Machine Learning

What is Machine Learning?

- **Machine Learning** is the field of study that gives computers the capability to learn without being explicitly programmed.

- ***Machine learning (ML) is defined as a discipline of artificial intelligence (AI) that provides machines the ability to automatically learn from data and past experiences to identify patterns and make predictions with minimal human intervention.***

- Machine learning is focused on building systems that can learn from historical data, identify patterns, and make logical decisions with little to no human intervention

What is Machine Learning?

- Machine Learning(ML) can be explained as automating and improving the learning process of computers based on their experiences without being actually programmed.

- The process starts with feeding good quality data and then training our machines(computers) by building machine learning models using the data and different algorithms. The choice of algorithms depends on what type of data do we have and what kind of task we are trying to automate.
- **Example: Training of students during exams.**

What is learning?

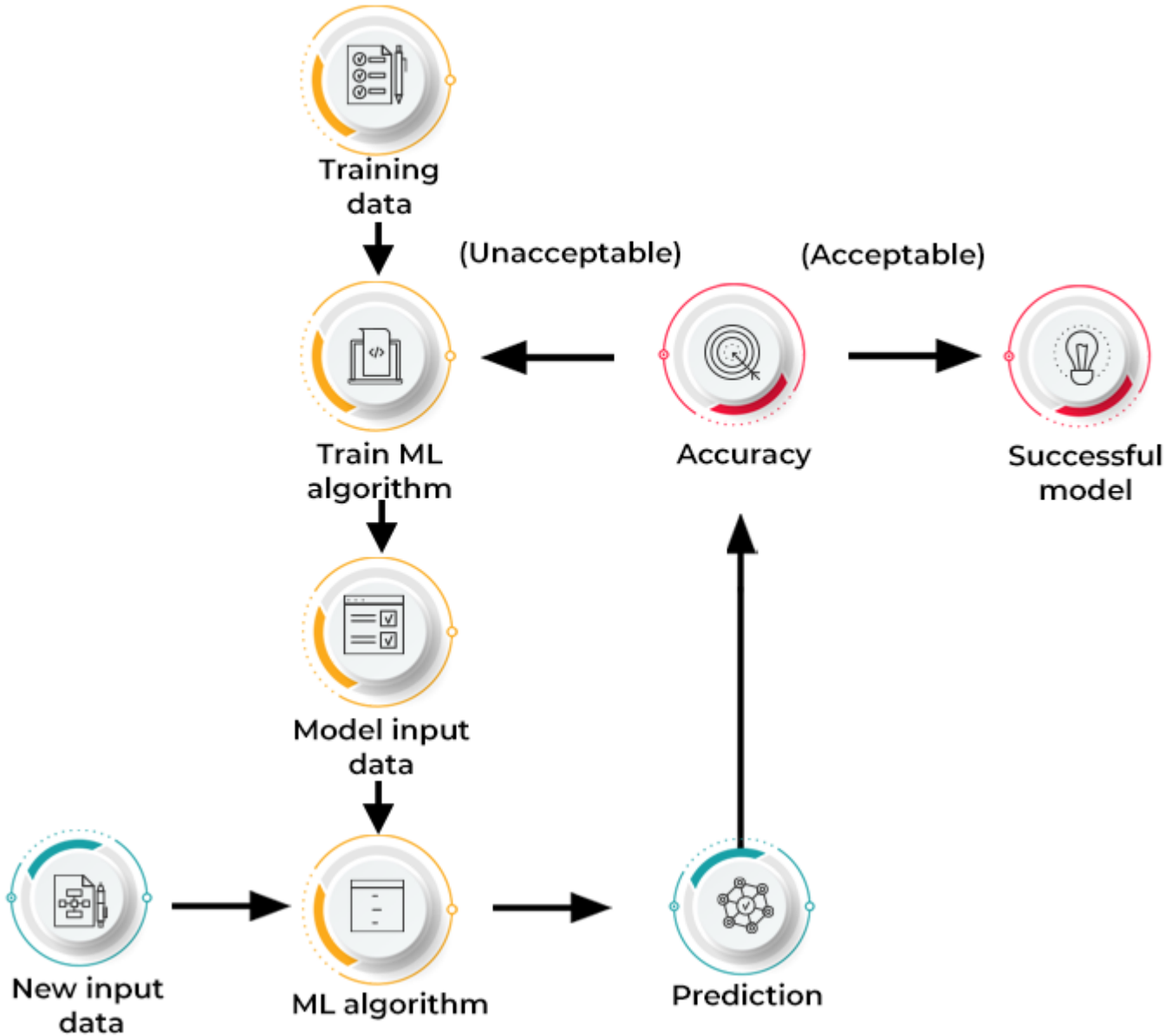
- A computer program is said to *learn* from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks T, as measured by P , improves with experience E.
- Handwriting recognition learning problem
  - Task T : Recognizing and classifying handwritten words within images
  - Performance P : Percent of words correctly classified
  - Training experience E : A dataset of handwritten words with given classifications

- The quality of a machine learning model is dependent on two major aspects:
- 1. The quality of the input data: “garbage in, garbage out”
  - **Volume**
  - **Variety**
  - **Velocity**
  - **Value**
  - **Veracity**
- 2. The model choice itself.

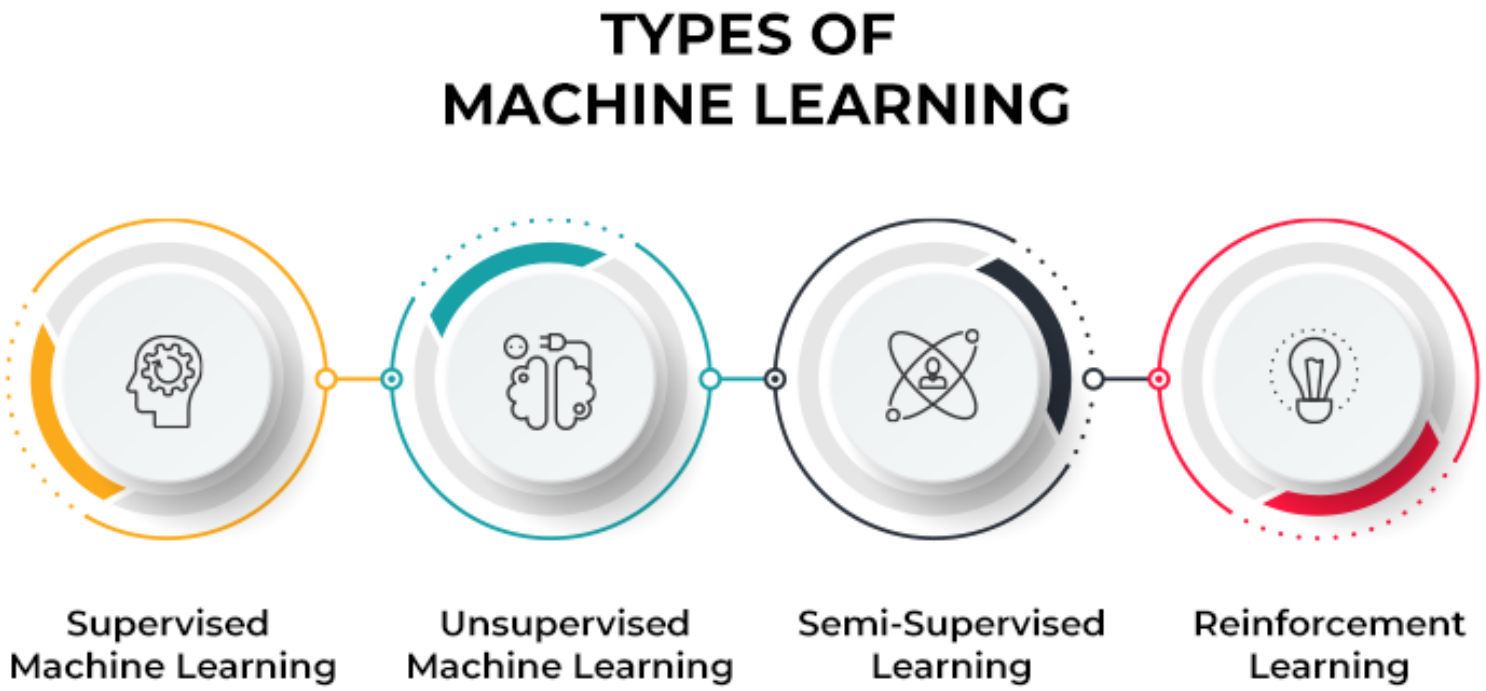
Why Machine Learning?

- Flood of available data(with the advent of internet)
- Increasing computational power
- Increasing support from industries(**trends in customer behavior and business operational patterns, as well as supports the development of new products)**

HOW DOES MACHINE LEARNING WORK?

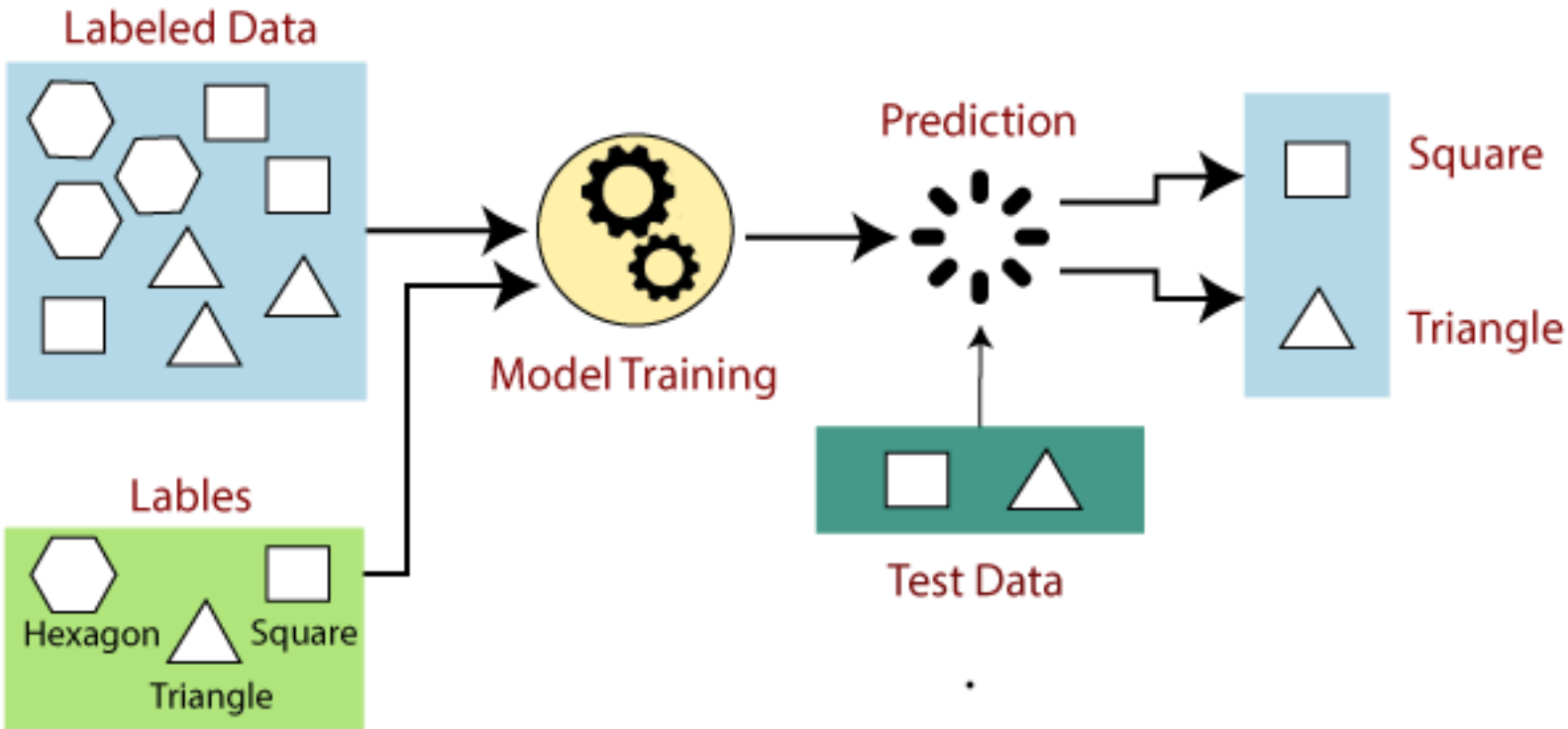


Types of Machine Learning



Supervised machine learning

- The model or algorithm is presented with example inputs and their desired outputs and then finding patterns and connections between the input and the output.
- The goal is to learn a general rule that maps inputs to outputs. The training process continues until the model achieves the desired level of accuracy on the training data.


Labeled Data
Lables
Hexagon
Square
Triangle
Model Training
Prediction
Test Data
Square
Triangle

Steps Involved in Supervised Learning

- First Determine the type of training dataset
- Collect/Gather the labelled training data.
- Split the training dataset into training **dataset, test dataset, and validation dataset**.
- Determine the input features of the training dataset, which should have enough knowledge so that the model can accurately predict the output.
- Determine the suitable algorithm for the model, such as support vector machine, decision tree, etc.
- Execute the algorithm on the training dataset. Sometimes we need validation sets as the control parameters, which are the subset of training datasets.
- Evaluate the accuracy of the model by providing the test set. If the model predicts the correct output, which means our model is accurate.

Types of Supervised Machine Algorithm

- **Regression**: It is used for the prediction of continuous variables, such as Weather forecasting, Market Trends, etc.

- Algorithms:
  - Linear Regression
  - Regression Trees
  - Non-Linear Regression
  - Bayesian Linear Regression
  - Polynomial Regression

Types of Supervised Machine Algorithm

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

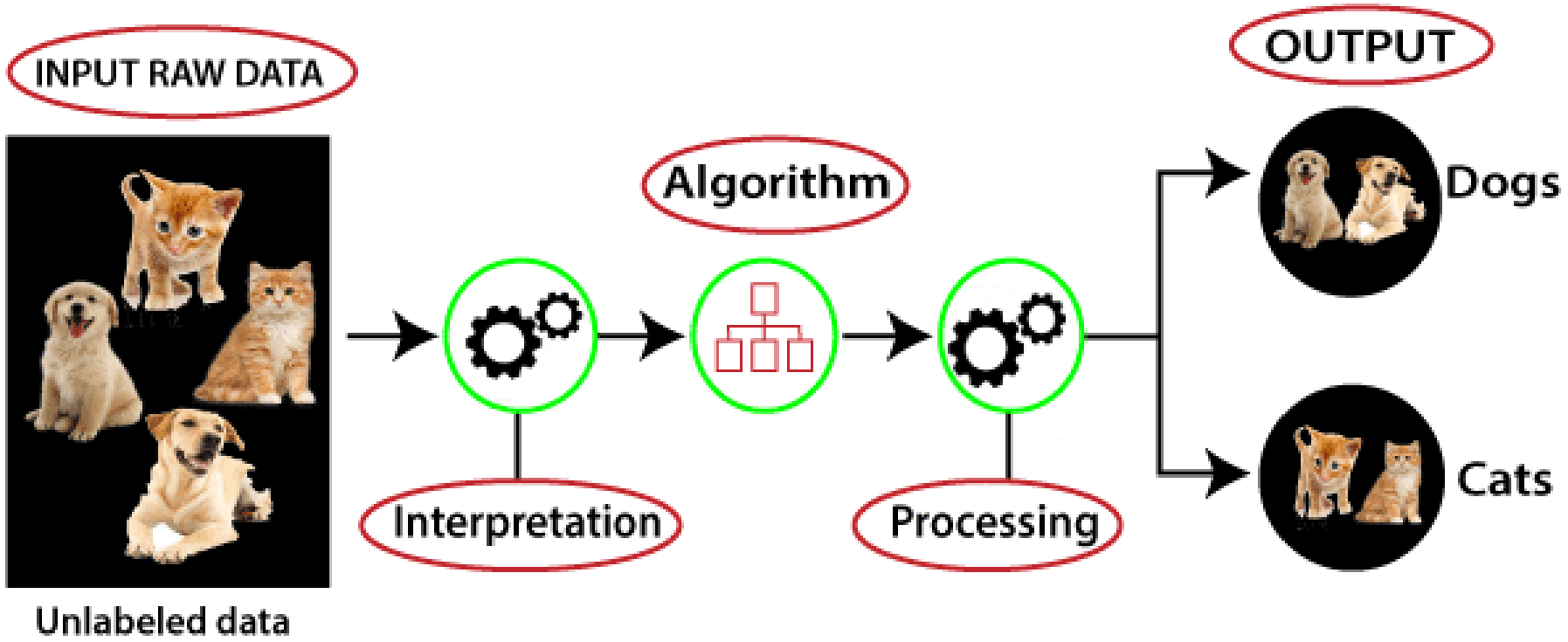
- Algorithms:
  - Random Forest
  - Decision Trees
  - Logistic Regression
  - Support vector Machines

Unsupervised machine learning

- No labels are given to the learning algorithm, leaving it on its own to find structure in its input.
- Models itself find the hidden patterns and insights from the given data. It can be compared to learning which takes place in the human brain while learning new things.
- **Clustering**: A clustering problem is where you want to discover the inherent groupings in the data, such as grouping customers by purchasing behavior.
- **Association**: An association rule learning problem is where you want to discover rules that describe large portions of your data, such as people that buy X also tend to buy Y.

Why Unsupervised Learning?

- Unsupervised learning is helpful for **finding useful insights** from the data.
- Unsupervised learning is much similar as a human learns to think by their own experiences, which makes it closer to the real AI.
- Unsupervised learning works on unlabeled and uncategorized data which make unsupervised learning more important.
- In real-world, we do not always have input data with the corresponding output so to solve such cases, we need unsupervised learning.


INPUT RAW DATA
Unlabeled data
Interpretation
Algorithm
Processing
OUTPUT
Dogs
Cats

Unsupervised Learning Algorithms

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

Supervised vs. Unsupervised learning

Feature	**Supervised learning**	**Unsupervised learning**
Labeled/Unlabeled Data	Algorithms are trained using labeled data.	Algorithms are trained using unlabeled data.
Data with/without Output	Input data is provided to the model along with the output.	Only input data is provided to the model.
Groups	Classes are already defined in data using labels	Groups are made based on similar characteristics.
Feedback of output	Supervised learning model takes direct feedback to check if it is predicting correct output or not.	Unsupervised learning model does not take any feedback.

Supervised vs. Unsupervised learning

Feature	**Supervised learning**	**Unsupervised learning**
Predicts output/pattern	Supervised learning model predicts the output.	Unsupervised learning model finds the hidden patterns in data.
Goal	The goal of supervised learning is to train the model so that it can predict the output when it is given new data.	The goal of unsupervised learning is to find the hidden patterns and useful insights from the unknown dataset.
Accuracy of results	Supervised learning model produces an accurate result.	Unsupervised learning model may give less accurate result as compared to supervised learning.

Supervised vs. Unsupervised learning

Feature	**Supervised learning**	**Unsupervised learning**
Algorithms	Linear Regression, Logistic Regression, Support Vector Machine, Multi-class Classification, Decision tree, Bayesian Logic, etc.	It includes various algorithms such as Clustering, KNN, and Apriori algorithm.

Semi-supervised machine learning

- This approach to machine learning involves a mix of the two preceding types. Data scientists may feed an algorithm mostly labeled training data, but the model is free to explore the data on its own and develop its own understanding of the data set.

- **Machine translation:** Teaching algorithms to translate language based on less than a full dictionary of words.

- **Fraud detection:** Identifying cases of fraud when you only have a few positive examples.

- **Labelling data:** Algorithms trained on small data sets can learn to apply data labels to larger sets automatically.

Reinforcement learning

- Reinforcement learning works by programming an algorithm with a distinct goal and a prescribed set of rules for accomplishing that goal.

- Data scientists also program the algorithm to seek positive rewards -- which it receives when it performs an action that is beneficial toward the ultimate goal -- and avoid punishments -- which it receives when it performs an action that gets it farther away from its ultimate goal.

- 1) Value-based
- 2) Policy-based and
- 3) Model based learning

Reinforcement learning

- **Robotics:** Robots can learn to perform tasks the physical world using this technique.
- **Video gameplay:** Reinforcement learning has been used to teach bots to play a number of video games.
- **Resource management:** Given finite resources and a defined goal, reinforcement learning can help enterprises plan out how to allocate resources.

Issues in Machine learning

- Lack Of Quality Data: noisy, incorrect, incomplete data
- Getting bad recommendations : proposal engines with complex algorithms may tend to provide wrong results
- Talent deficit: rare experts are available
- Implementation: slow deployment, data security, lack of data
- Making the wrong assumptions
- Having algorithms become obsolete when data grows
- Absence of skilled resources
- Complexity

Application of Machine Learning

Application of Machine Learning

- **Automation**
- **Banking and Finance**
- **Transportation and Traffic Prediction**
- **Image Recognition**
- **Speech Recognition**
- **Product Recommendation**
- **Virtual Personal Assistance**
- **Email Spam and Malware detection & Filtering**
- **Language Translation**
- **Self-driving cars**

Recent trends in ML