# Machine Learning Introduction

What is Machine Learning

Machine learning is a subset of artificial intelligence (AI) that focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed for those tasks.

In simpler terms, it's about creating systems that can learn and improve from experience.
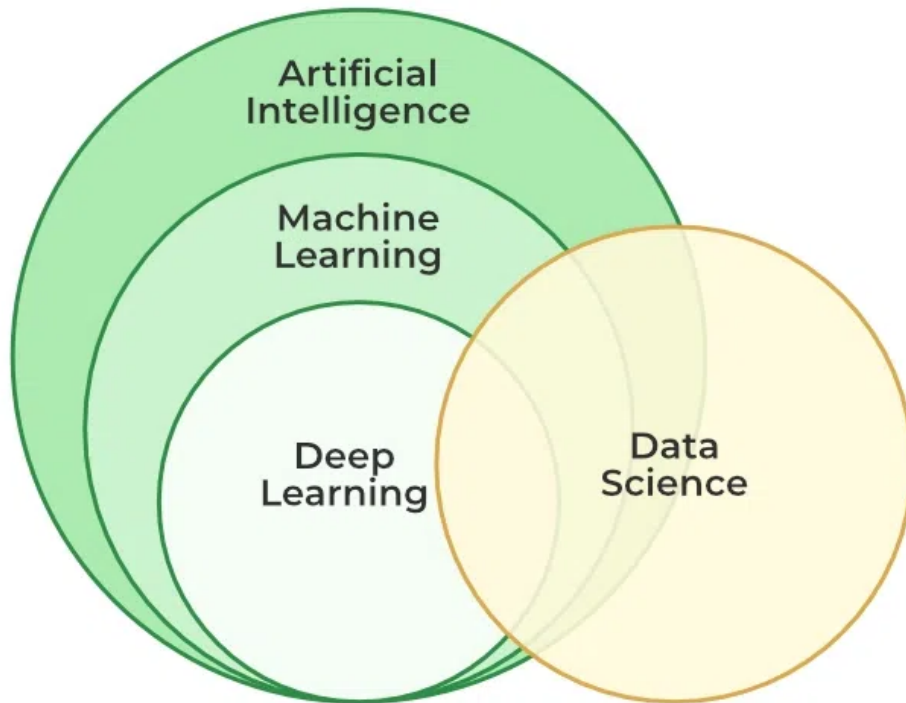
Machine Learning lifecycle

1. Study the Problems:

- The first step is to study the problem. This step involves understanding the business problem and defining the objectives of the model.

2. Data Collection:

- When the problem is well-defined, we can collect the relevant data required for the model.

3. Data Preparation:

When our problem-related data is collected. then it is a good idea to check the data properly and make it in the desired format so that it can be used by the model to find the hidden patterns. This can be done in the following steps:

- Data cleaning
- Data Transformation
- Explanatory Data Analysis and Feature Engineering
- Split the dataset for training and testing.

4. Model Selection:

- The next step is to select the appropriate machine learning algorithm that is suitable for our problem.
- This step requires knowledge of the strengths and weaknesses of different algorithms.

5. Model building and Training:

After selecting the algorithm, we have to build the model.

- In the case of traditional machine learning building mode is easy it is just a few hyperparameter tunings.
- In the case of deep learning, we have to define layer-wise architecture along with input and output size, number of nodes in each layer, loss function, gradient descent optimizer, etc.
- After that model is trained using the preprocessed dataset.

6. Model Evaluation:

Once the model is trained, it can be evaluated on the test dataset to determine its accuracy and performance using different techniques like

- classification report
- F1 score
- precision
- recall
- ROC Curve
- Mean Square error
- absolute error, etc.

7. Model Tuning:

- Based on the evaluation results, the model may need to be tuned or optimized to improve its performance.
- This involves tweaking the hyperparameters of the model.

8.Deployment:

- Once the model is trained and tuned, it can be deployed in a production environment to make predictions on new data.

9. Monitoring and Maintenance:

- Finally, it is essential to monitor the model's performance in the production environment and perform maintenance tasks as required.

Types of Machine Learning
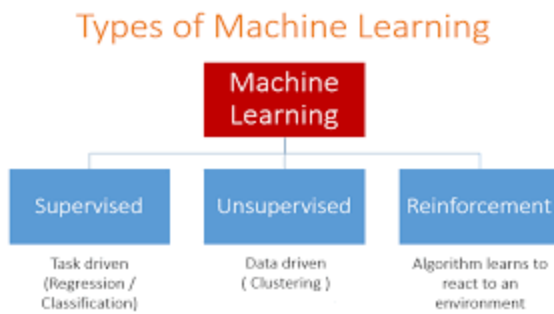
1. Supervised Machine Learning

- Supervised learning is the most common type of machine learning.
- In this approach, the model is trained on a labeled dataset. ###### or
- where machines are trained on labeled datasets and enabled to predict outputs based on the provided training.
- The labeled dataset specifies that some input and output parameters are already mapped.

Supervised machine learning is further classified into two broad categories

1) Classification:

- Classification algorithms are used to predict a categorical output.
- For example, a classification algorithm could be used to predict whether an email is spam or not.

2) Regression:

- Regression algorithms are used to predict a continuous numerical output.
- For example, a regression algorithm could be used to predict the price of a house based on its size, location, and other features.
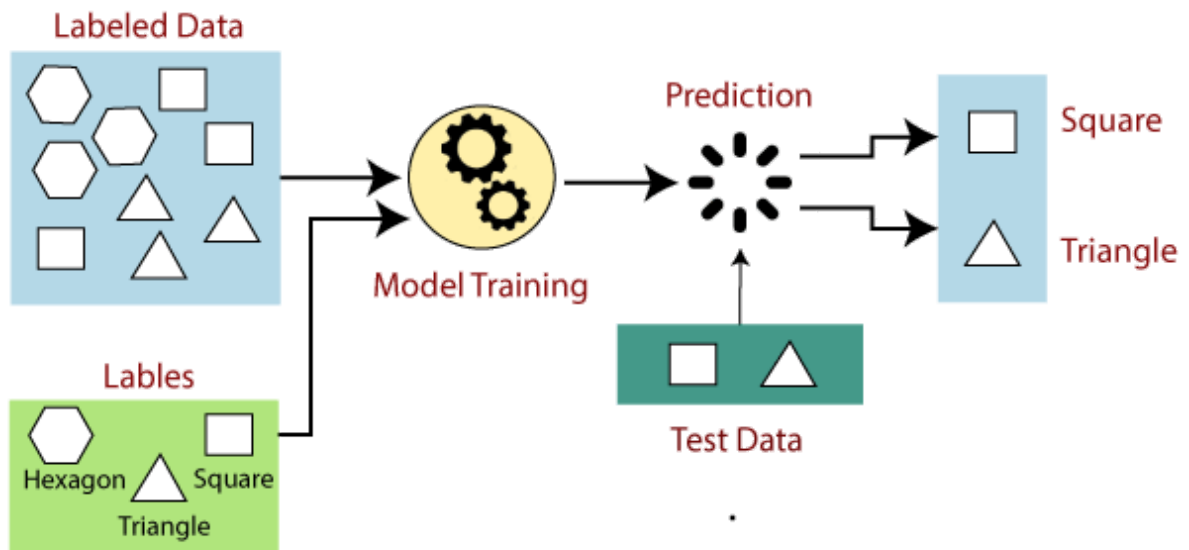
2. Unsupervised machine learning

- The training is provided to the machine with the set of data that has not been labeled, classified, or categorized, and the algorithm needs to act on that data without any supervision.
- The goal of unsupervised learning is to restructure the input data into new features or a group of objects with similar patterns.

It can be further classifieds into two categories of algorithms:

1) Clustering:

  - The clustering technique refers to grouping objects into clusters based on parameters such as similarities or differences between objects.
  - For example, grouping customers by the products they purchase.

2)Association:

- Association learning refers to identifying typical relations between the variables of a large dataset.
- It determines the dependency of various data items and maps associated variables.
- Typical applications include web usage mining and market data analysis.

Raw Data   Output
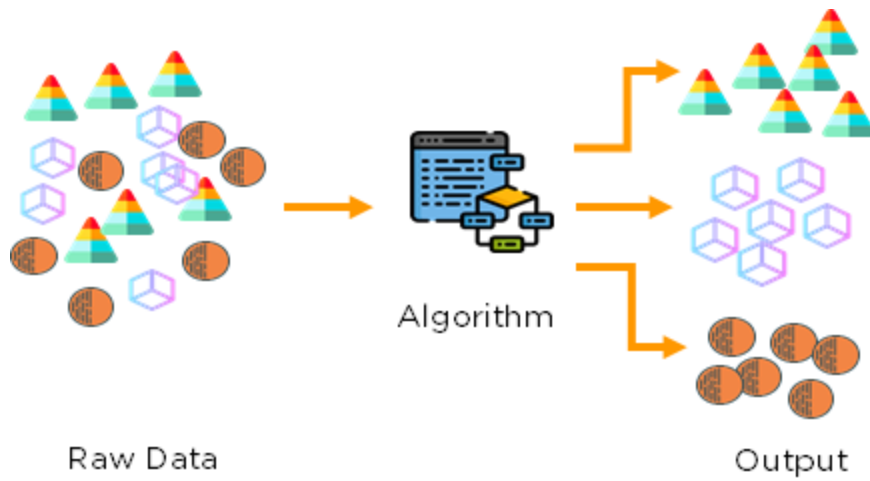
3. Reinforcement learning

- 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 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.

In [ ]: