Sanjivani Rural Education Society's

Sanjivani College of Engineering, Kopargaon-423603

(An Autonomous Institute Affiliated to Savitribai Phule Pune University, Pune)

NAAC 'A' Grade Accredited

Department of Information Technology

NBA Accredited-UG Programme

Machine Learning

Ms. K. D. Patil

**Assistant Professor**

Contents

- Introduction: Definition, Real life applications, Introduction to Data in Machine Learning Types of Learning: Supervised Learning Unsupervised Learning, Semi-Supervised Learning, Reinforcement Learning, Concept of Feature, Feature Construction, Feature Selection and Transformation, Curse of Dimensionality. Dataset Preparation: Training Vs. Testing Dataset, Dataset Validation Techniques – Hold-out, kfold Cross validation, Leave-One-Out Cross- Validation (LOOCV)

Course Outcome

- **CO1:** To recognize the characteristics of machine learning that makes it useful to real-world problems.

Introduction to Machine Learning (ML)
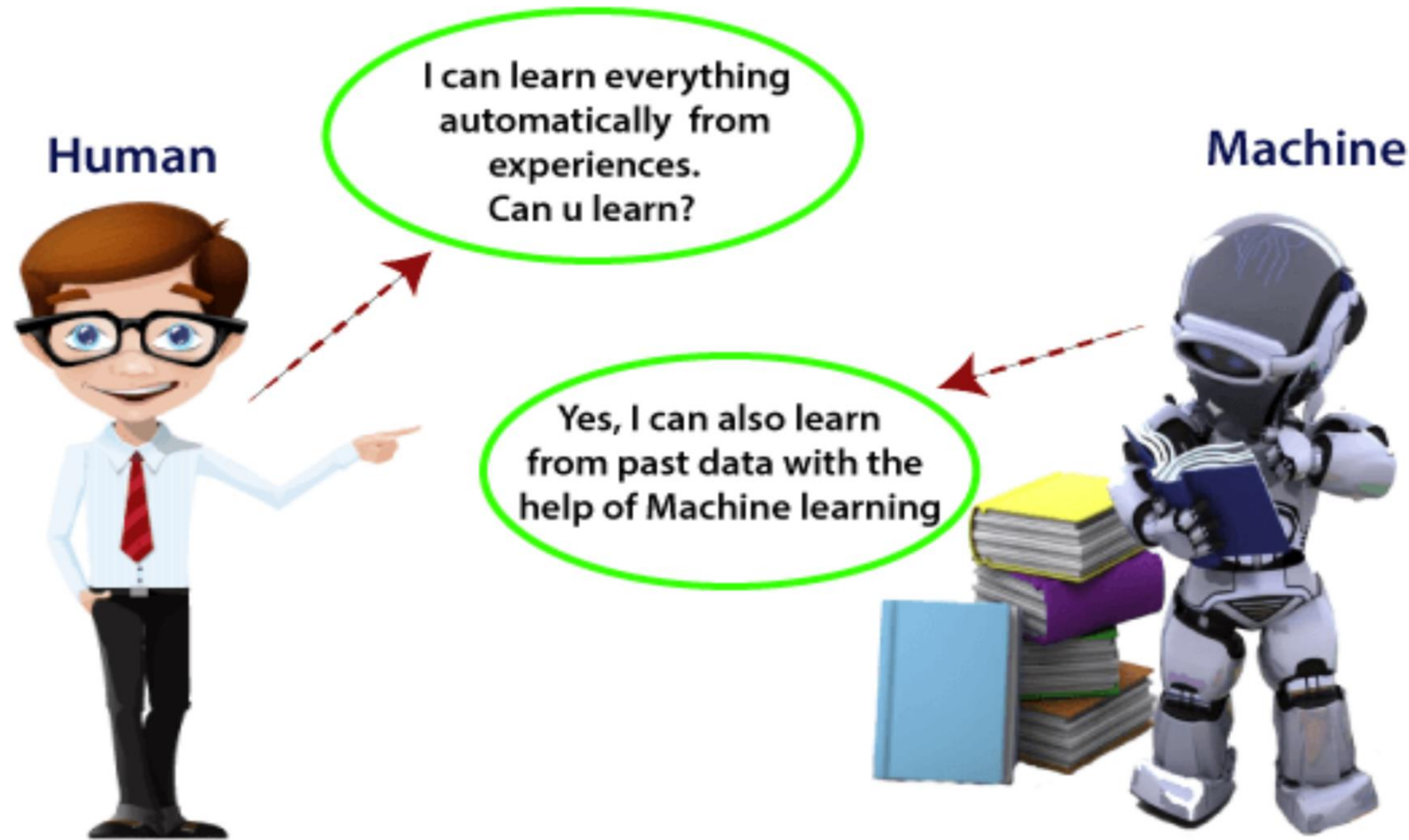
- Machine Learning is said as a subset of **Artificial Intelligence (AI)** that is mainly concerned with the development of algorithms which allow a computer to learn from the data and past experiences on their own.

- The term machine learning was first introduced by **Arthur Samuel** in **1959**.

"Machine learning enables a machine to automatically learn from data, improve performance from experiences, and predict things without being explicitly programmed."

Introduction to Machine Learning (ML)

- With the help of sample historical data, which is known as **Training Data**, machine learning algorithms build a **mathematical model** that helps in making predictions or decisions without being explicitly programmed.
- Machine learning brings computer science and statistics together for creating predictive models.
- Machine learning constructs or uses the algorithms that learn from historical data. The more we will provide the information, the higher will be the performance.

A machine has the ability to learn if it can improve its performance by gaining more data.
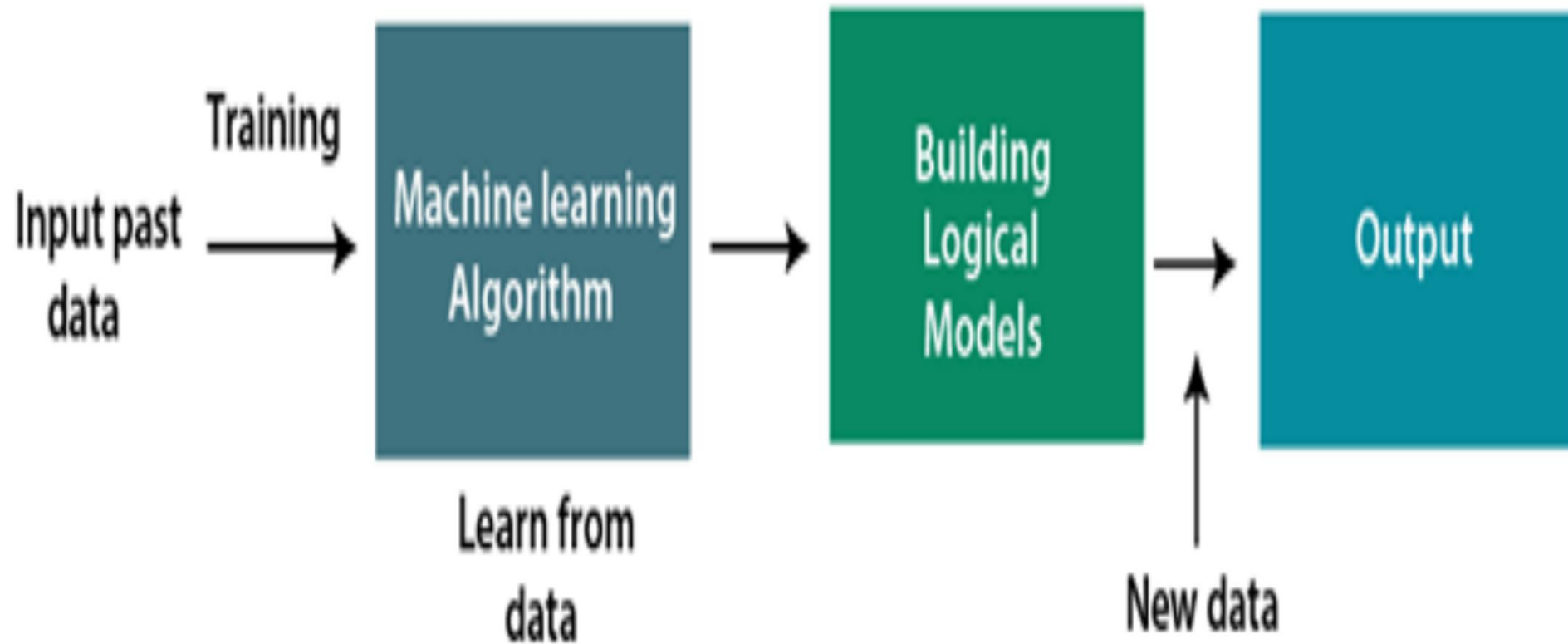
Introduction to Machine Learning (ML)

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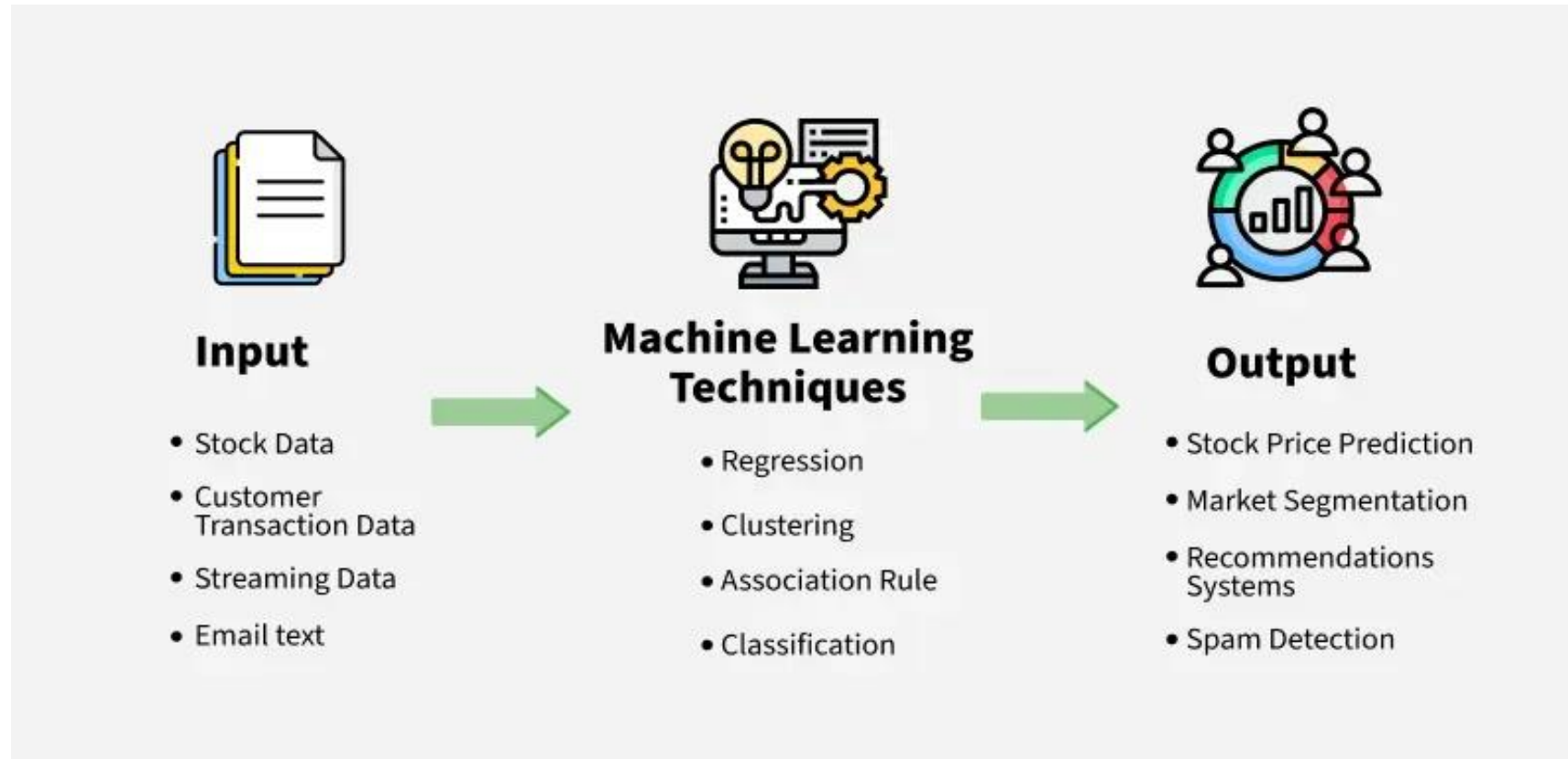- Machine learning has changed our way of thinking about the problem.

Introduction to Machine Learning (ML)

- 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.

- Machine learning has changed our way of thinking about the problem.
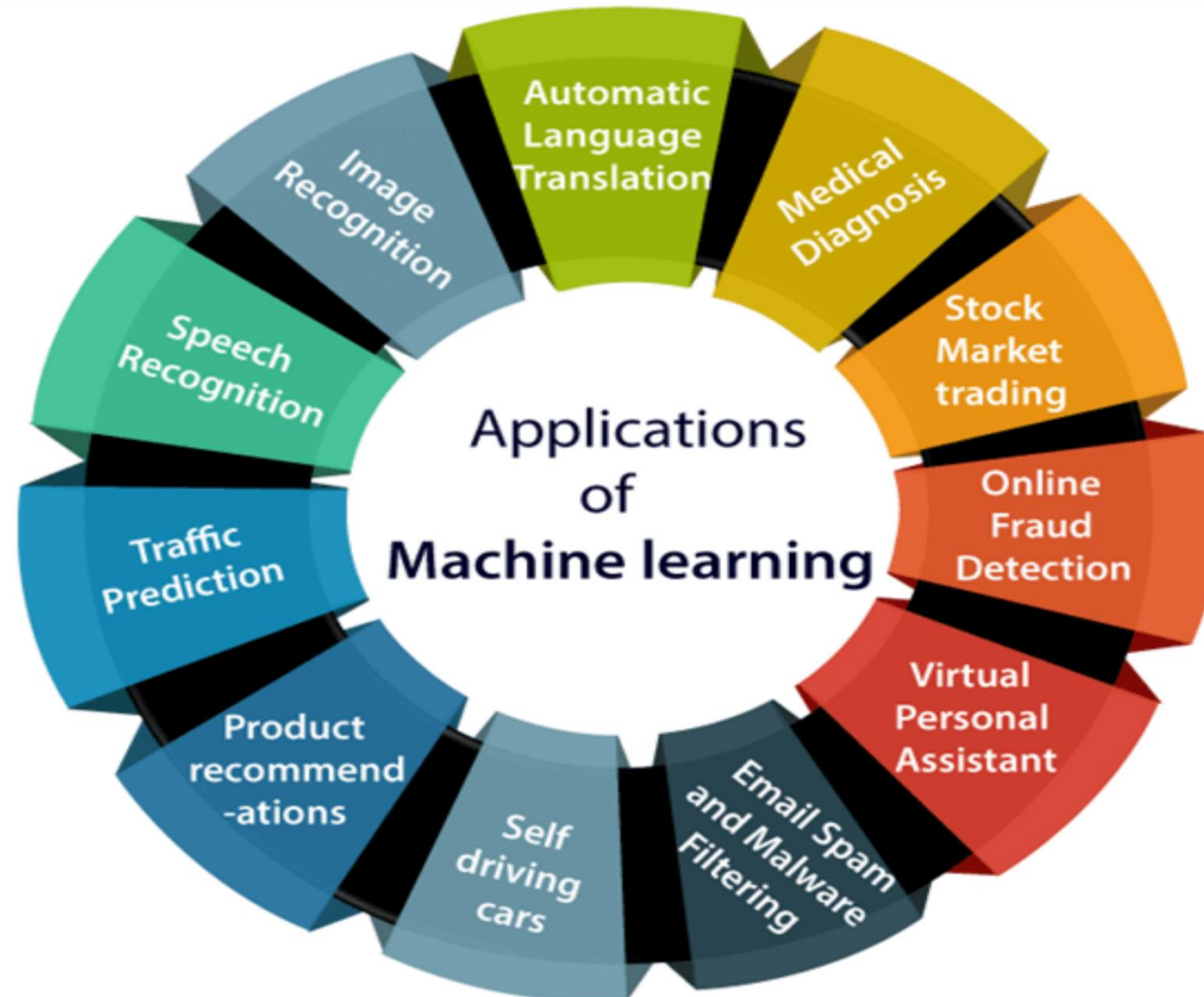
How does Machine Learning works?

How does Machine Learning works?

Input
- Stock Data
- Customer Transaction Data
- Streaming Data
- Email text

Machine Learning Techniques
- Regression
- Clustering
- Association Rule
- Classification

Output
- Stock Price Prediction
- Market Segmentation
- Recommendations Systems
- Spam Detection

Ref: https://www.geeksforgeeks.org/machine-learning/ml-machine-learning/

Applications of Machine Learning

Applications of Machine Learning

- **Virtual Personal Assistants:**
    - Siri, Alexa, Google Now are some of the popular examples of virtual personal assistants. As the name suggests, they assist in finding information, when asked over voice.
    - All you need to do is activate them and ask
        - Q1. What is my schedule for today?
        - Q2. What are the flights from Germany to London, or similar questions.
    - For answering, your personal assistant looks out for the information, recalls your related queries, or send a command to other resources (like phone apps) to collect info.

Applications of Machine Learning

- **Videos Surveillance:**
    - Imagine a single person monitoring multiple video cameras! Certainly, a difficult job to do and boring as well. This is why the idea of training computers to do this job makes sense.
    - The video surveillance system nowadays are powered by AI that makes it possible to detect crime before they happen. They track unusual behaviour of people like standing motionless for a long time, stumbling, or napping on benches etc.
    - The system can thus give an alert to human attendants, which can ultimately help to avoid mishaps. And when such activities are reported and counted to be true, they help to improve the surveillance services. This happens with machine learning doing its job at the backend.

Applications of Machine Learning

- **Social Media Services:**
  - *People You May Know*: Machine learning works on a simple concept: understanding with experiences.
  - Facebook continuously notices the friends that you connect with, the profiles that you visit very often, your interests, workplace, or a group that you share with someone etc.
  - On the basis of continuous learning, a list of Facebook users are suggested that you can become friends with.

Applications of Machine Learning

- **Email Spam and Malware Filtering:**
  - Over 325, 000 malwares are detected everyday and each piece of code is 90-98% similar to its previous versions.
  - The system security programs that are powered by machine learning understand the coding pattern.
  - Therefore, they detects new malware with 2-10% variation easily and offer protection against them.

Applications of Machine Learning

- **Online Customer Support:**
    - A number of websites nowadays offer the option to chat with customer support representative while they are navigating within the site.
    - However, not every website has a live executive to answer your queries.
    - In most of the cases, you talk to a **chatbot**. These bots tend to extract information from the website and present it to the customers.
    - Meanwhile, the chatbots advances with time. They tend to understand the user queries better and serve them with better answers, which is possible due to its machine learning algorithms.

Applications of Machine Learning

- **Search Engine Result Refining:**
    - Google and other search engines use machine learning to improve the search results for you.
    - Every time you execute a search, the algorithms at the backend keep a watch at how you respond to the results.
    - If you open the top results and stay on the web page for long, the search engine assumes that the results it displayed were in accordance to the query.
    - Similarly, if you reach the second or third page of the search results but do not open any of the results, the search engine estimates that the results served did not match requirement. This way, the algorithms working at the backend improve the search results.

Applications of Machine Learning

- **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

Applications of Machine Learning

- **Self-driving Cars:**
  - One of the most exciting applications of machine learning is self-driving cars.
  - Machine learning plays a significant role in self-driving cars.
  - Tesla, the most popular car manufacturing company is working on self-driving car. It is using unsupervised learning method to train the car models to detect people and objects while driving.
  - Eg. Waymo (https://waymo.com/)

Applications of Machine Learning

- **Optical Character Recognition:**
    - Optical character recognition problem, which is the problem of recognizing character codes from their images, is an example of classification problem.
    - This is an example where there are multiple classes, as many as there are characters we would like to recognize.
    - Especially interesting is the case when the characters are handwritten.
    - People have different handwriting styles; and there are many possible images corresponding to the same character.

# Applications of Machine Learning

- **Face Recognition:**
    - In the case of face recognition, the input is an image, the classes are people to be recognized, and the learning program should learn to associate the face images to identity.
    - This problem is more difficult than optical character recognition because there are more classes, input image is larger, and a face is three-dimensional and differences in pose and lighting cause significant changes in the image.

Types of Machine Learning

- Machine learning contains a set of algorithms that work on a huge amount of data. Data is fed to these algorithms to train them, and on the basis of training, they build the model & perform a specific task.

- These ML algorithms help to solve different business problems like Regression, Classification, Forecasting, Clustering, and Associations etc.

1) Supervised Machine Learning
2) Unsupervised Machine Learning
3) Semi-Supervised Machine Learning
4) Reinforcement Learning

Types of Machine Learning
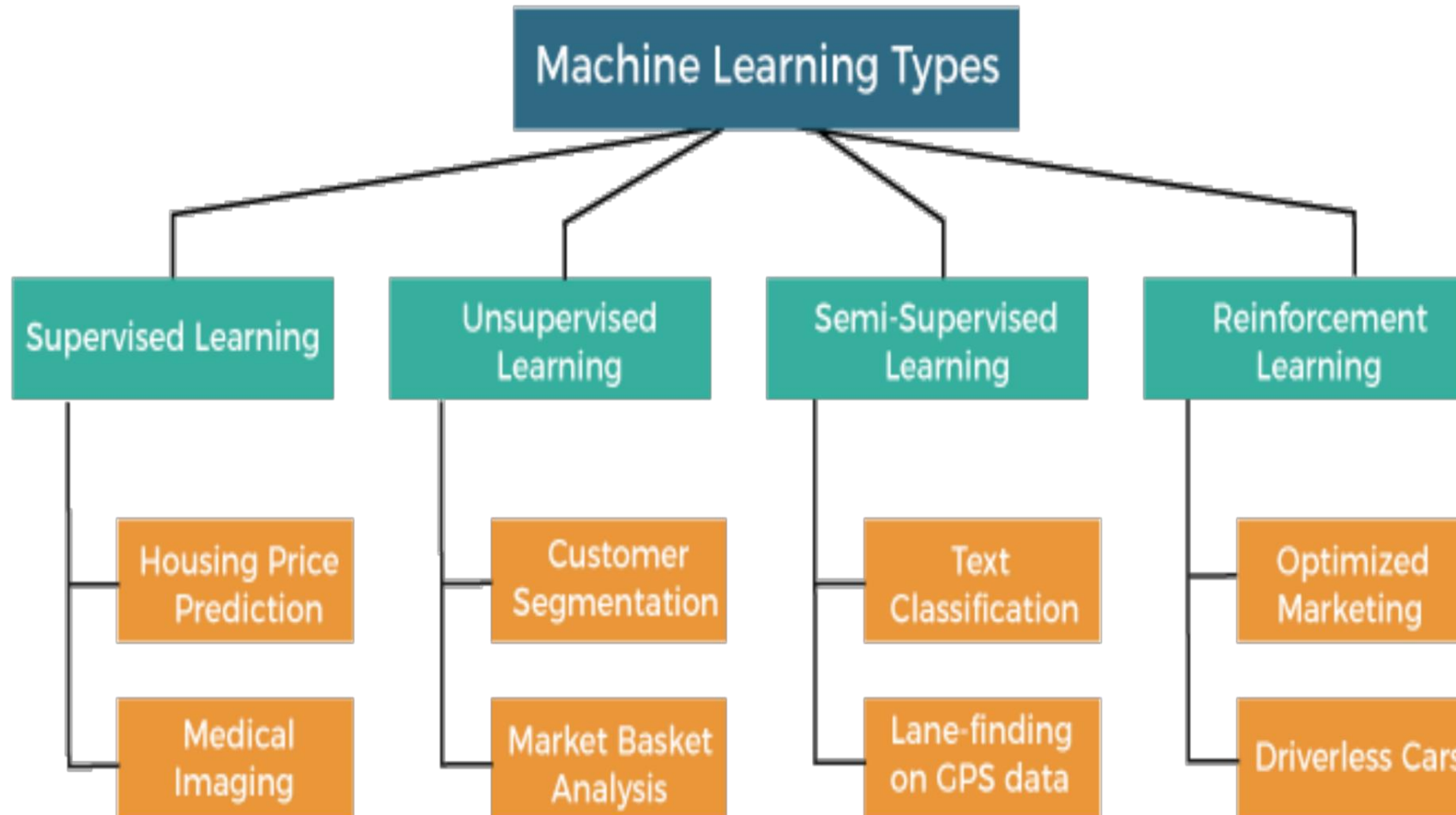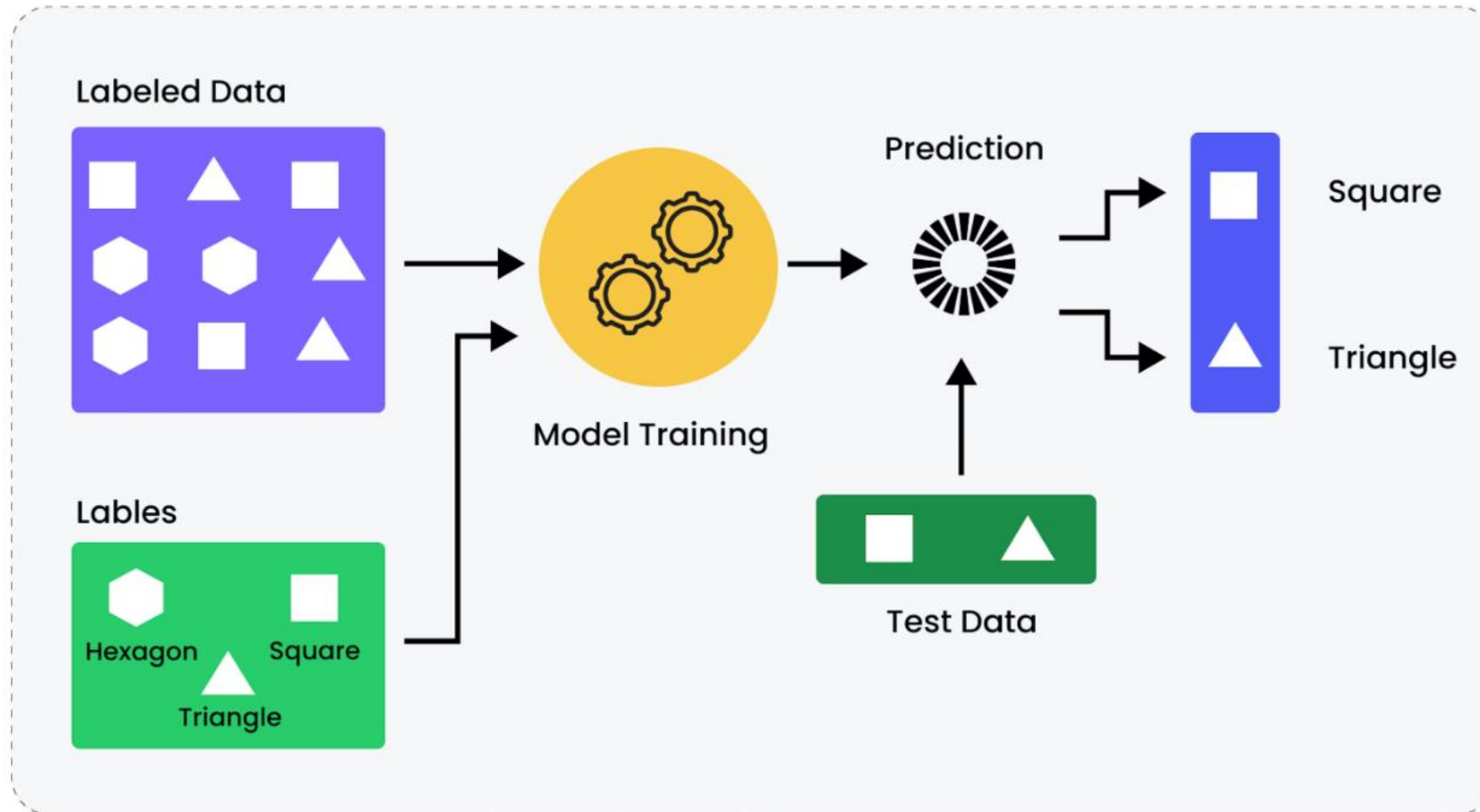
Supervised Machine Learning

- Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs.

- In supervised learning, each example in the training set is a pair consisting of an input object (typically a vector) and an output value.

- A supervised learning algorithm analyzes the training data and produces a function, which can be used for mapping new examples.

Supervised Machine Learning

- In the optimal case, the function will correctly determine the class labels for unseen instances.
- Both classification and regression problems are supervised learning problems.
- A **"supervised learning"** is so called because the process of an algorithm learning from the training dataset can be thought of as a teacher supervising the learning process.
- We know the correct answers (that is, the correct outputs), the algorithm iteratively makes predictions on the training data and is corrected by the teacher.
- Learning stops when the algorithm achieves an acceptable level of performance.

Supervised Machine Learning

Example

Consider the following data regarding patients entering a clinic. The data consists of the gender and age of the patients and each patient is labeled as "healthy" or "sick".

gender	age	label
M	48	sick
M	67	sick
F	53	healthy
M	49	healthy
F	34	sick
M	21	healthy

Based on this data, when a new patient enters the clinic, how can one predict whether he/she is healthy or sick?

Supervised Machine Learning

- In the given example, supervised learning is **to use this labeled data** to train a model that can **predict the label ("healthy" or "sick")** for new patients **based on their gender and age.**

- For example if a new patient i.e. Male with 25 years old visits the clinic, model can classify whether the patient is "healthy" or "sick" based on the patterns it learned during training.("healthy")

Supervised Machine Learning

- The main goal of the supervised learning technique is to map the input variable(x) with the output variable(y).
- Some real-world applications of supervised learning are Risk Assessment, Fraud Detection, Spam filtering, etc.

- **Two Categories:**
  - Classification
  - Regression

Supervised Machine Learning

- **Classification:**
  - **Binary Classification:**
    - When we have to categorize given data into 2 distinct classes.
    - Example – On the basis of given health conditions of a person, we have to determine whether the person has a certain disease or not.
  - **Multiclass Classification:**
    - The number of classes is more than 2.
    - For Example – On the basis of data about different species of flowers, we have to determine which specie our observation belongs.
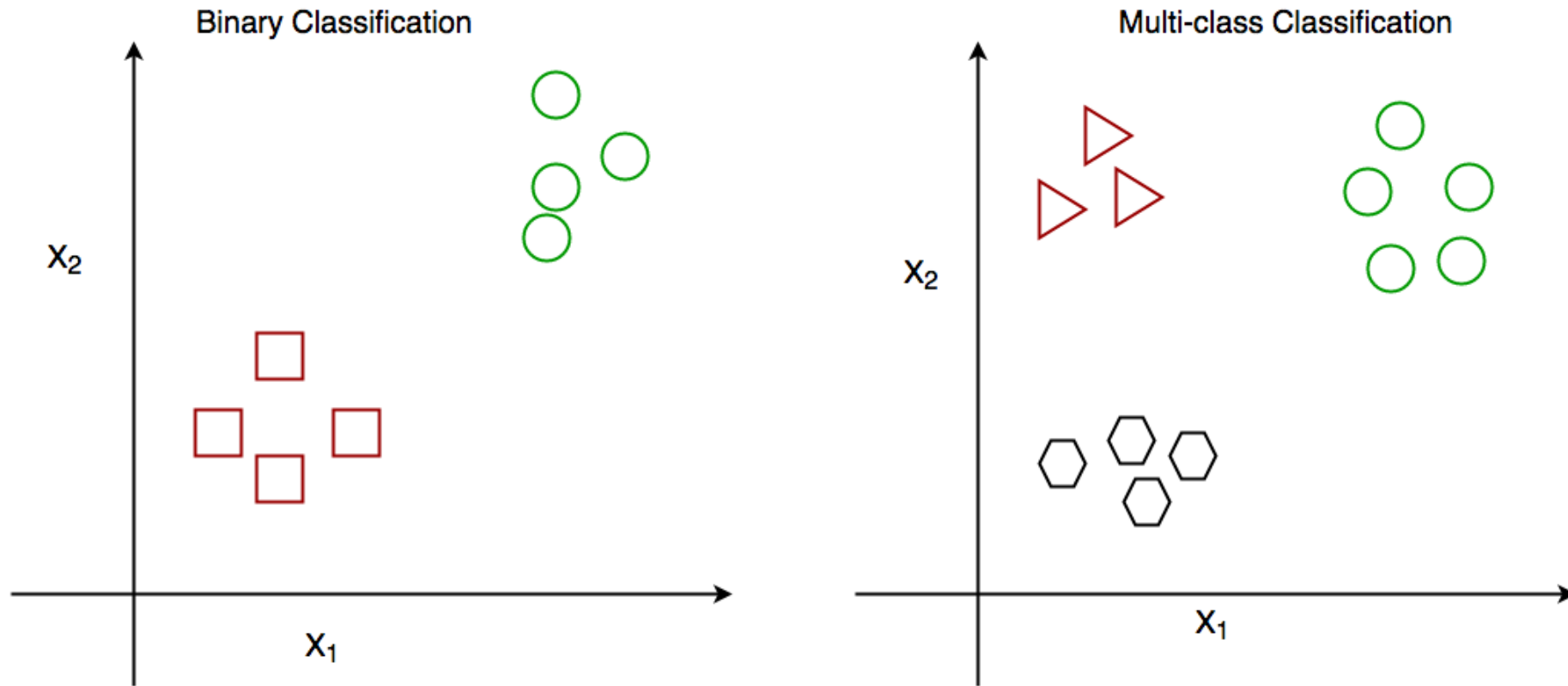
Supervised Machine Learning

Figure 1: Binary and Multiclass Classification. Here x1 and x2 are the variables upon which the class is predicted.

Supervised Machine Learning

- **There are various types of classifiers:**

  - Linear Classifiers: Logistic Regression
  - Tree-Based Classifiers: Decision Tree Classifier
  - Support Vector Machines
  - Artificial Neural Networks
  - Bayesian Regression
  - Gaussian Naive Bayes Classifiers
  - Stochastic Gradient Descent (SGD) Classifier
  - Ensemble Methods: Random Forests, AdaBoost, Bagging Classifier, Voting Classifier, ExtraTrees Classifier

- **Regression:**
  - Regression algorithms are used to solve regression problems in which there is a **linear relationship between input and output variables**.
  - These are used to predict continuous output variables, such as market trends, weather prediction, etc.
  - Types:
    - Simple Linear Regression Algorithm
    - Multivariate Regression Algorithm
    - Decision Tree Algorithm
    - Lasso Regression

Supervised Machine Learning

- **Advantages:**
  - Since supervised learning work with the labelled dataset so we can have an exact idea about the classes of objects.
  - These algorithms are helpful in predicting the output on the basis of prior experience.

- **Disadvantages:**
  - These algorithms are not able to solve complex tasks.
  - It may predict the wrong output if the test data is different from the training data.
  - It requires lots of computational time to train the algorithm.

Unsupervised Machine Learning

- Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of **input data without labeled responses.**

- In unsupervised learning algorithms, a classification or categorization is not included in the observations.

- There are **no output values** and so there is **no estimation of functions**.

- Since the examples given to the learner are **unlabeled**, the accuracy of the structure that is **output** by the algorithm **cannot be evaluated**.

- The most common unsupervised learning method is **cluster analysis**, which is used for exploratory data analysis to find hidden patterns or grouping in data.

Unsupervised Machine Learning
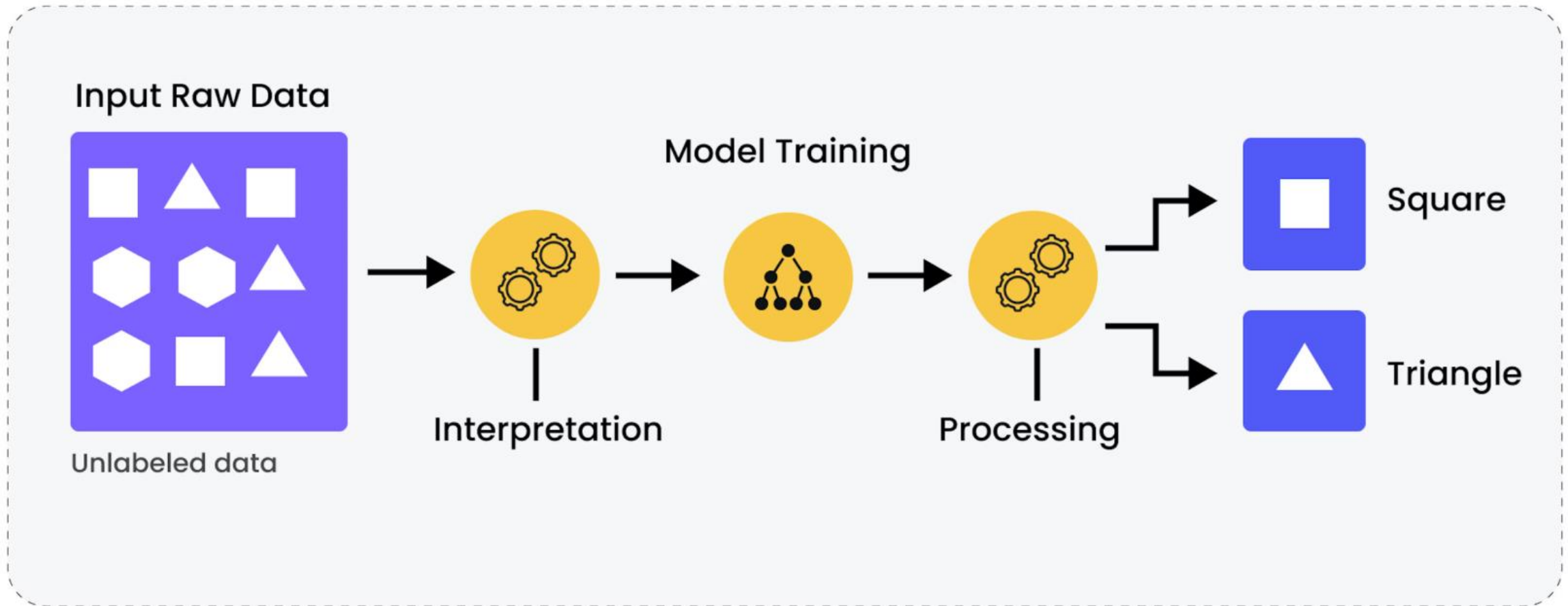
Example

Consider the following data regarding patients entering a clinic. The data consists of the gender and age of the patients.

gender	age
M	48
M	67
F	53
M	49
F	34
M	21

Based on this data, can we infer anything regarding the patients entering the clinic?

Unsupervised Machine Learning

- In the given example, unsupervised learning looks for **patterns or groups** within the data on its own.
- For example it might cluster patients by age or gender and grouping them into categories like "younger healthy patients" or "older patients" without knowing their health status.

Unsupervised Machine Learning

- The machine is trained using the unlabeled dataset, and the machine predicts the output without any supervision.
- The models are trained with the data that is neither classified nor labeled, and the model acts on that data without any supervision.
- The main aim of the unsupervised learning algorithm is to group or categories the unsorted dataset according to the similarities, patterns, and differences.
- Machines are instructed to find the hidden patterns from the input dataset.

- **Two Categories:**
  - Clustering
  - Association

Unsupervised Machine Learning

- **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.
    - Cluster analysis finds the commonalities between the data objects and categorizes them as per the presence and absence of those commonalities.

Unsupervised Machine 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 occurs together in the dataset.
  - Association rule makes 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 Machine Learning

- **Popular Algorithms:**
  - K-means clustering
  - KNN (k-nearest neighbors)
  - Hierarchal clustering
  - Anomaly detection
  - Neural Networks
  - Principle Component Analysis
  - Independent Component Analysis
  - Apriori algorithm
  - Singular value decomposition

Semi-supervised Machine Learning

- Semi-Supervised learning is a type of Machine Learning algorithm that represents the intermediate ground between Supervised and Unsupervised learning algorithms.

- It uses the combination of labeled and unlabeled datasets during the training period.

- Semi-Supervised learning is a type of Machine Learning algorithm that represents the intermediate ground between Supervised and Unsupervised learning algorithms.

- It uses the combination of labeled and unlabeled datasets during the training period.

Semi-supervised Machine Learning

- The basic disadvantage of supervised learning is that it requires hand-labeling by ML specialists or data scientists, and it also requires a high cost to process.
- Further unsupervised learning also has a limited spectrum for its applications.
- To overcome these drawbacks of supervised learning and unsupervised learning algorithms, the concept of Semi-supervised learning is introduced.
- Example:
  - Text Categorization in Newspaper as Sports, Bollywood, Education etc.

Semi-supervised Machine Learning

- Example:
- Handwritten text recognition. It allows models to adapt to different styles of handwritten text, improving their ability to recognize and interpret different variants of letter and word writing by using methods like Variational Autoencoders (VAE) that can generate new data samples based on learned features.
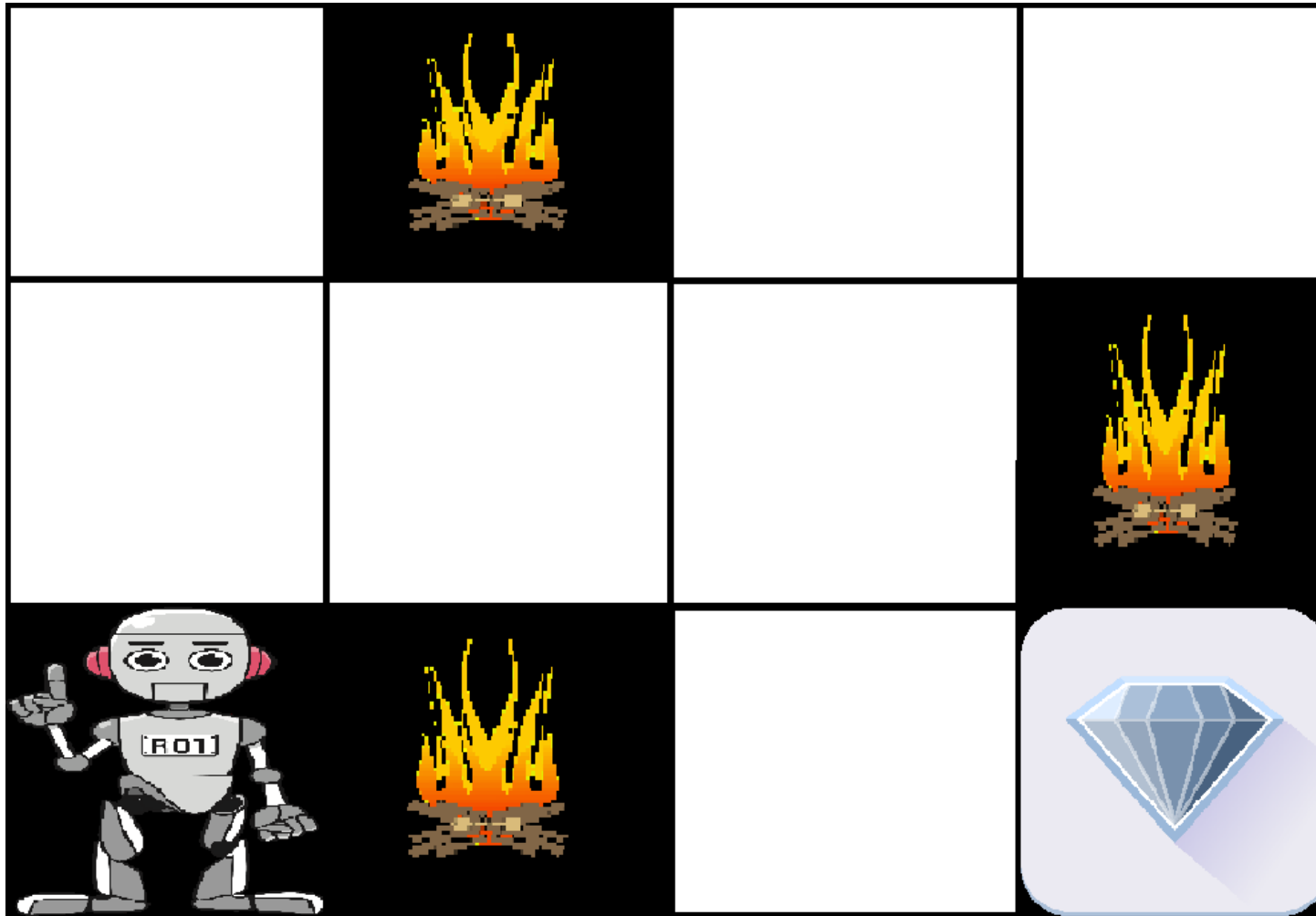
Reinforcement Machine Learning

- Reinforcement Learning is a **feedback-based** Machine learning technique in which an **agent learns to behave** in an environment by performing the actions and seeing the results of actions.

- For each good action, the agent gets positive feedback, and for each bad action, the agent gets negative feedback or penalty.

- In Reinforcement Learning, the agent learns automatically using feedbacks without any labeled data, unlike supervised learning.

Reinforcement Machine 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 problem is as follows: We have an agent and a reward, with many hurdles in between.
- The agent is supposed to find the best possible path to reach the reward.

- **Applications:**

- Robotics: Robots with pre-programmed behavior are useful in structured environments, such as the assembly line of an automobile manufacturing plant, where the task is repetitive in nature.

- A master chess player makes a move. The choice is informed both by planning, anticipating possible replies and counter replies.

- An adaptive controller adjusts parameters of a petroleum refinery's operation in real time.

References

- Ethem Alpaydin, Introduction to Machine Learning, PHI 4th Edition-2020, The MIT Press,ISBN:9780262043793.
- Peter Flach, "Machine Learning The Art and Science of Algorithms that Make Sense of Data", Cambridge University Press India. ISBN 13: 9781107422223

Thank You!!!

Happy Learning!!!