



# Machine Learning

---

**Course Instructor : Dr. Sarwar Shah Khan**

---

# Introduction to Machine Learning

Lecture # 01

---

# Machine Learning

---

- ❑ Machine learning is a growing technology which enables computers to learn automatically from past data.
- ❑ Machine learning uses various algorithms for **building mathematical models and making predictions using historical data or information.**

Currently, it is being used for various tasks such as

- ❑ Image recognition
- ❑ Speech recognition
- ❑ Email filtering,
- ❑ Facebook auto-tagging,
- ❑ Recommender system, and many, many more.

# Machine Learning

---

- Machine learning along with the wide range of machine learning techniques such as
  - Supervised
  - Unsupervised
  - Reinforcement learning

da copy

# What is Machine Learning

---

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

# How machine learning works

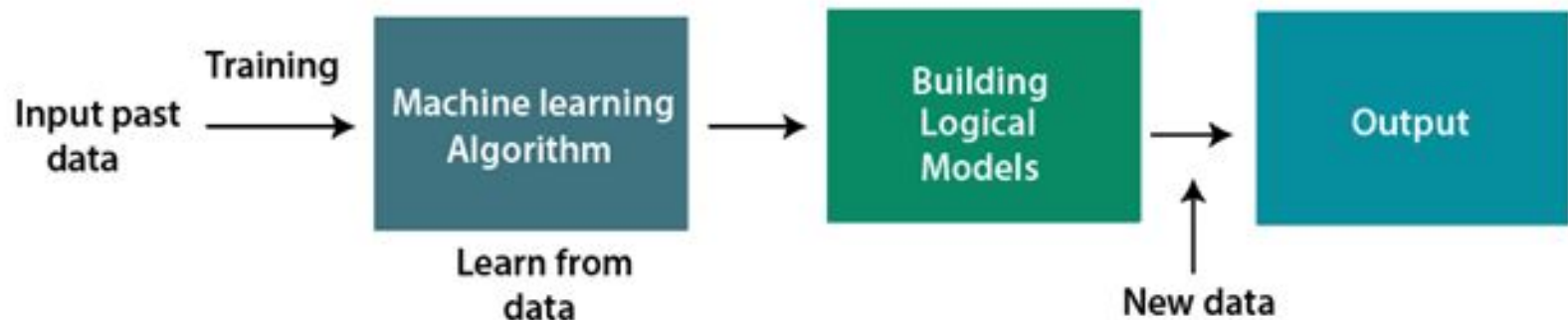
---

- ❑ A Machine Learning system learns from **historical data**, builds the prediction models, and whenever it receives new data, **predicts** the output for it.
- ❑ The accuracy of predicted output depends upon the amount of data, as the huge amount of data helps to build a better model which predicts the output more accurately.
- ❑ Suppose we have a complex problem, where we need to perform some predictions, so instead of writing a code for it, we just need to feed the data to generic algorithms, and with the help of these algorithms, machine builds the logic as per the data and predict the output.

# How machine learning works

---

Machine learning has changed our way of thinking about the problem.



COPY

# Features of Machine Learning

---

- Machine learning uses data to detect various patterns in a given dataset.
- It can learn from past data and improve automatically.
- It is a data-driven technology.
- Machine learning is much similar to data mining as it also deals with the huge amount of the data.



WhatsApp

# Need for Machine Learning

---

- ❑ The need for machine learning is increasing day by day. The reason behind the need for machine learning is that it is capable of doing tasks that are too complex for a person to implement directly.
  - ❑ As a human, we have some limitations as we cannot access the huge amount of data manually, so for this, we need some computer systems and here comes the machine learning to make things easy for us.
  - ❑ We can train machine learning algorithms by providing them the huge amount of data and let them explore the data, construct the models, and predict the required output automatically.
  - ❑ With the help of machine learning, we can save both time and money.
-

# Need for Machine Learning (cont...)

---

- ❑ Various top companies such as **Netflix** and **Amazon** have build machine learning models that are using a vast amount of data to analyze the user **interest** and **recommend** product accordingly.
- ❑ Following are some key points which show the importance of Machine Learning:
  - ❑ Rapid increment in the production of data
  - ❑ Solving complex problems, which are difficult for a human
  - ❑ Decision making in various sector including finance
  - ❑ Finding hidden patterns and extracting useful information from data.

# Why & When to Make Machines Learn?

---

## ❑ Lack of human expertise

- ❑ A machine to learn and take data-driven decisions, can be the domain where there is a lack of human expertise.
- ❑ The examples can be **navigations** in **unknown territories** or **spatial planets**.

## ❑ Dynamic scenarios

- ❑ There are some scenarios which are dynamic in nature i.e. they keep changing over time. In case of these scenarios and behaviors, we want a machine to learn and take data-driven decisions.
- ❑ Some of the examples can be **network connectivity** and **availability of infrastructure** in an organization.

## ❑ Difficulty in translating expertise into computational tasks

- ❑ There can be various domains in which humans have their expertise,; however, they are unable to translate this expertise into computational tasks. In such circumstances we want machine learning. The examples can be the domains of speech recognition, cognitive tasks etc.

# Machine Learning Model

---

Before discussing the machine learning model, we must need to understand the following formal definition of ML given by professor Mitchell –

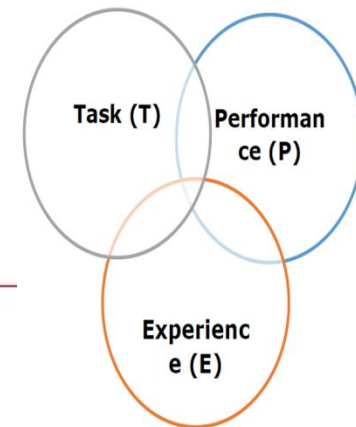
*“A computer program is said to learn from experience  $E$  with respect to some class of tasks  $T$  and performance measure  $P$ , if its performance at tasks in  $T$ , as measured by  $P$ , improves with experience  $E$ . ”*

The above definition is basically focusing on three parameters, also the main components of any learning algorithm, namely Task ( $T$ ), Performance ( $P$ ) and experience ( $E$ ). In this context, we can simplify this definition as –

ML is a field of AI consisting of learning algorithms that –

- Improve their performance ( $P$ )
  - At executing some task ( $T$ )
  - Over time with experience ( $E$ )
-

# Machine Learning Model



## □ Task (T)

- From the perspective of problem, we may define the **task T** as the **real-world problem** to be solved.
- The problem can be anything like finding best **house price in a specific location** or to **find best marketing strategy** etc.
- On the other hand, if we talk about machine learning, the definition of task is different because it is difficult to solve ML based tasks by conventional programming approach.
- A task T is said to be a ML based task when it is based on the process and the system must follow for operating on data points.
- The examples of ML based tasks are Classification, Regression, Clustering, etc.

# Machine Learning Model

---

## □ Experience (E)

- As name suggests, it is the knowledge gained from data points provided to the algorithm or model.
- Once provided with the dataset, the model will run iteratively and will learn some inherent pattern.
- The learning thus acquired is called experience (E).
- Making an analogy with human learning, we can think of this situation as in which a human being is learning or gaining some experience from various attributes like situation, relationships etc.
- The experience gained by our ML model or algorithm will be used to solve the task T.

# Machine Learning Model

---

## ☐ **Performance (P)**

- An ML algorithm is supposed to perform task and gain experience with the passage of time.
- The measure which tells whether ML algorithm is performing as per expectation or not is its performance (P).
- P is basically a quantitative metric that tells how a model is performing the task, T, using its experience, E. There are many metrics that help to understand the ML performance, such as
  - ☐ Accuracy
  - ☐ F1 score
  - ☐ Confusion matrix
  - ☐ Precision
  - ☐ Recall
  - ☐ Sensitivity etc.

COPY

# Challenges in Machines Learning

---

- The challenges that ML is facing currently are –
  - **Quality of data** – Having good-quality data for ML algorithms is one of the biggest challenges. Use of low-quality data leads to the problems related to data preprocessing and feature extraction.
  - **Time-Consuming task** – Another challenge faced by ML models is the consumption of time especially for data acquisition, feature extraction and retrieval.
  - **Lack of specialist persons** – As ML technology is still in its infancy stage, availability of expert resources is a tough job.
-



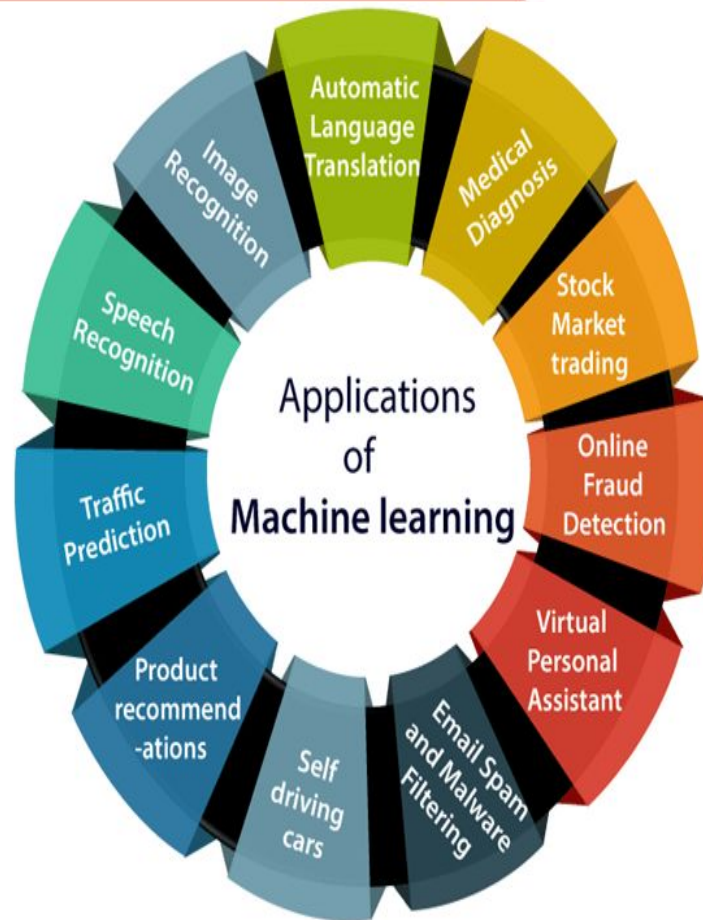
# Challenges in Machines Learning

---

- ❑ **No clear objective for formulating business problems** – Having no clear objective and well-defined goal for business problems is another key challenge for ML because this technology is not that mature yet.
  - ❑ **Issue of overfitting & underfitting** – If the model is overfitting or underfitting, it cannot be represented well for the problem.
  - ❑ **Curse of dimensionality** – Another challenge ML model faces is too many features of data points. This can be a real hindrance.
  - ❑ **Difficulty in deployment** – Complexity of the ML model makes it quite difficult to be deployed in real life.
-

# Applications of Machines Learning

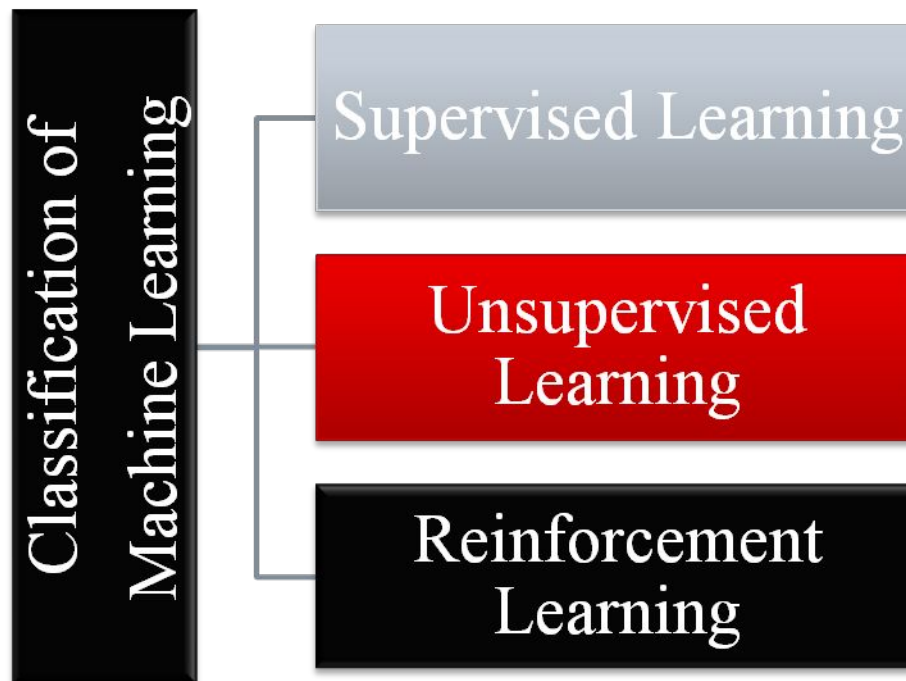
---



Copy

# Classification of Machine Learning

---



# 1) Supervised Learning

---

- Supervised learning is a type of machine learning method in which we provide sample labeled data to the machine learning system in order to train it, and on that basis, it predicts the output.
- The system creates a model using labeled data to understand the datasets and learn about each data, once the training and processing are done then we test the model by providing a sample data to check whether it is predicting the exact output or not.
- Supervised learning can be grouped further in two categories of algorithms:

- Classification

- Regression

## 2) Unsupervised Learning

---

- Unsupervised learning is a learning method in which a machine learns without any supervision.
- 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.
- In unsupervised learning, we don't have a predetermined result. The machine tries to find useful insights from the huge amount of data. It can be further classified into two categories of algorithms:

- 
- Clustering
  - Association

### 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.
- The goal of an agent is to get the most reward points, and hence, it improves its performance.
- The robotic dog, which automatically learns the movement of his arms, is an example of Reinforcement learning.

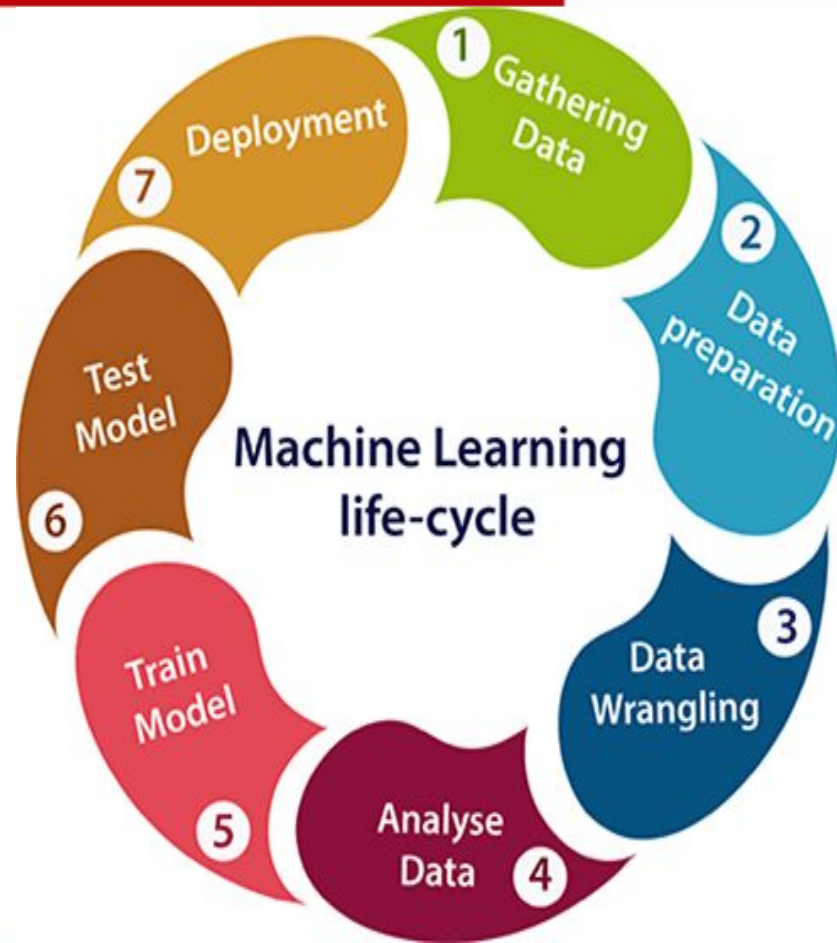
# Machine learning Life Cycle

---

- ❑ Machine learning has given the computer systems the abilities to automatically learn without being explicitly programmed.
- ❑ But how does a machine learning system work?
- ❑ So, it can be described using the life cycle of machine learning.
- ❑ Machine learning life cycle is a cyclic process to build an efficient machine learning project.
- ❑ The main purpose of the life cycle is to find a solution to the problem or project.

# Machine learning Life Cycle

---





# 1) Gathering Data

---

- ❑ The goal of this Data gathering is to identify and obtain all data-related problems.
  - ❑ In this step, we need to identify the different data sources, as data can be collected from various sources such as files, database, internet, or mobile devices..
  - ❑ The quantity and quality of the collected data will determine the efficiency of the output. The more will be the data, the more accurate will be the prediction.
    - ❑ Identify various data sources
    - ❑ Collect data
    - ❑ Integrate the data obtained from different sources
-

## 2. Data preparation

---

- In this step, first, we put all data together, and then randomize the ordering of data.
- This step can be further divided into two processes:
- **Data exploration**

It is used to understand the nature of data that we have to work with. We need to understand the characteristics, format, and quality of data. A better understanding of data leads to an effective outcome.

  - In this, we find Correlations, general trends, and outliers.
- **Data pre-processing**

Now the next step is preprocessing of data for its analysis.

# 3. Data Wrangling

---

- ☐ Data wrangling is the process of cleaning and converting raw data into a useable format.
- ☐ It is the process of cleaning the data, selecting the variable to use, and transforming the data in a proper format to make it more suitable for analysis in the next step.
- ☐ Cleaning of data is required to address the quality issues.
- ☐ In real-world applications, collected data may have various issues, including:
  - ☐ Missing Values
  - ☐ Duplicate data
  - ☐ Invalid data
  - ☐ Noise
- ☐ So, we use various filtering techniques to clean the data.
- ☐ ~~It is mandatory to detect and remove the above issues because it can negatively affect the quality of the outcome.~~

## 4. Data Analysis

---

- Now the cleaned and prepared data is passed on to the analysis step. This step involves:
  - Selection of analytical techniques
  - Building models
  - Review the result
- The aim of this step is to build a machine learning model to analyze the data using various analytical techniques and review the outcome.
- It starts with the determination of the type of the problems, where we select the machine learning techniques such as **Classification, Regression, Cluster analysis, Association**, etc.
- Hence, in this step, we take the data and use machine learning algorithms to build the model.

## 5. Train Model

---

- Now the next step is to train the model, in this step we train our model to improve its performance for better outcome of the problem.
  - We use datasets to train the model using various machine learning algorithms. Training a model is required so that it can understand the various patterns, rules, and, features.
  - **6. Test Model**
  - Once our machine learning model has been trained on a given dataset, then we test the model. In this step, we check for the accuracy of our model by providing a test dataset to it.
  - Testing the model determines the percentage accuracy of the model as per the requirement of project or problem.
-

# 7. Deployment

---

- The last step of machine learning life cycle is deployment, where we deploy the model in the real-world system.
- If the above-prepared model is producing an accurate result as per our requirement with acceptable speed, then we deploy the model in the real system.
- But before deploying the project, we will check whether it is improving its performance using available data or not.
- The deployment phase is similar to making the final report for a project.