machine Learning
- Types of Machine Learning
- Applications of Machine Learning ,
- Tools and Technology for Machine Learning



What is Human Learning ?



What is Machine Learning ?



What Is Machine Learning?

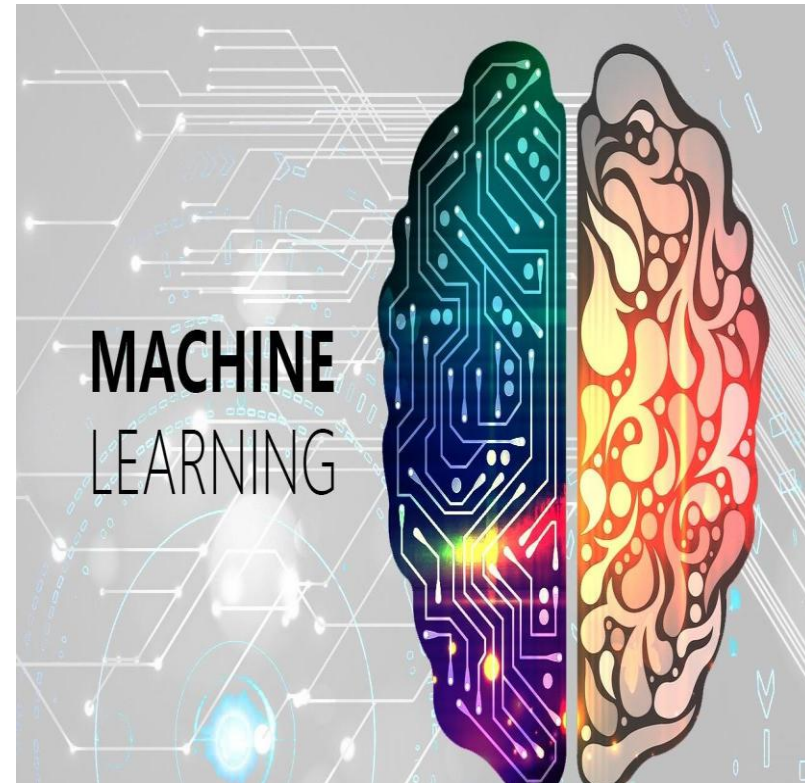
- Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.
- **Machine learning focuses on development of computer programs** that can access data and use it learn for themselves.
- **The primary aim is to allow computers learn automatically** without human intervention or assistance and adjust actions accordingly.



What Is Machine Learning?

Definition:

- 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.

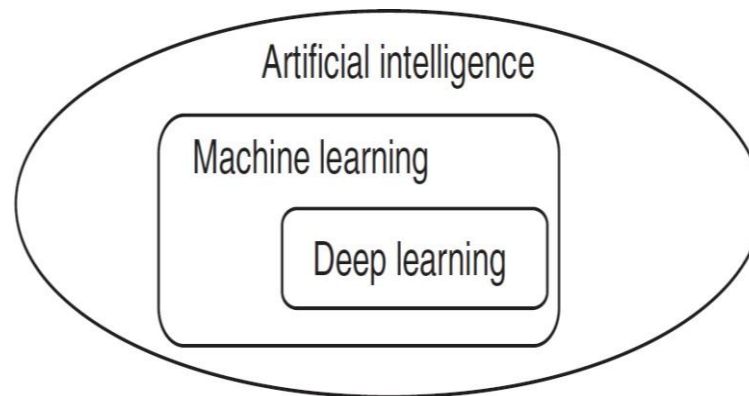


Introduction To Analytics, Machine Learning and deep learning

Artificial Intelligence: Algorithms and systems that exhibit human-like intelligence.

Machine Learning: Subset of AI that can learn to perform a task with extracted data and/or models.

Deep Learning: Subset of machine learning that imitates the functioning of human brain to solve problems.



Types of Machine learning algorithms

1. Supervised Learning Algorithms: Require the knowledge of both the outcome variable (dependent variable) and the features (independent variable or input variables). E.G., Linear regression, logistic regression, discriminant analysis.

In the case of multiple linear regression, the regression parameters are estimated by $\sum_{i=1}^n (y_i - \hat{y}_i)^2$

2. Unsupervised Learning Algorithms: Set of algorithms which do not have the knowledge of the outcome variable in the dataset. E.g., clustering, principal component analysis.

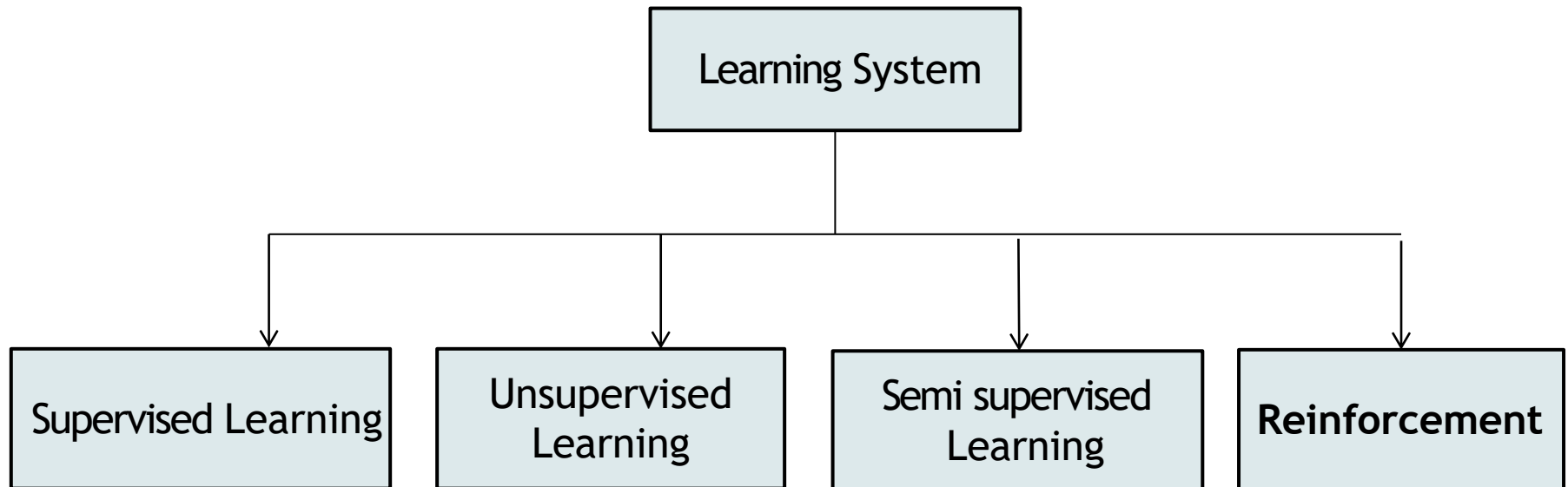
3. Reinforcement Learning Algorithms: Algorithms that have to take sequential actions (decisions) to maximize a cumulative reward. E.g., techniques such as Markov chain and Markov decision process.

4. Evolutionary Learning Algorithms: Algorithms that imitate natural evolution to solve a problem. E.g., techniques such as genetic algorithm and ant colony optimization.



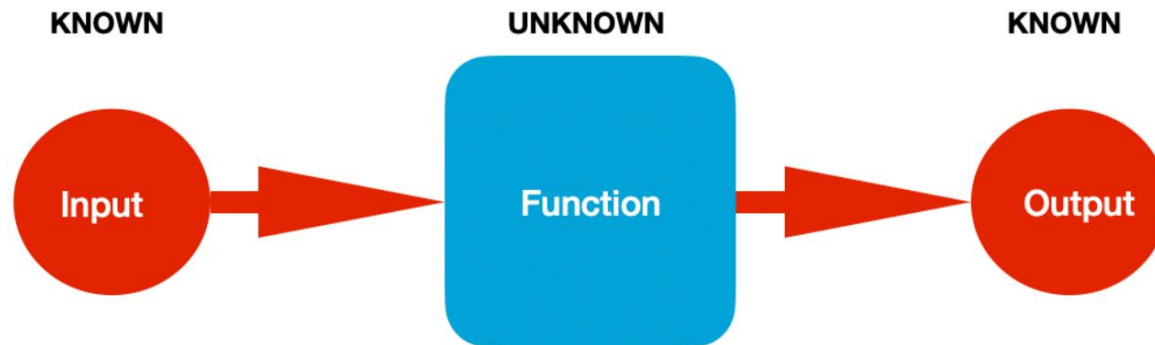
Types of Machine learning algorithms

Machine learning algorithms are often categorized as supervised or unsupervised.



Supervised machine learning algorithms

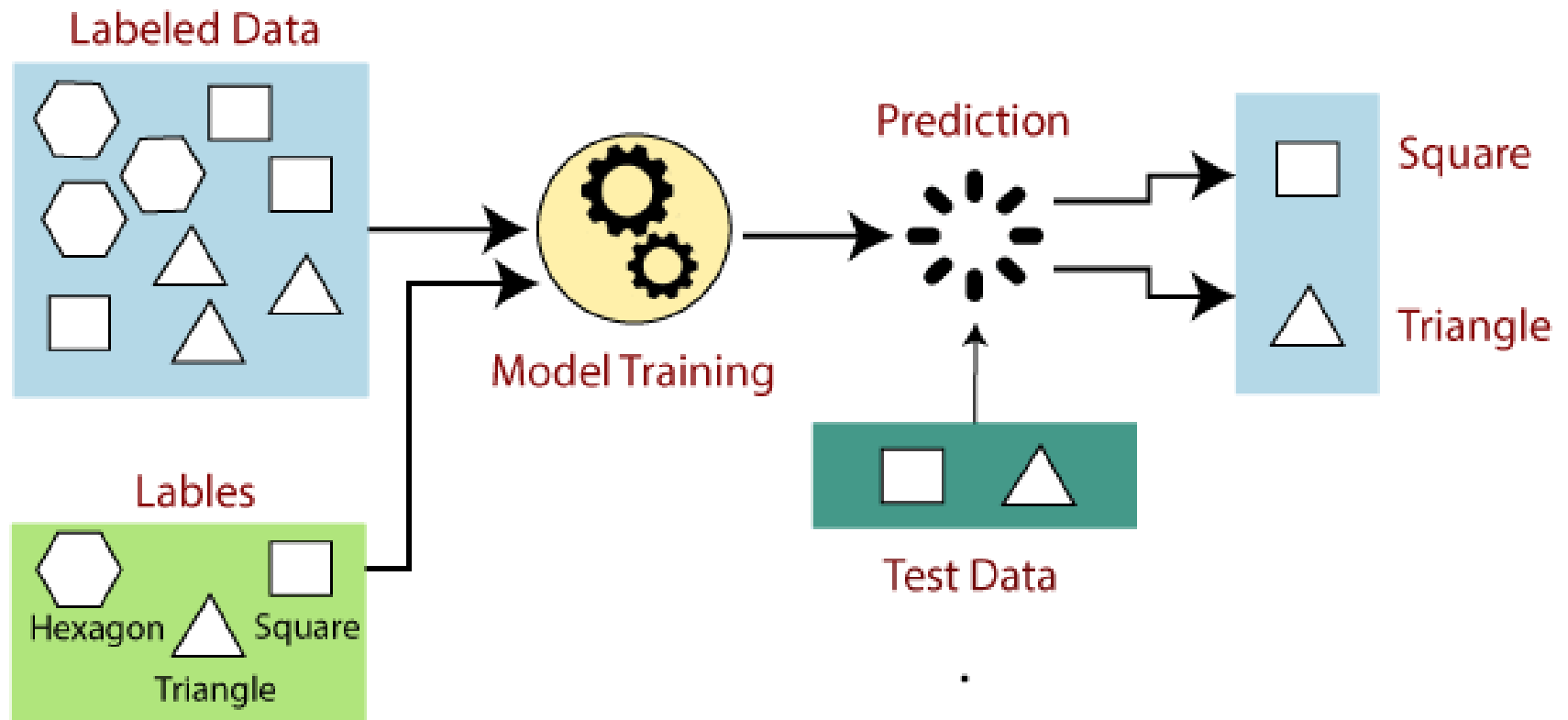
- **Supervised machine learning algorithms** can apply what has been learned in the past to new data using labeled examples to predict future events.
- Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values.
- The system is able to provide targets for any new input after sufficient training.
- The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.



Supervised Learning

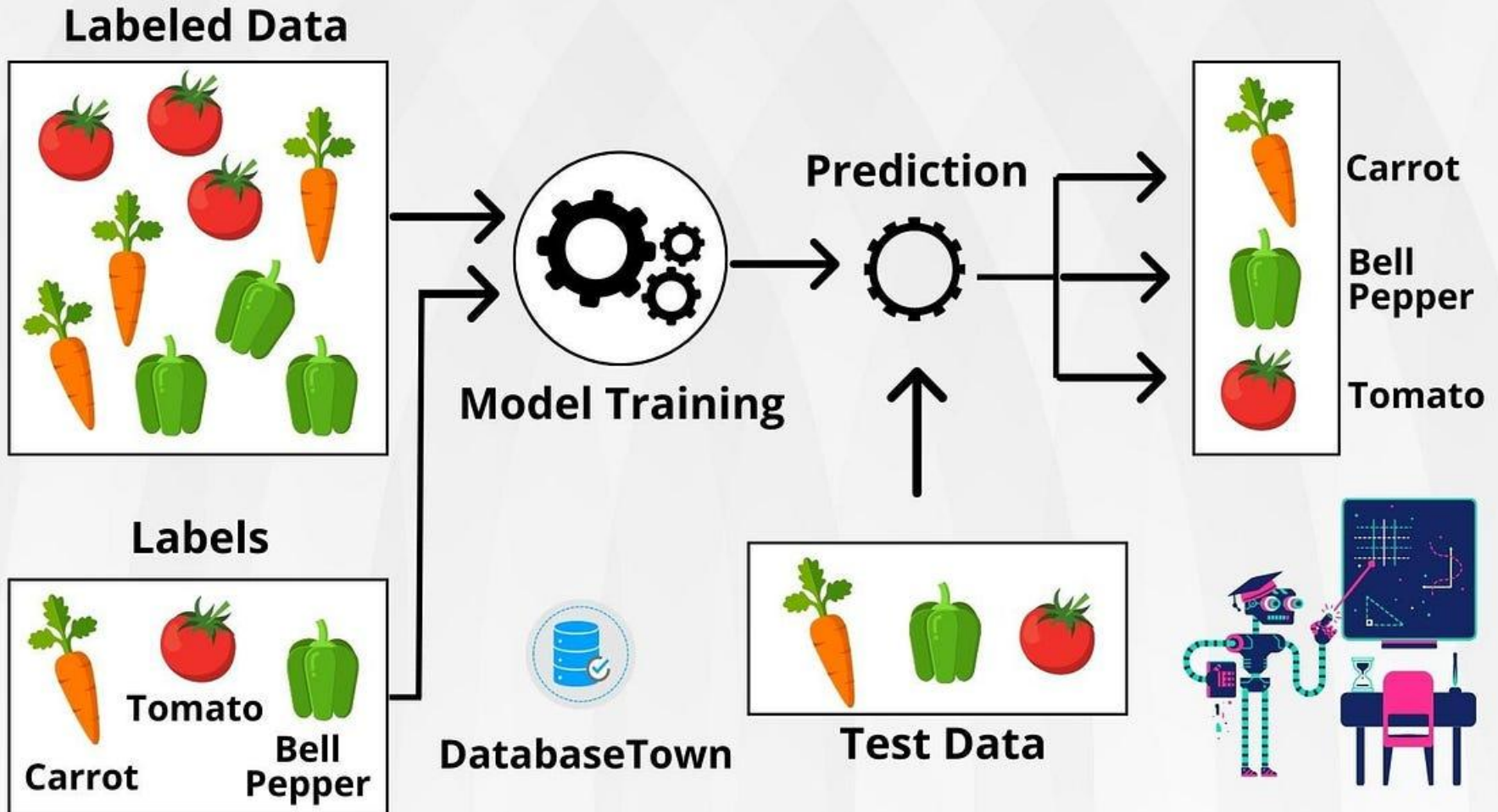


Supervised machine learning algorithms



SUPERVISED LEARNING

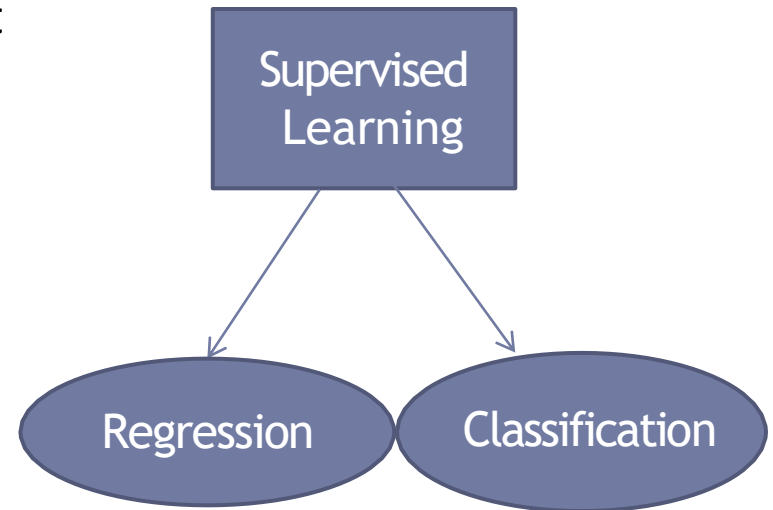
Supervised machine learning is a branch of artificial intelligence that focuses on training models to make predictions or decisions based on labeled training data.



Types of supervised Machine learning

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:

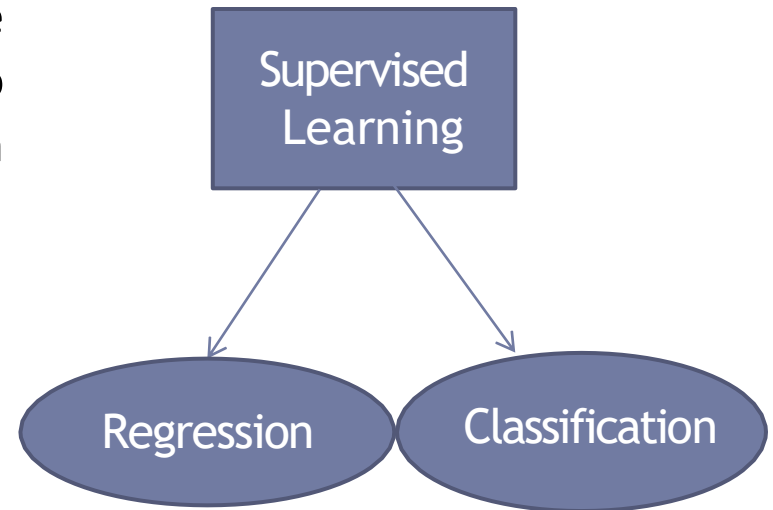
- Linear Regression
- Regression Trees
- Non-Linear Regression
- Bayesian Linear Regression
- Polynomial Regression



Types of supervised Machine learning

Classification : Classification algorithms are used when the output variable is two categorical, which means there are classes such that Yes-No, Male-Female, True-False, etc.

- Spam Filtering,
- Random Forest
- Decision Trees
- Logistic Regression
- Support Vector Machines



Advantages and Disadvantages

Advantages:

- 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.

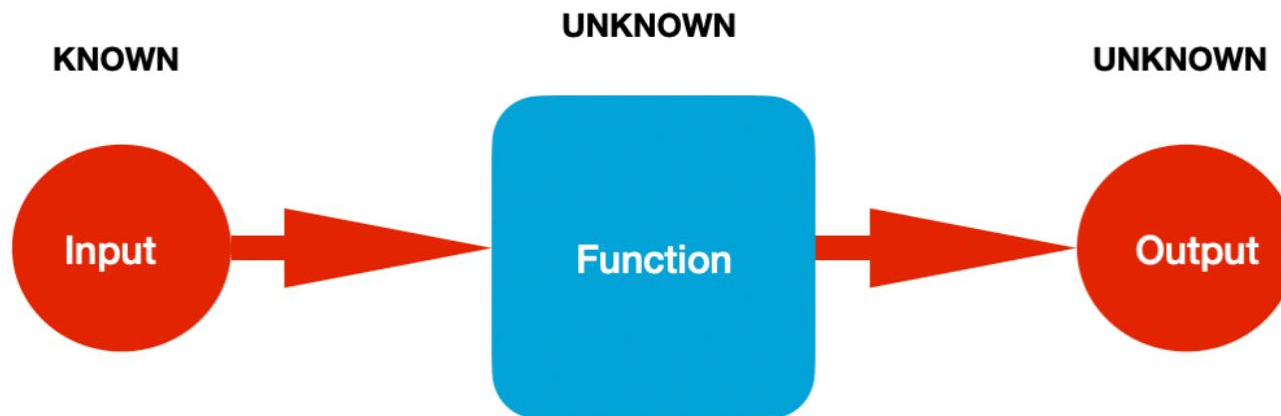
Disadvantages :

- 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 machine learning Algorithms

- In contrast, **unsupervised machine learning algorithms** are used when the information used to train is neither classified nor labeled.
- Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data.
- The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.
- *"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."*

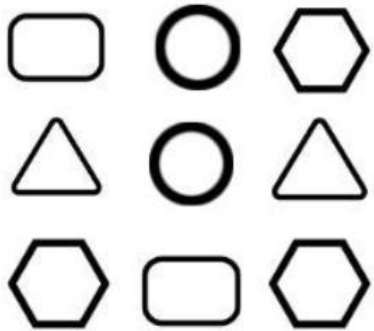


Unsupervised Learning

Unsupervised Learning



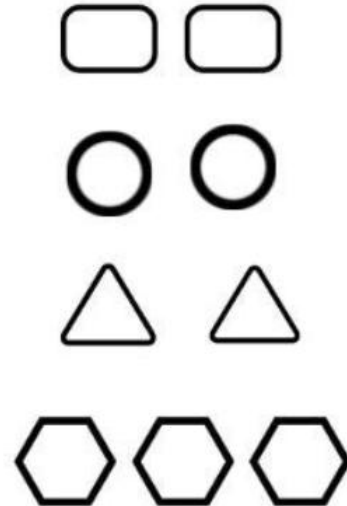
Unlabelled Data



Machine

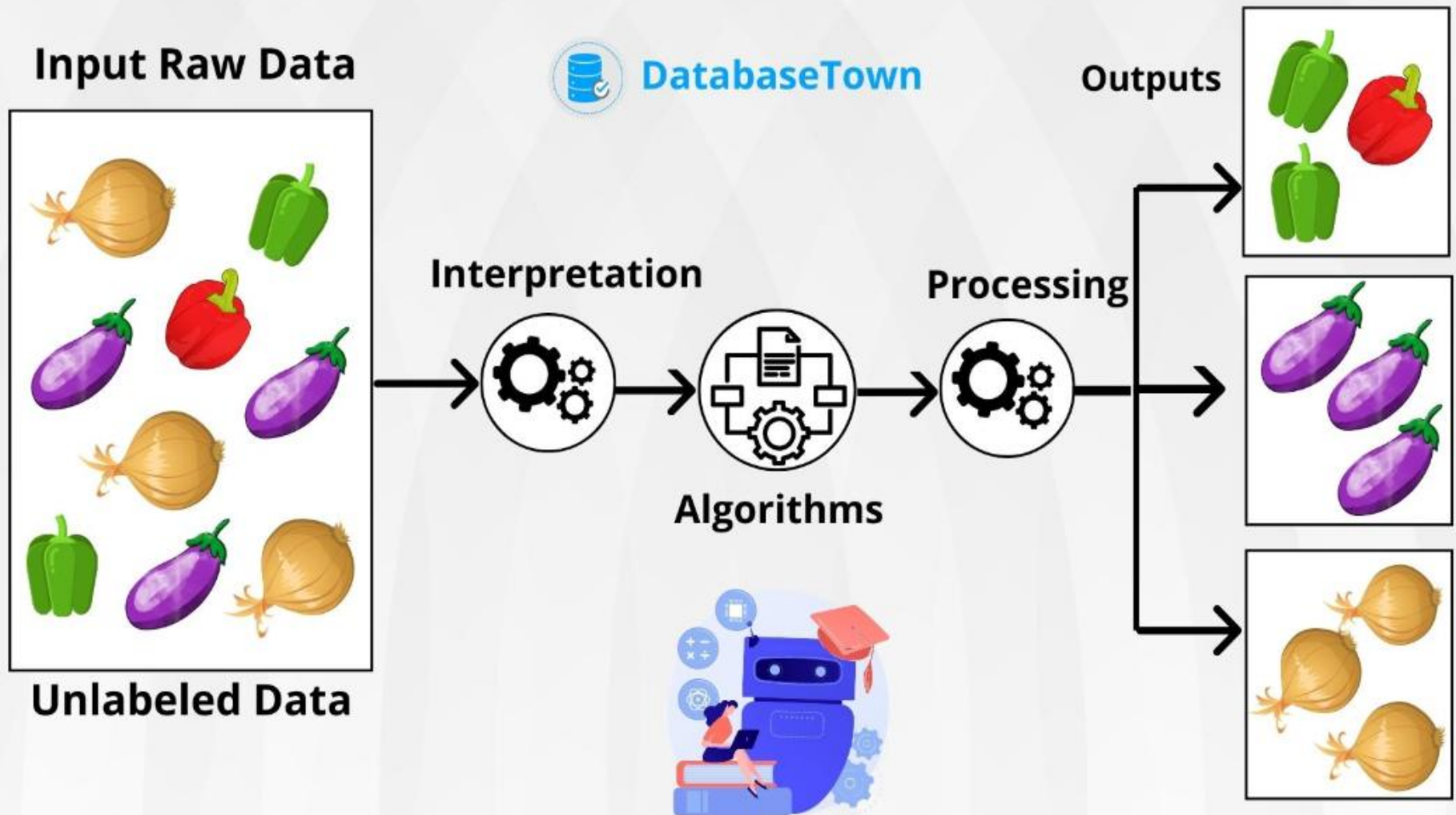


Results



UNSUPERVISED LEARNING

Unsupervised learning is a type of machine learning where the algorithm learns from unlabeled data without any predefined outputs or target variables.

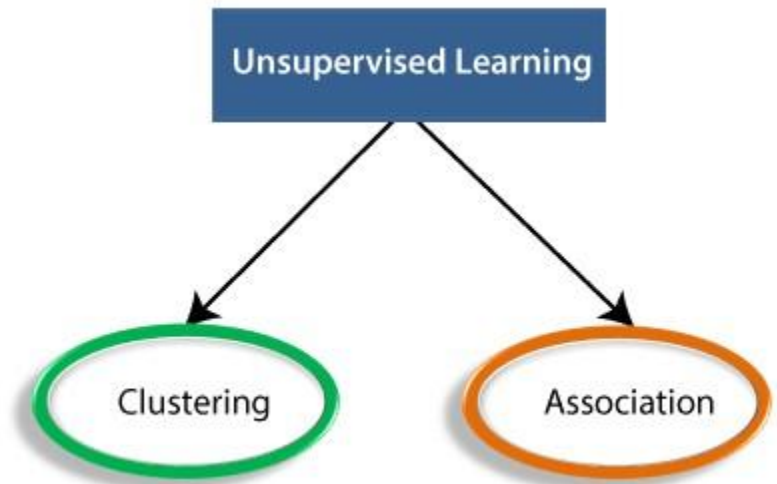


Types of Unsupervised Learning

- **Clustering:**

Clustering is a method of grouping the objects into clusters such that object with most similarities remains into group and has less or no similarity 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.



Types of Unsupervised Learning

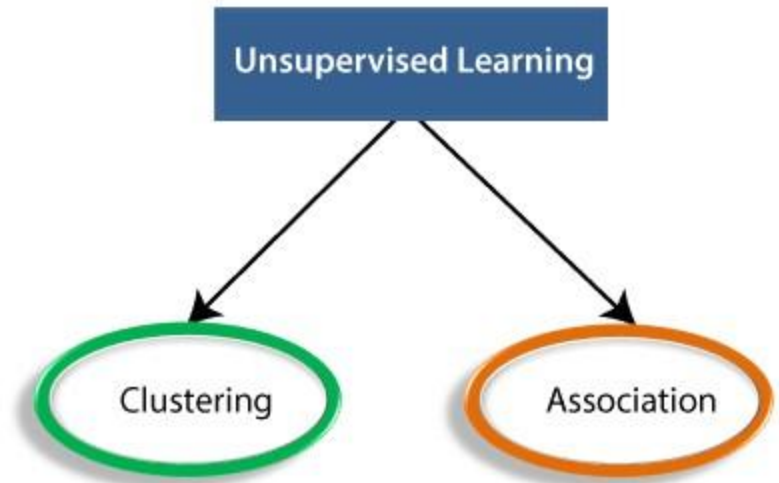
- **Association :**

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 occur together in the data Association rule making 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.



Unsupervised Learning Algorithm

Below is the list of some popular unsupervised learning algorithms:

- **K-means clustering**
- **KNN (k-nearest neighbors)**
- **Hierarchical clustering**
- **Anomaly detection**
- **Neural Networks**
- **Principle Component Analysis**
- **Independent Component Analysis**



Advantages & Disadvantages

Advantages :

- 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 :

- 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**.



Difference between SL and USL

supervised learning

Input data



Annotations

These are apples



Model

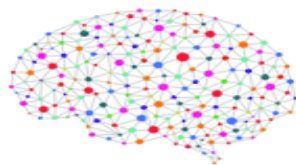


Prediction

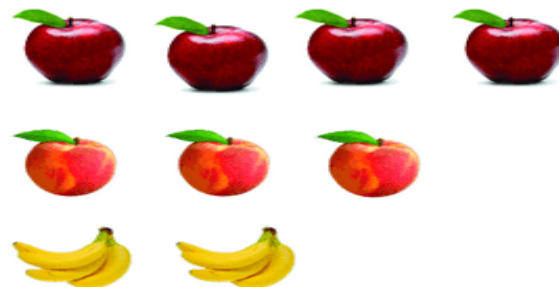
Its an apple!

unsupervised learning

Input data



Model



Difference between SL and USL

Overview comparison between these methods

Category	Supervised	Unsupervised	Semi-supervised	Reinforcement
Input data	All data is labelled	All data is unlabelled	Partially labelled	No predefined data
Training?	External supervision	No supervision	(External supervision)	No supervision
Use	Calculate outcomes	Discover underlying patterns	Improve learning performance	Learn a series of outcomes
Computational complexity	Simple	Complex	Depends	Complex
Accuracy	Higher	Lesser	Lesser	Good for trial/error situations

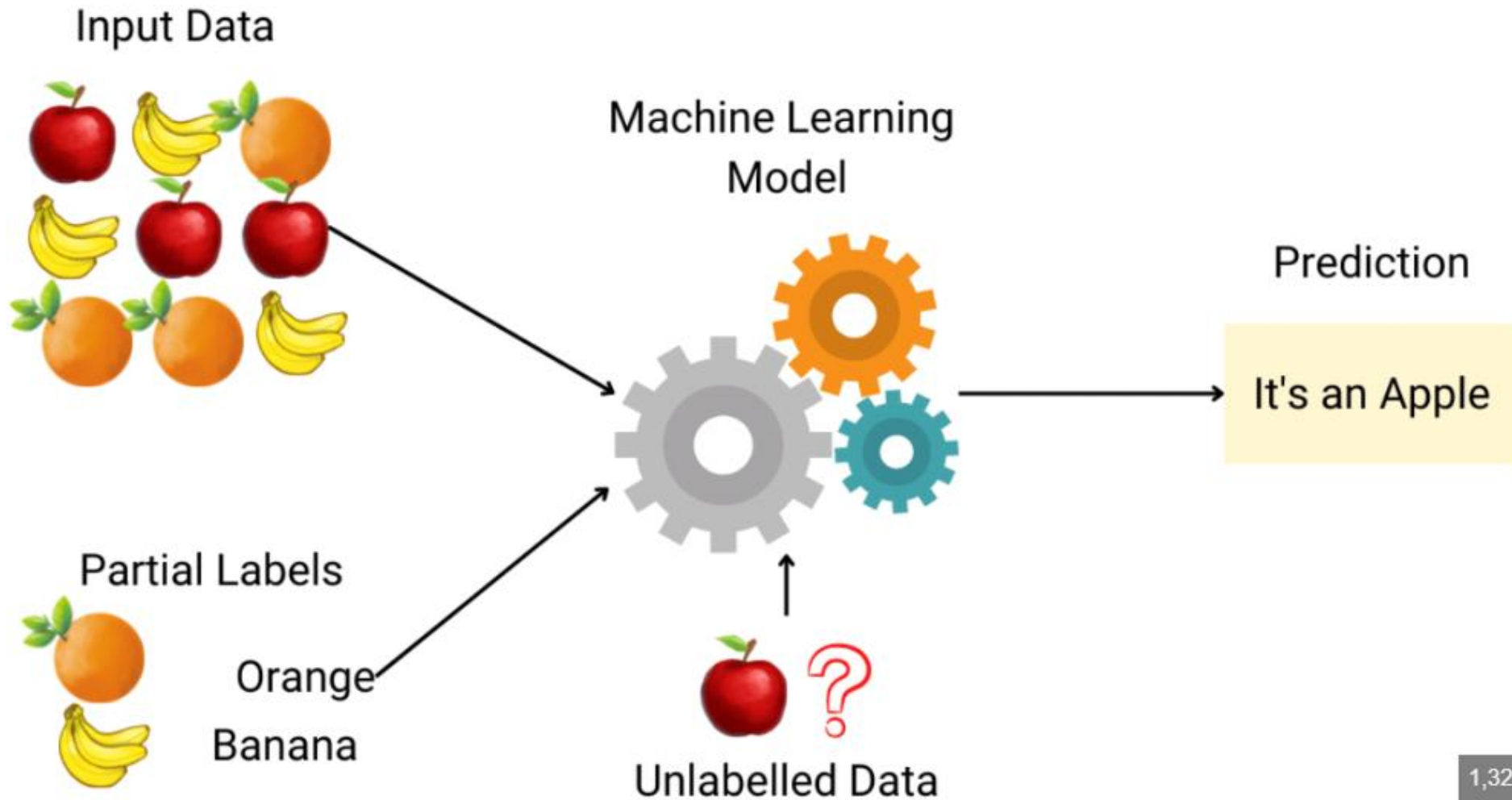


Semi-supervised machine learning algorithms

- **Semi-supervised machine learning algorithms** fall somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training – typically a small amount of labeled data and a large amount of unlabeled data.
- The systems that use this method are able to considerably improve learning accuracy.
- Usually, semi-supervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it.
- Otherwise, acquiring unlabeled data generally doesn't require additional resources.



Semi-supervised machine learning algorithms



Reinforcement machine learning algorithms

- **Reinforcement** machine learning algorithms is a learning method that interacts with its environment by producing actions and discovers errors or rewards.
- Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning.
- This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance.
- Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal.

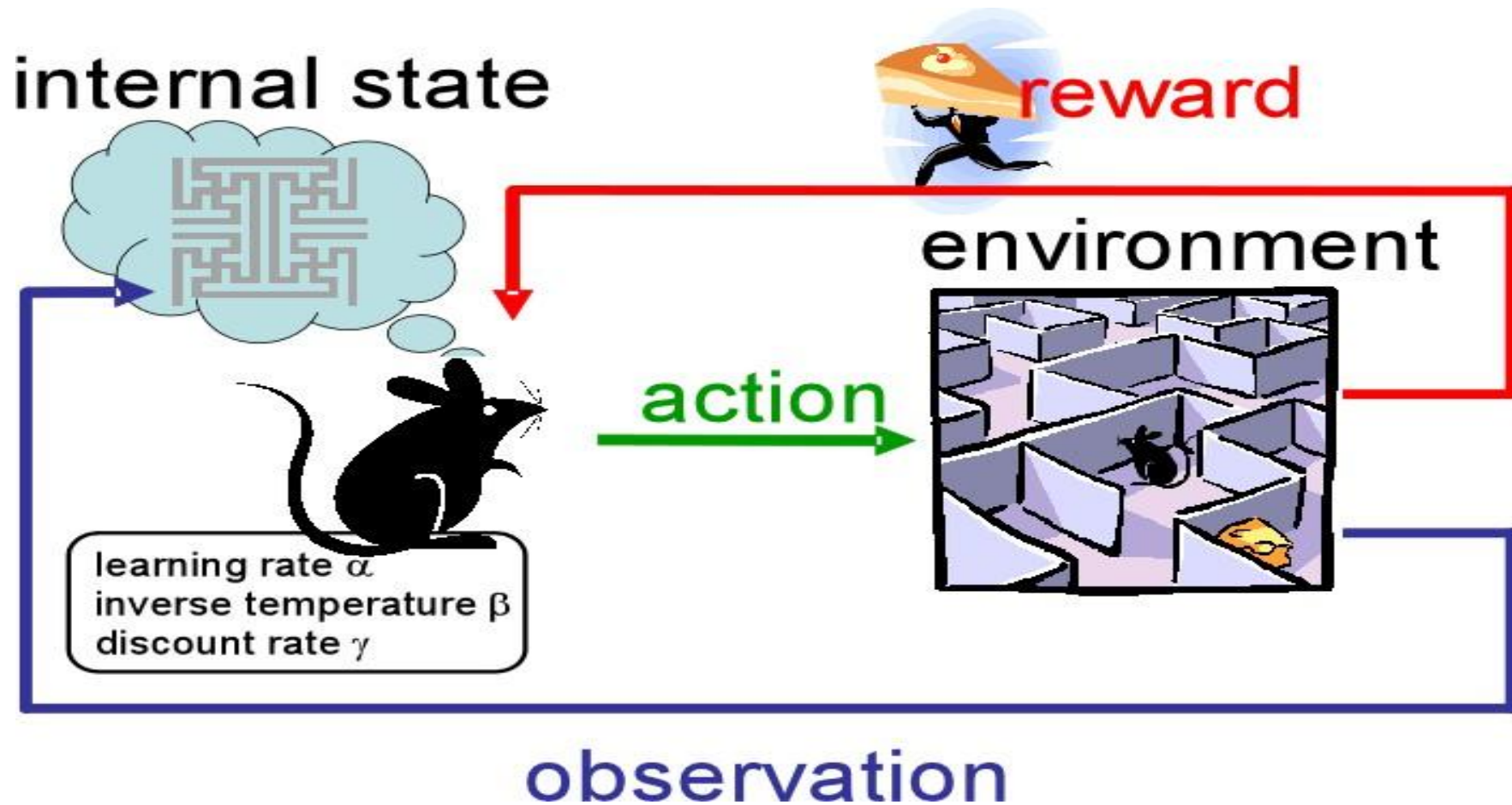


Reinforcement machine learning algorithms

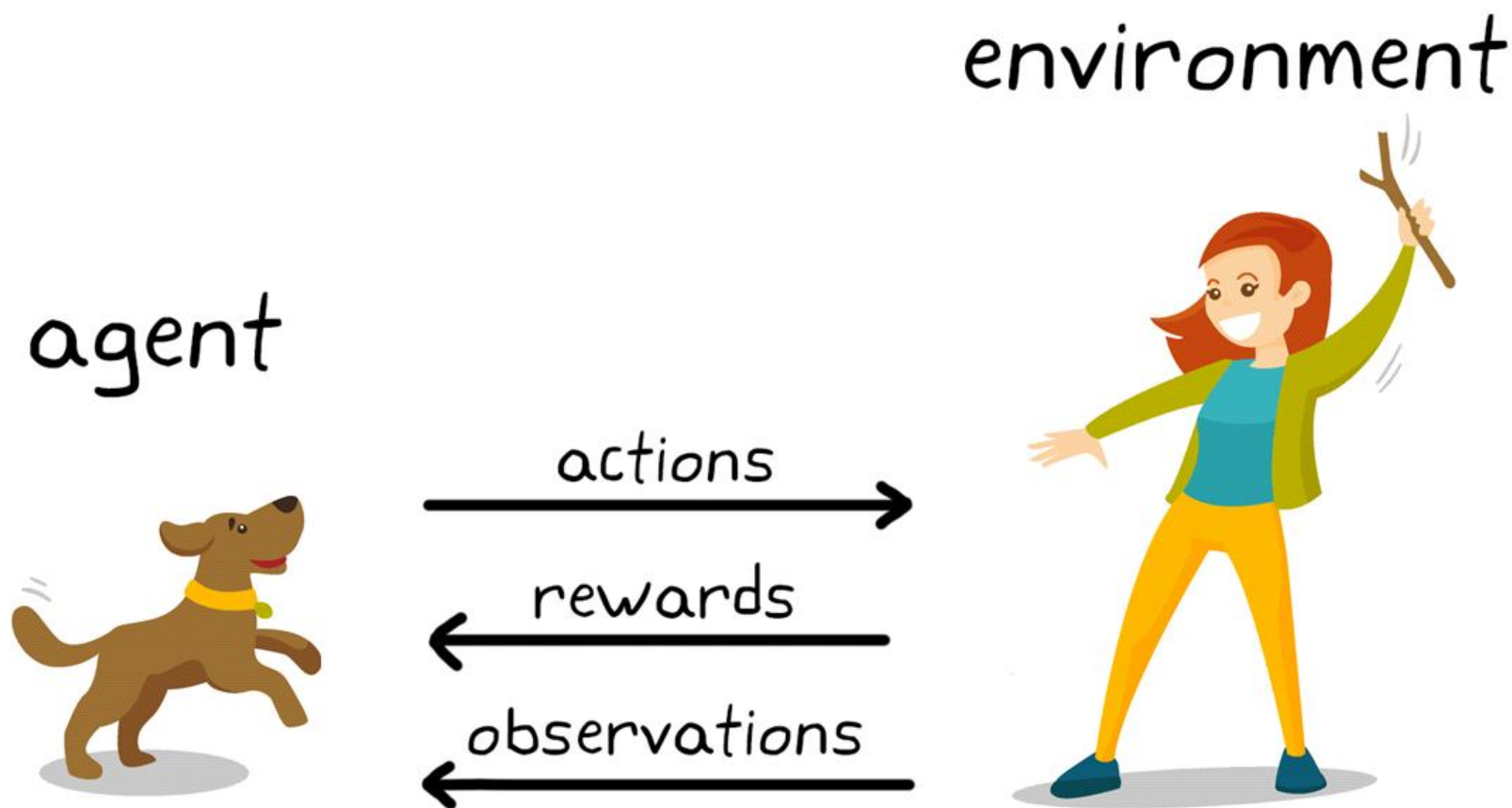
- Though both supervised and reinforcement learning use mapping between input and output, unlike supervised learning where the feedback provided to the agent is **correct set of actions** for performing a task, reinforcement learning uses **rewards and punishments** as signals for positive and negative behavior.
- **Application:**
 - RL is quite widely used in building AI for playing **computer games**.
 - In robotics and industrial automation, RL is used to **enable the robot to create an efficient adaptive control system** for itself which learns from its own experience and behavior.



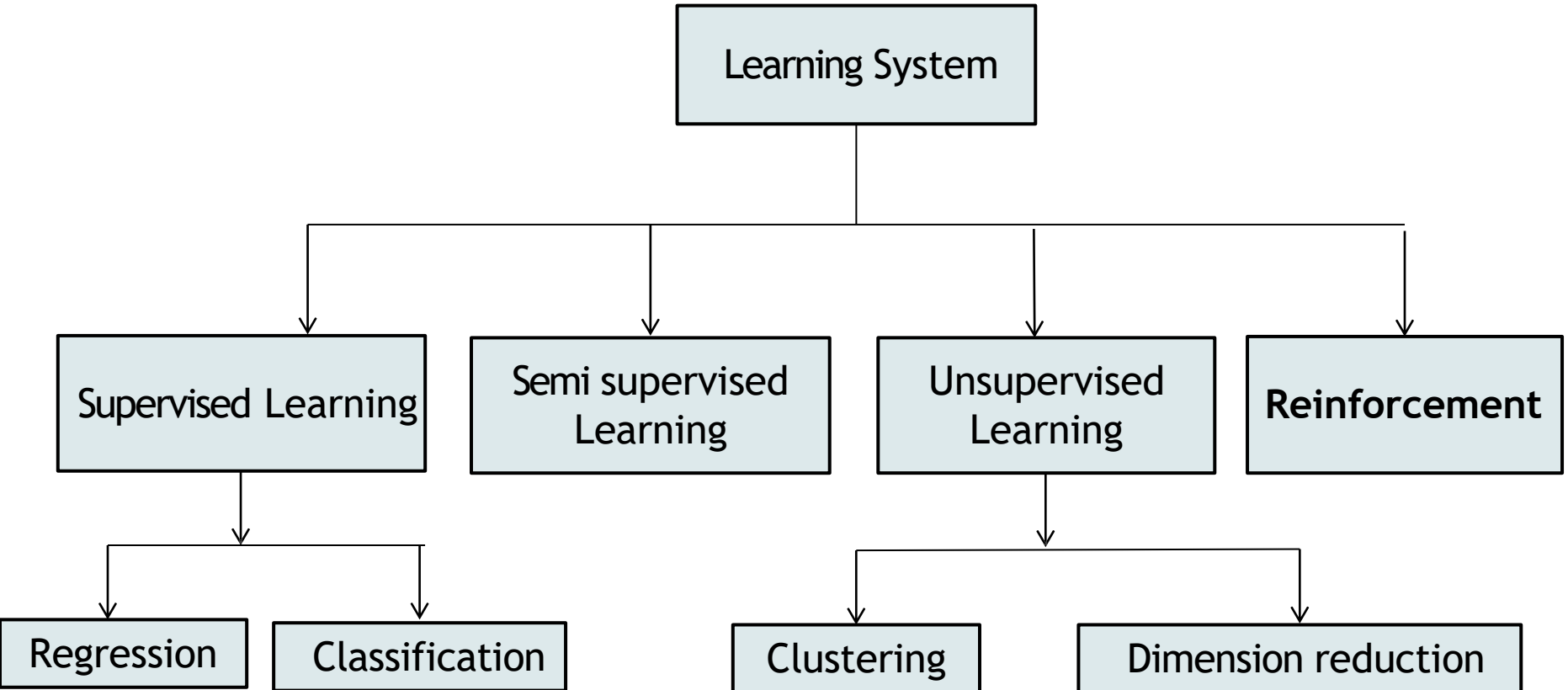
Reinforcement machine learning



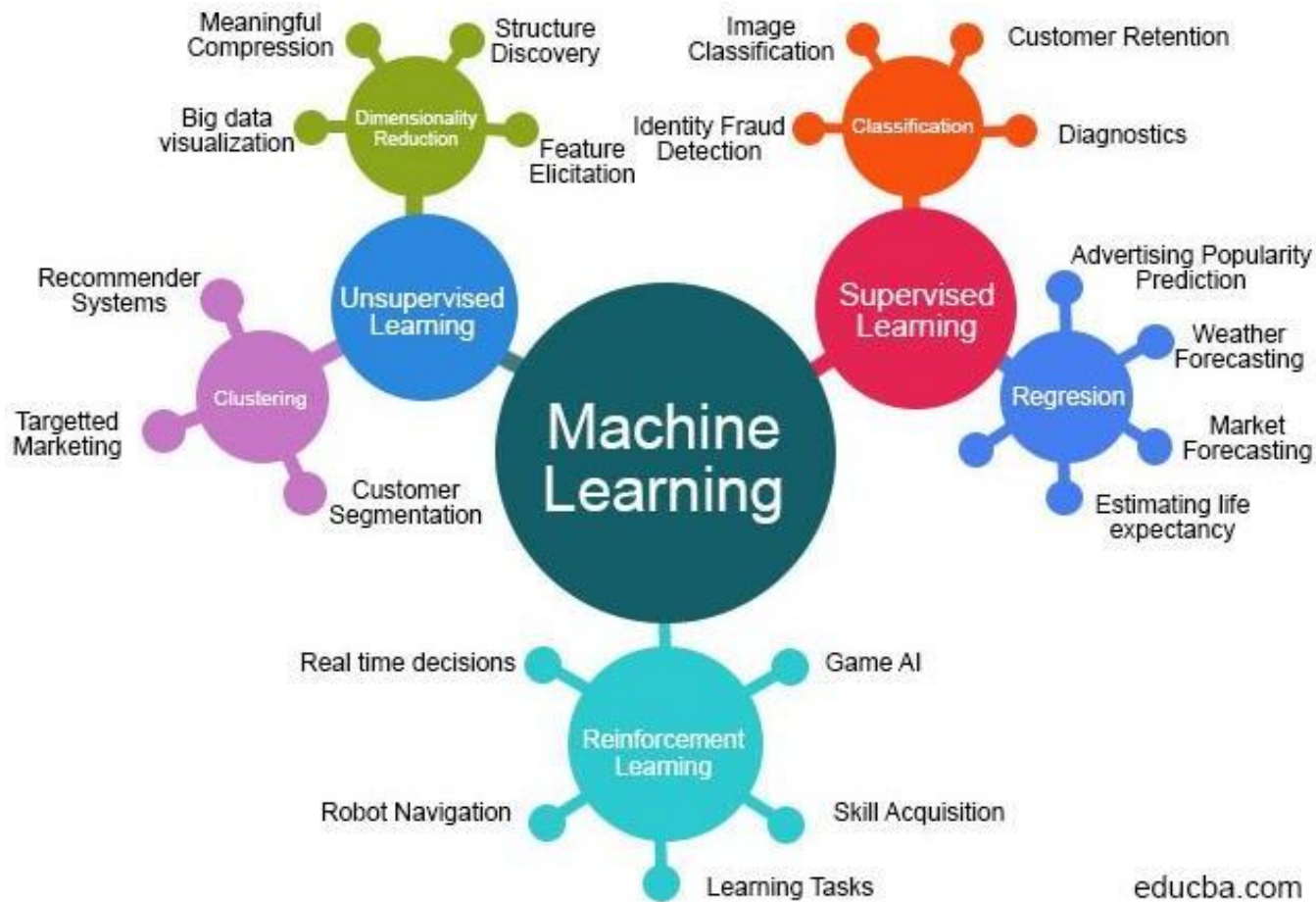
Reinforcement machine learning



Methods of ML



Applications of ML



Applications of ML

1. 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.



Applications of ML

2.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.



Applications of ML

3. 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 from **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.



Applications of ML

4. Product recommendations:

- Machine learning is widely used by various e-commerce and entertainment companies such as **Amazon**, **Netflix**, etc., for product recommendation to the user. Whenever we search for some product on Amazon, then we started getting an advertisement for the same product while internet surfing on the same browser and this is because of machine learning.
- Google understands the user interest using various machine learning algorithms and suggests the product as per customer interest.
- As similar, when we use Netflix, we find some recommendations for entertainment series, movies, etc., and this is also done with the help of machine learning.



Applications of ML

5. 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.



Applications of ML

6. Self-driving cars:

7. Virtual Personal Assistant:

Google assistant, Alexa, Cortana, Siri.

8. Online Fraud Detection:

9. Stock Market trading:

10. Medical Diagnosis:

11. Automatic Language Translation:

Google's GNMT (Google Neural Machine Translation)



Tools for Machine Learning

- **Scikit-learn:**

- Scikit-learn is for machine learning development in python. It provides a
 - library for the Python programming language.

- **PyTorch:**

- PyTorch is a Torch based, Python machine learning library. The torch is a Lua based computing framework, scripting language, and machine learning library.

- **TensorFlow:**

- TensorFlow provides a JavaScript library which helps in machine learning. APIs will help you to build and train the models.

- **Weka**

- **Jupyter Notebook**

- **Python**

- **R programming**



Core Python Libraries for Machine Learning

Areas of Application	Library	Description	Documentation Website
Statistical Computations	SciPy	SciPy contains modules for optimization and computation. It provides libraries for several statistical distributions and statistical tests.	www.scipy.org
Statistical Modelling	StatsModels	StatsModels is a Python module that provides classes and functions for various statistical analyses.	www.statsmodels.org/stable/index.html
Mathematical Computations	NumPy	NumPy is the fundamental package for scientific computing involving large arrays and matrices. It provides useful mathematical computation capabilities.	www.numpy.org
Data Structure Operations (Dataframes)	Pandas	Pandas provides high-performance, easy-to-use data structures called DataFrame for exploration and analysis. DataFrames are the key data structures that feed into most of the statistical and machine learning models.	pandas.pydata.org
Visualization	Matplotlib	It is a 2D plotting library.	matplotlib.org
More elegant Visualization	Seaborn	According to seaborn.pydata.org , Seaborn is a Python visualization library based on matplotlib. It provides a high-level interface for drawing attractive statistical graphics.	seaborn.pydata.org
Machine Learning Algorithm	Scikit-learn (aka sklearn)	Scikit-learn provides a range of supervised and unsupervised learning algorithms.	scikit-learn.org
IDE (Integrated Development Environment)	Jupyter Notebook	According to jupyter.org , the Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and explanatory text.	jupyter.org

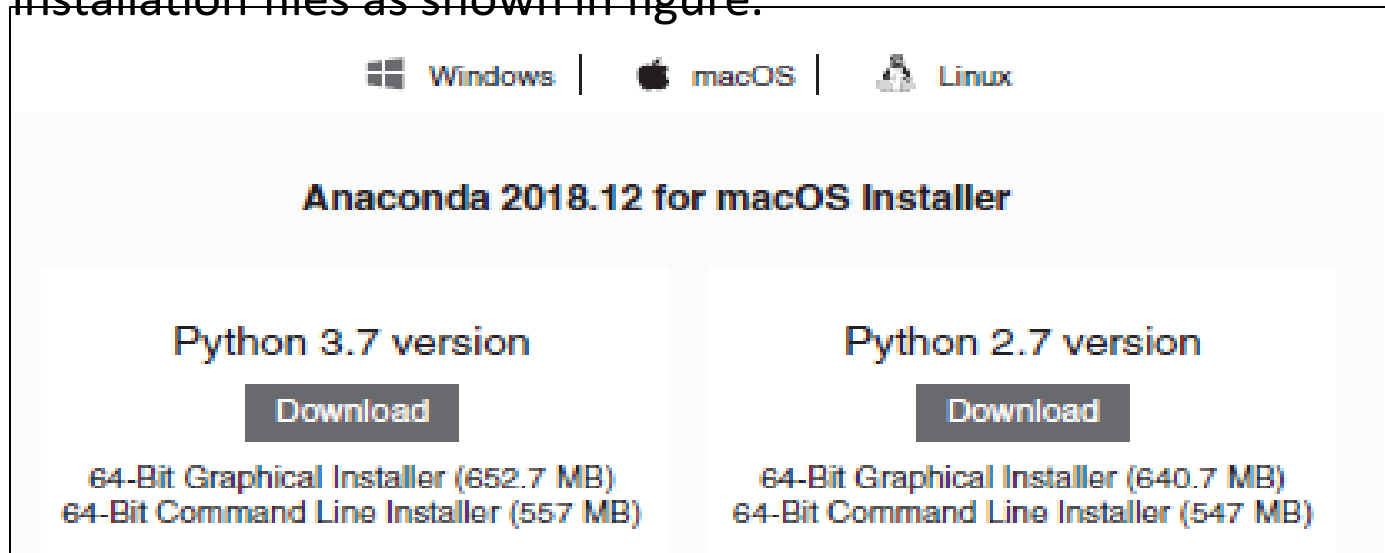
Getting Started With Anaconda Platform

▮ **Step 1:** Go to Anaconda Site

▮ Go to <https://www.anaconda.com/distribution/> using your browser window.

▮ **Step 2:** Download Anaconda Installer for your Environment

Select your OS environment and choose Python 3.7 version to download the installation files as shown in figure.



Getting Started With Anaconda Platform

▮ **Step 3:** Install Anaconda

- ▮ Double click on the downloaded file and follow the on-screen installation instructions, leaving options as set by default. This will take a while and complete the installation process.

▮ **Step 4:** Start Jupyter Notebook

- ▮ Open the command terminal window as per your OS environment and type the following command:

```
jupyter notebook -- ip=*
```

This should start the Jupyter notebook. Open a browser window in your default browser software.

▮ **Step 5:** Create a New Python Program

- ▮ On the browser window, select “New” for a menu. Clicking on the “Folder” will create a directory in the current directory. To create a Python program, click on “Python 3”. It will open a new window, which will be the program editor for the new Python program as shown in the figure.



Getting Started With Anaconda Platform

Step 6: Rename the Program

- By default, the program name will be “Untitled”. Click on it to rename the program and name as per your requirement. For example, we have renamed it to “My First Program”.

Step 7: Write and Execute Code

- Write Python code in the cell and then press SHIFT+ENTER to execute the cell.

Step 8: Basic Commands for Working with Jupyter Notebook

- Click on “User Interface Tour” for features. Or click on “Keyboard Shortcuts” for basic editor commands as shown in the figure.

