Advanced Data Science

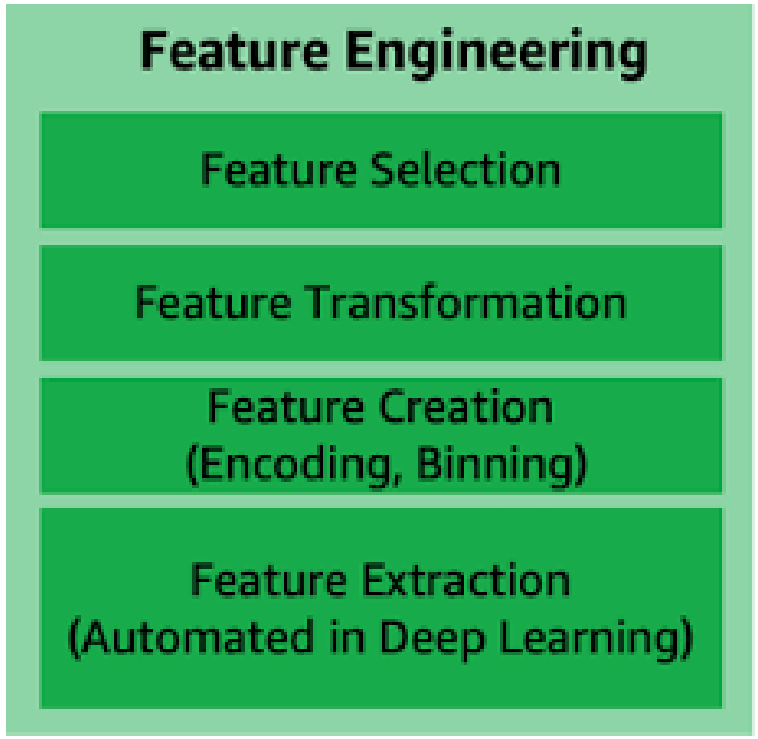
Data Modelling

Session 4

Kiran Waghmare
Program Manager
C-DAC Mumbai

Agenda

- **Data Modelling**
- **Feature Engineering**



Unit I:
Advance Methods of Data Science and Algorithms
Statistics
Central Limit Theorem;
AB Testing;
Linear Regression

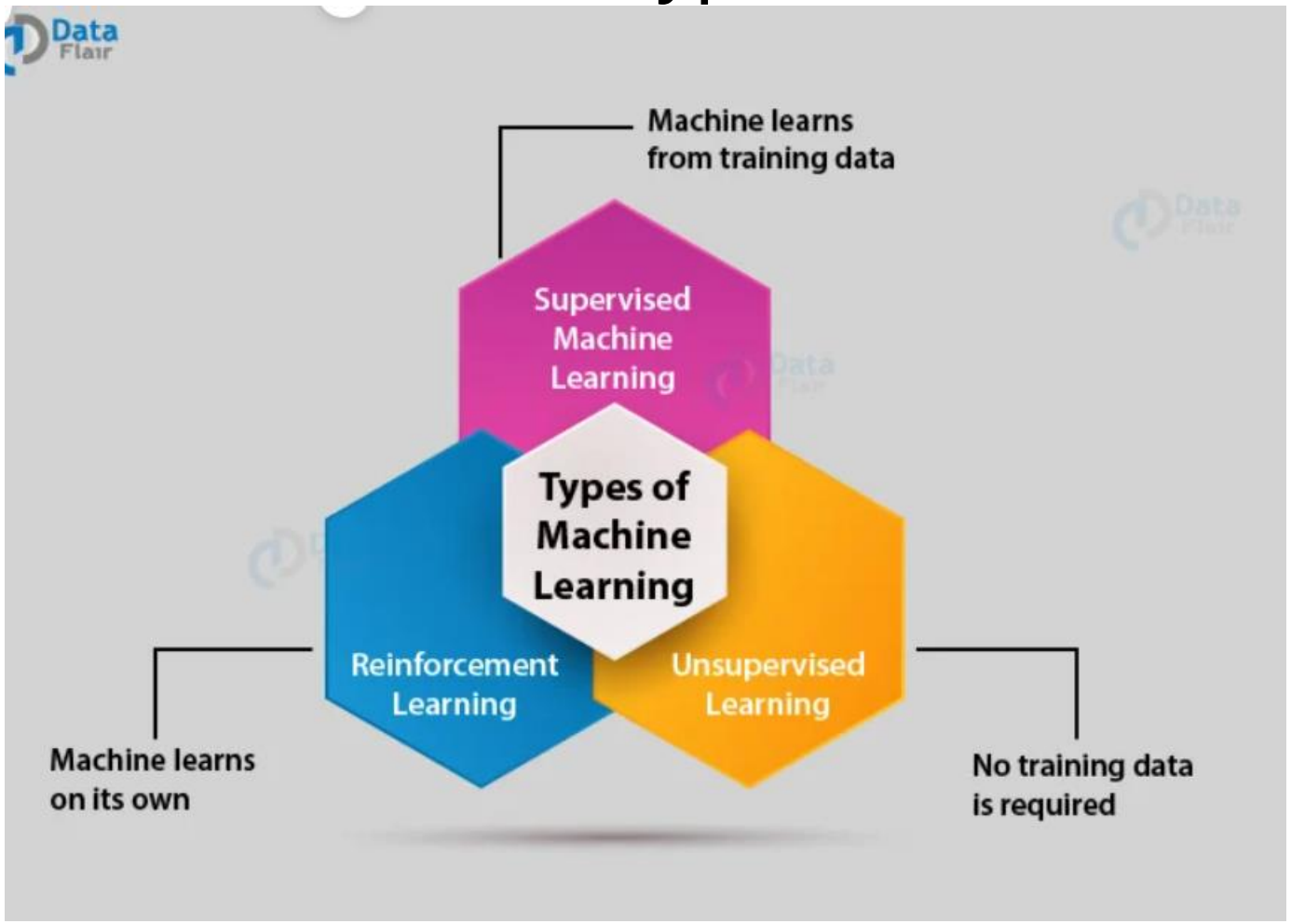
Comparison

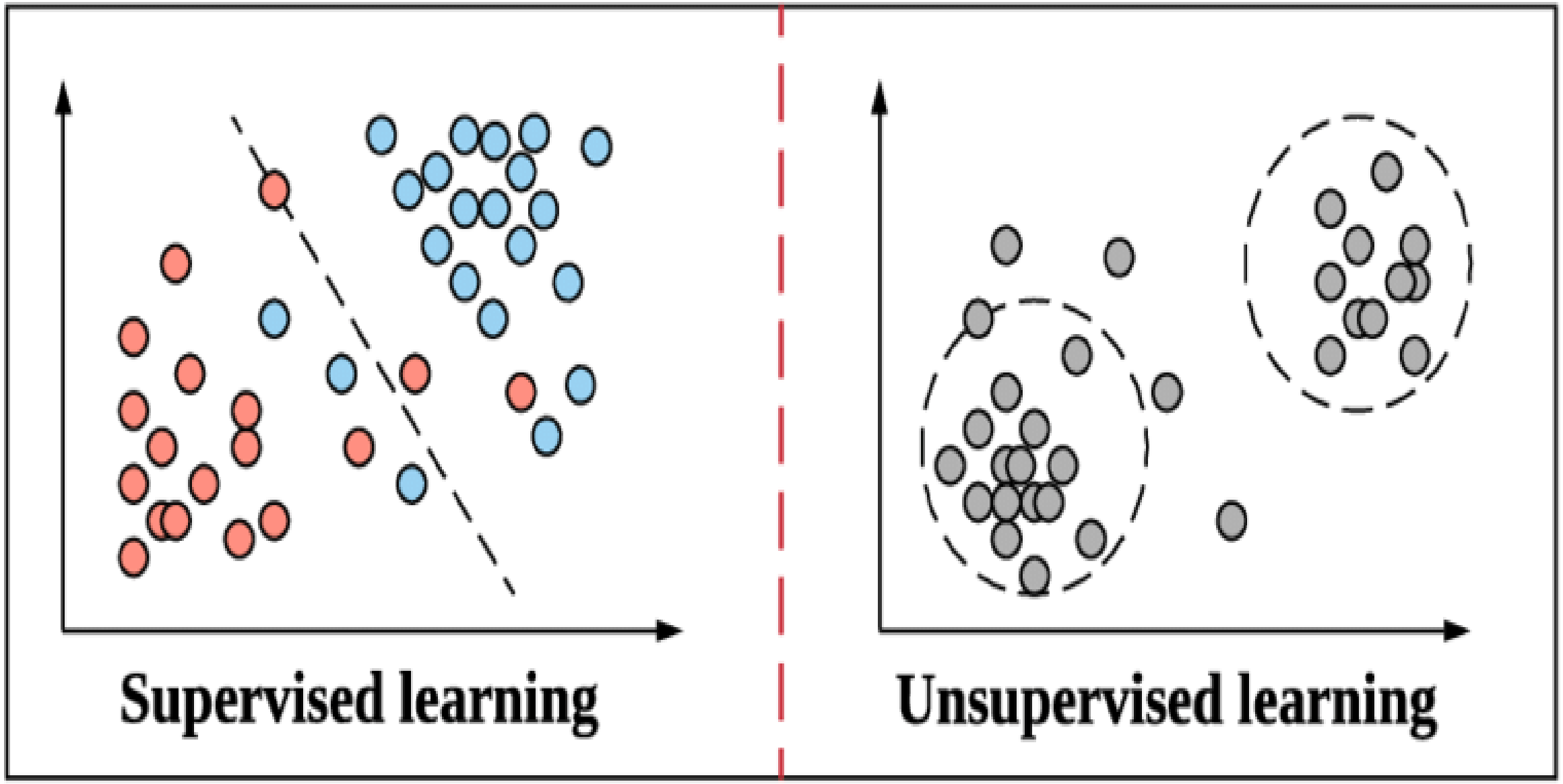
- **Traditional Programming**

- **Machine Learning**

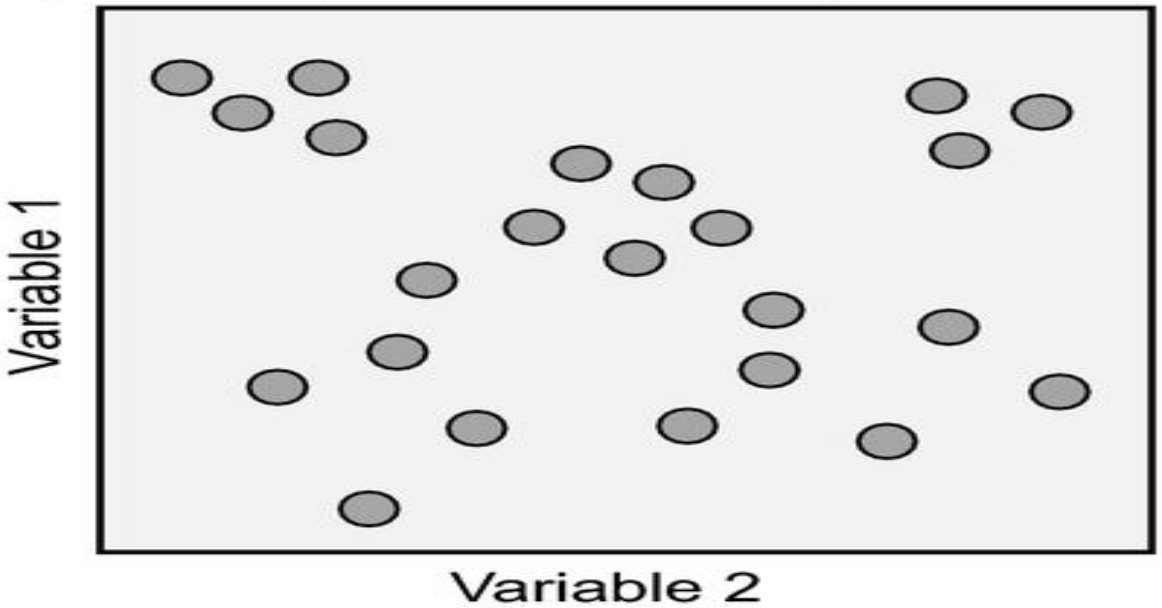
Types of Learning algorithms

- Learning Algorithms can be classified into 3 types as follows –
  - **Supervised Learning**
  - **Unsupervised Learning**
  - **Reinforcement Learning**




Supervised learning
Unsupervised learning

a) **Unsupervised learning**



Types of Machine Learning

MATER DYNAMICS

Supervised Learning

Train an algorith on a labeled data set to predict the crrect output value for unseen inputs.

- ✓ Input / Output
- ✓ Labeled data
- ✓ "Replicate the right answers"
- ✓ Classification, prediction

Unsupervised Learning

Train an algorithm to find similari-ties on abnormalities in a data set.

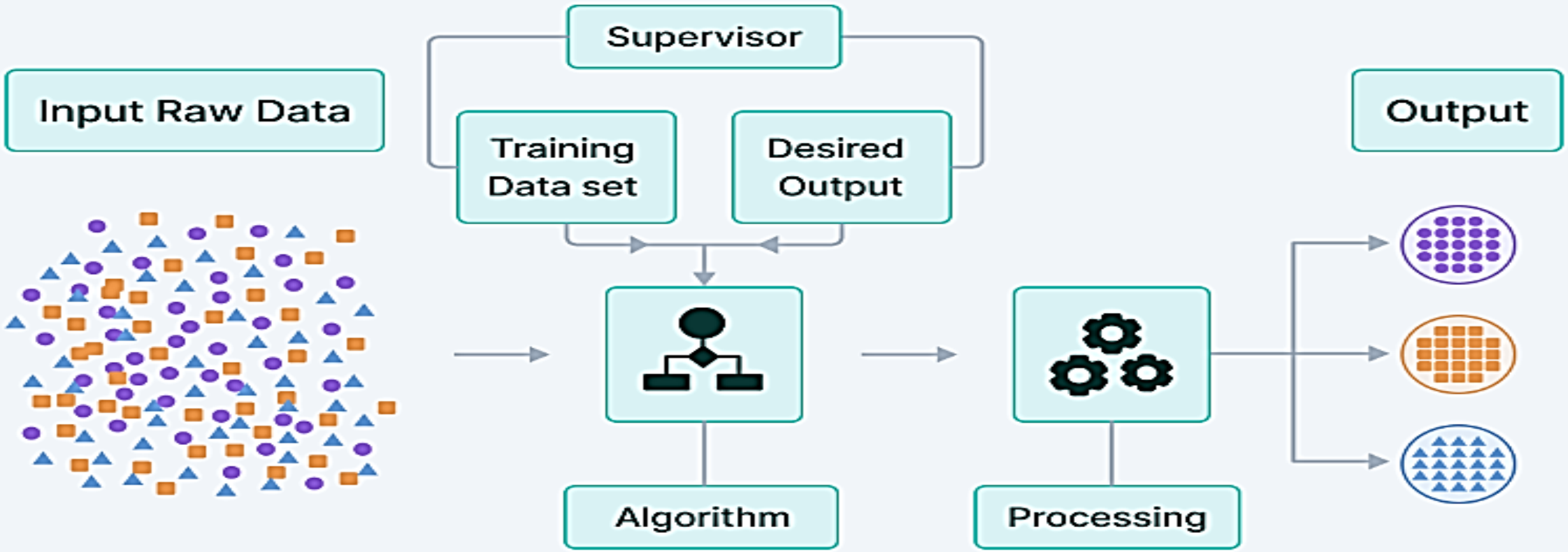
- ✓ Input
- ✓ Unlabeled data
- ✓ "Find patterns in data"
- ✓ Clustering association
- ✓ Anomaly detection, custo-

Reinforcement Learning

Learn through trial and error from interaction with an environment.

- ✓ States & actions
- ✓ No data set
- ✓ "Find actions that maximize reward"
- ✓ Decision making

Supervised Learning



Supervised Learning process

Supervised Learning

Input

Annotations

These are grapes

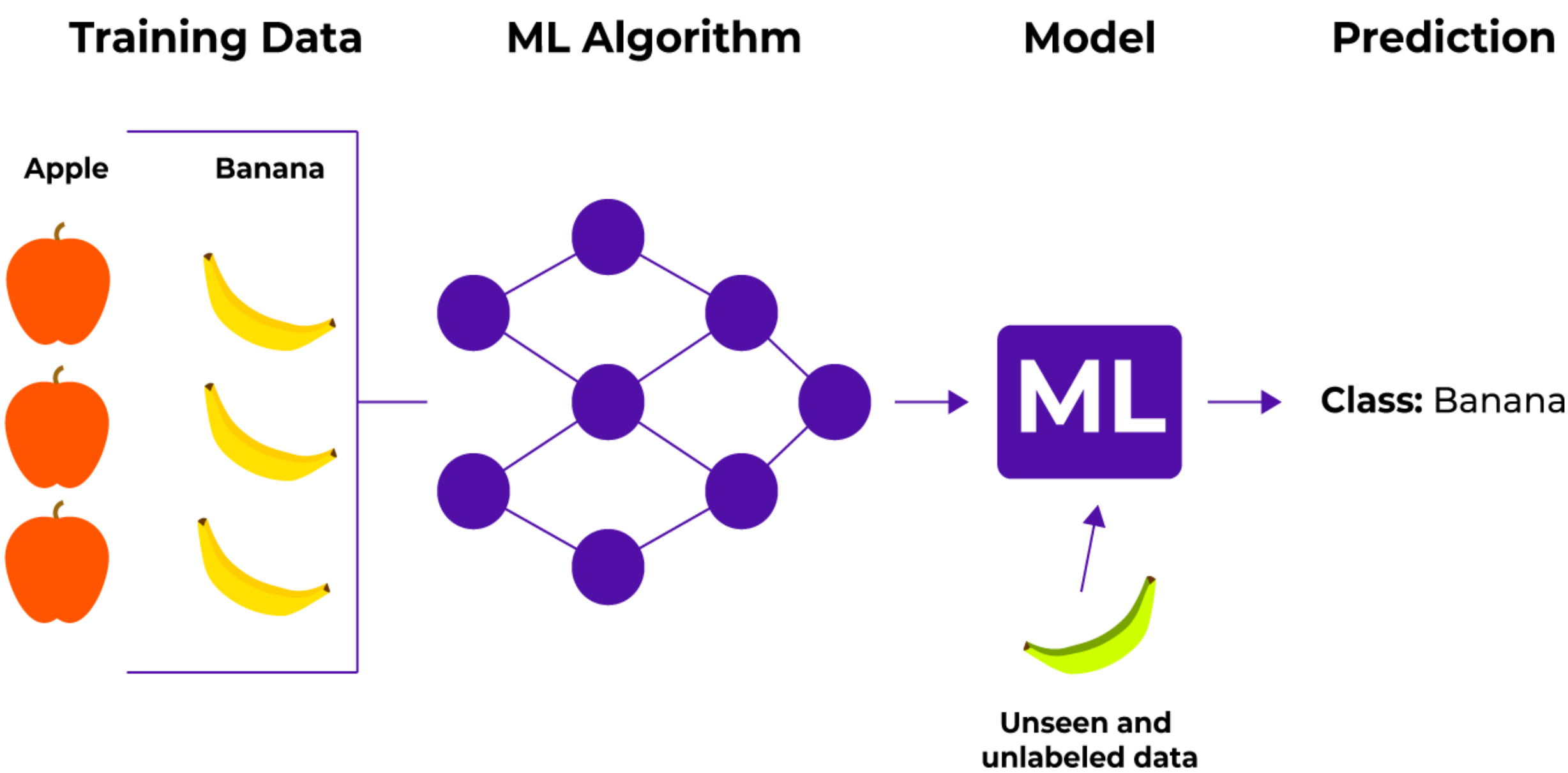
Model

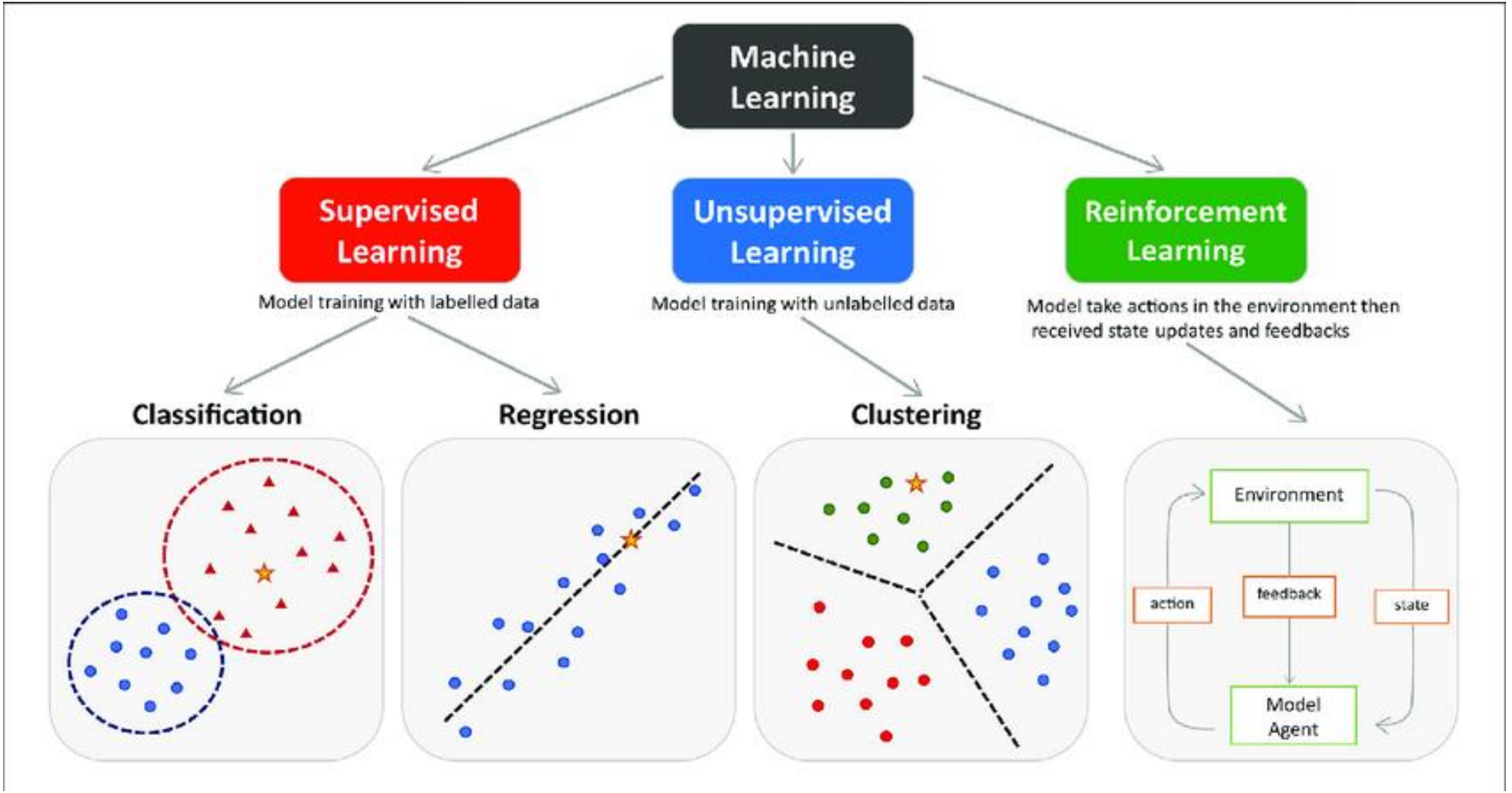
It's Grapes

Prediction

Supervised learning

- machine learning task of learning a function that maps an **input** to an **output** supported example input-output pairs.
- In Supervised Learning, the dataset on which we **train our model is labeled**. There is a clear and **distinct mapping** of input and output. Based on the example inputs, the model is able to get **trained** in the **instances**.
- An example of supervised learning is **spam filtering**.
- Based on the **labeled data**, the model is able to determine if the data is **spam** or **ham.** This is an easier form of **training**.
- Spam filtering is an example of this type of **machine learning algorithm**.


Training Data
ML Algorithm
Model
Prediction
Apple
Banana
ML
Class: Banana
Unseen and
unlabeled data


Machine Learning
Supervised Learning
Model training with labelled data
Unsupervised Learning
Model training with unlabelled data
Reinforcement Learning
Model take actions in the environment then received state updates and feedbacks
Classification
Regression
Clustering
Environment
action
feedback
state
Model Agent

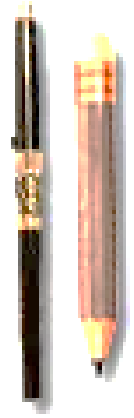
Unsupervised Learning

Input

Annotations

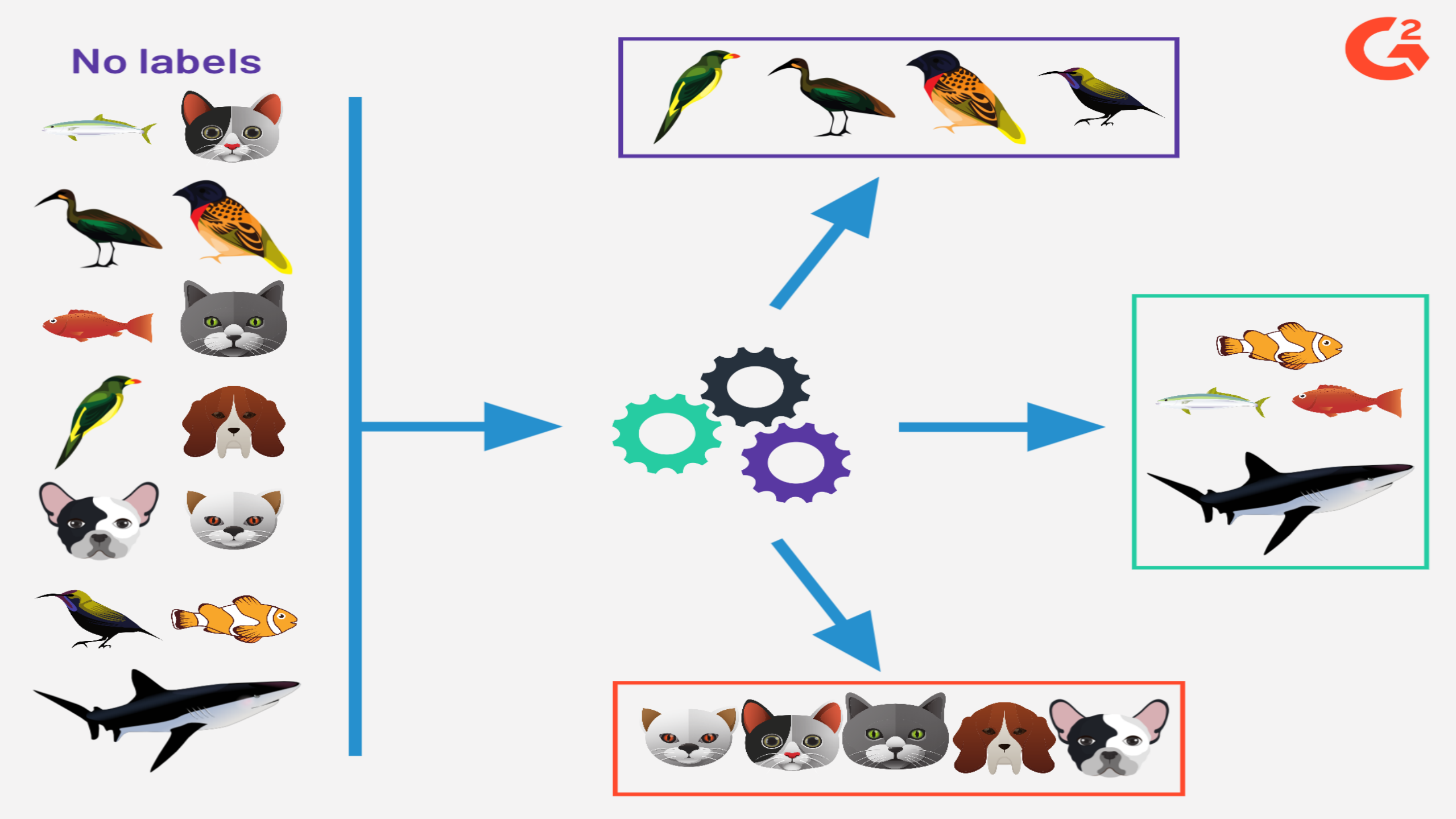
Unsorted Information

Model



Pen & Pencil

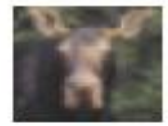
Predictions


No labels

Supervised Learning

data	label
	Dog
	Bird
	Airplane
	Deer
	Cat
	Truck
	Ship

Semi-Supervised Learning

data	label
	Dog
	Bird
	No label
	No label
	No label
	No label
	No label

Reinforcement Learning

- **Reinforcement learning is one among three basic machine learning paradigms.**
- **Reinforcement Learning is an emerging and most popular type of Machine Learning Algorithm.**
- **It is used in various autonomous systems like cars and industrial robotics.**
- **The aim of this algorithm is to reach a goal in a dynamic environment. It can reach this goal based on several rewards that are provided to it by the system.**
- **It is most heavily used in programming robots to perform autonomous actions.**
- **It is also used in making intelligent self-driving cars.**

Reinforcement Learning