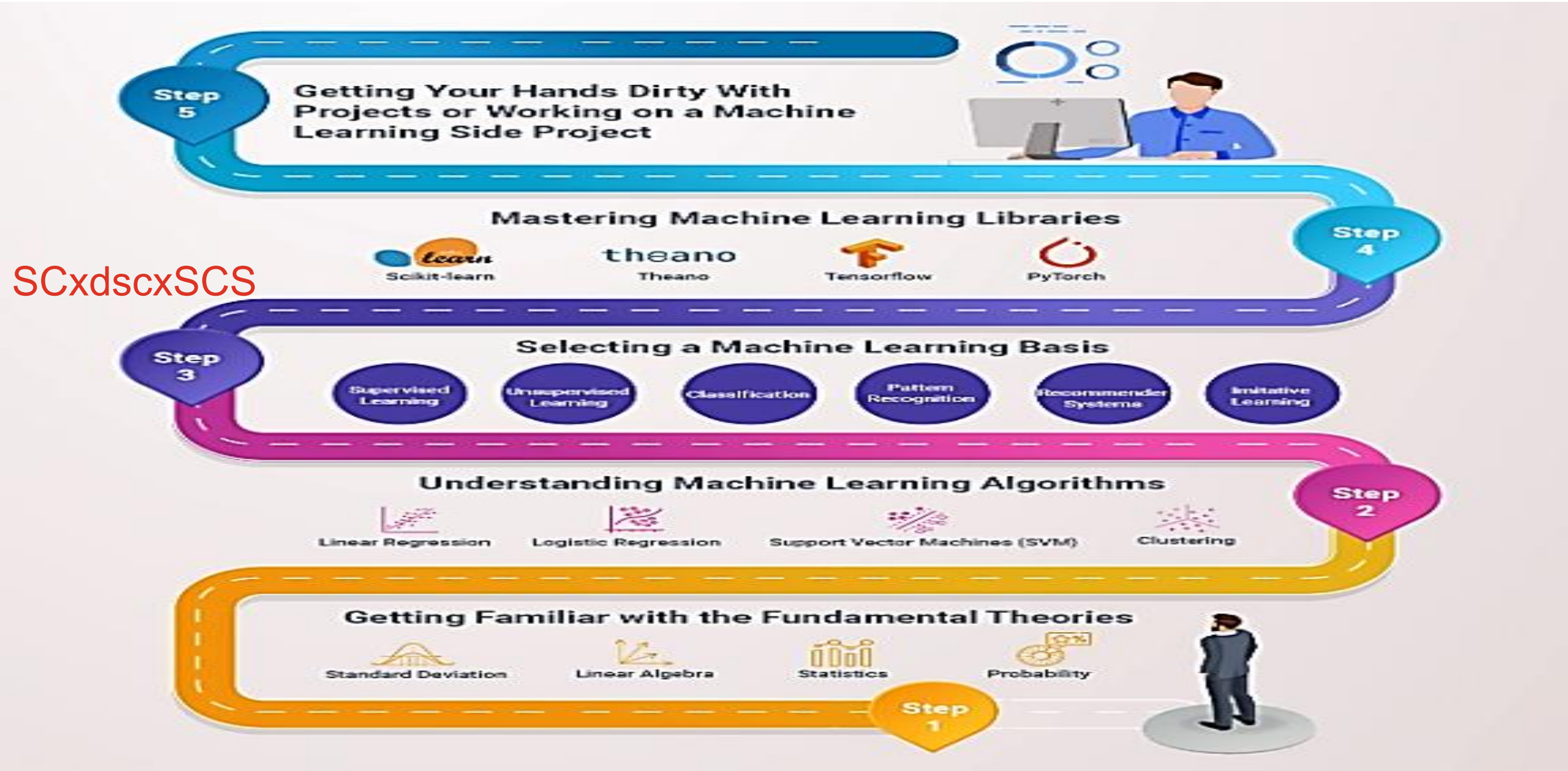
Practical Machine Learning

Day 1: MAR24 DBDA

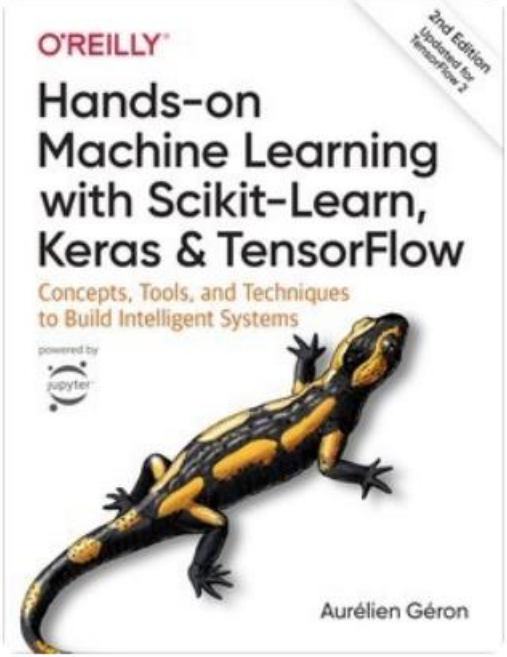
Kiran Waghmare

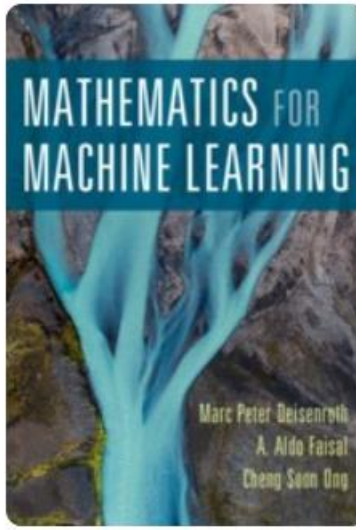
CDAC Mumbai

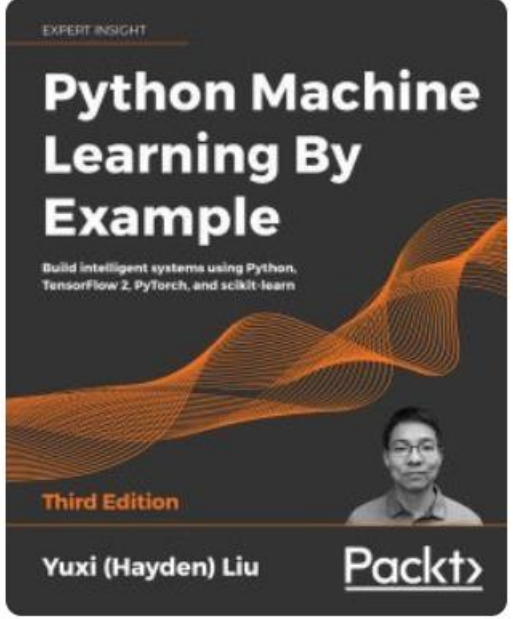
Machine Learning Roadmap

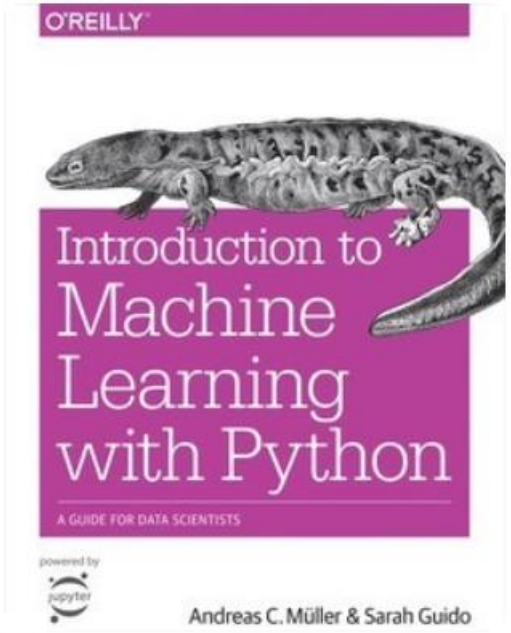


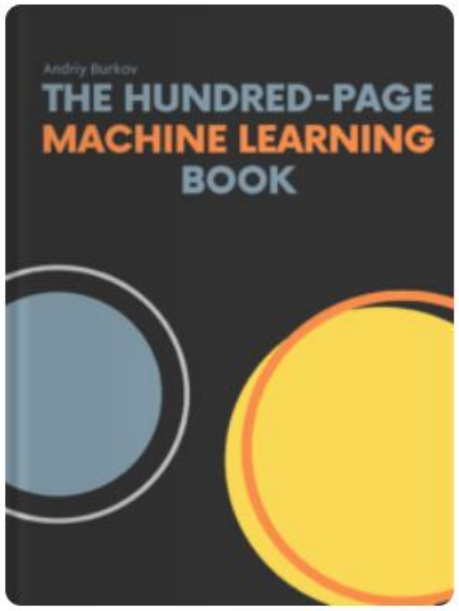
SCxdscxSCS

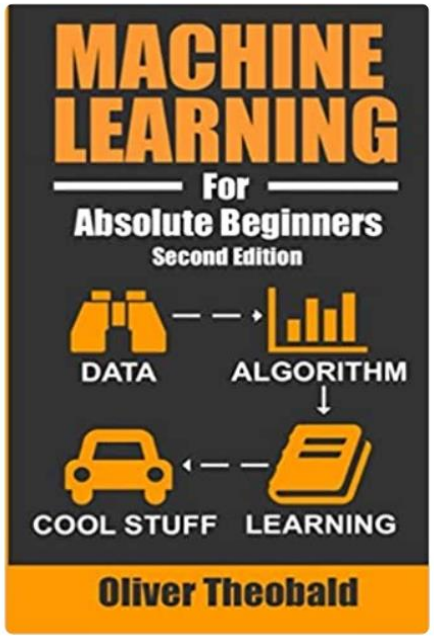

O'REILLY
2nd Edition
Updated for TensorFlow 2
Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow
Concepts, Tools, and Techniques to Build Intelligent Systems
powered by jupyter
Aurélien Géron

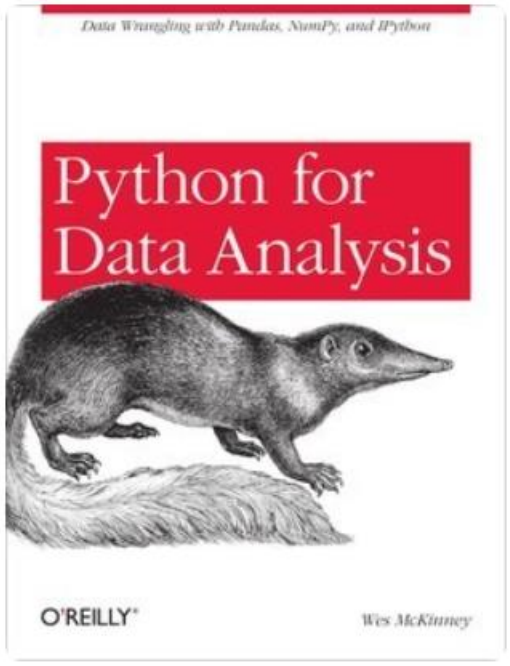

MATHEMATICS FOR MACHINE LEARNING
Marc Peter Deisenroth
A. Aldo Faisal
Cheng Soon Ong

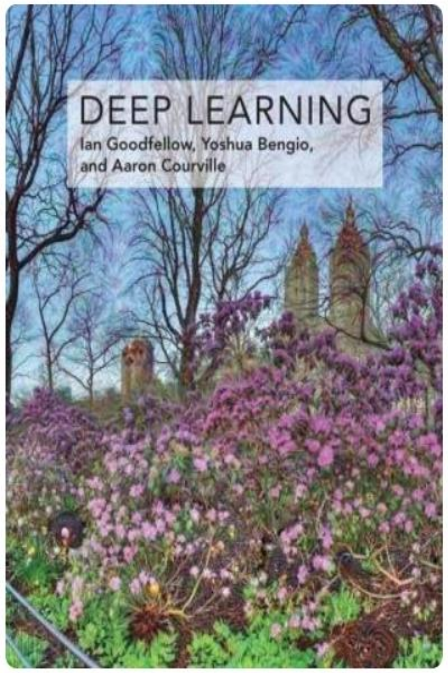

EXPERT INSIGHT
Python Machine Learning By Example
Build intelligent systems using Python, TensorFlow 2, PyTorch, and scikit-learn
Third Edition
Yuxi (Hayden) Liu
Packt


O'REILLY
Introduction to Machine Learning with Python
A GUIDE FOR DATA SCIENTISTS
powered by jupyter
Andreas C. Müller & Sarah Guido


Andriy Burkov
THE HUNDRED-PAGE MACHINE LEARNING BOOK


MACHINE LEARNING
For
Absolute Beginners
Second Edition
DATA
ALGORITHM
COOL STUFF
LEARNING
Oliver Theobald


Data Wrangling with Pandas, NumPy, and IPython
Python for Data Analysis
O'REILLY
Wes McKinney


DEEP LEARNING
Ian Goodfellow, Yoshua Bengio, and Aaron Courville

Agenda

- What is machine learning?
- Algorithm types of Machine learning
- Supervised and Unsupervised Learning
- Uses of Machine learning
- Evaluating ML techniques
- Introduction to Scikit Learn

Why Machine Learning?

- Develop systems that can automatically adapt and customize themselves to individual users.
  - Personalized news or mail filter
- Discover new knowledge from large databases (***data mining***).
  - Market basket analysis (e.g. diapers and beer)
- Ability to mimic human and replace certain monotonous tasks - which require some intelligence.
    - like recognizing handwritten characters
- Develop systems that are too difficult/expensive to construct manually because they require specific detailed skills or knowledge tuned to a specific task (knowledge engineering bottleneck).

Machine Learning

- **Herbert Alexander Simon**:
  "Learning is any process by which a system improves performance from experience."
- "Machine Learning is concerned with computer programs that automatically improve their performance through experience. "

Herbert Simon
Turing Award 1975
Nobel Prize in Economics 1978

What is Machine Learning?

- [Arthur Samuel, 1959]
  - Field of study that gives computers
  - the ability to learn without being explicitly programmed

- [Kevin Murphy] algorithms that
  - automatically detect patterns in data
  - use the uncovered patterns to predict future data or other outcomes of interest

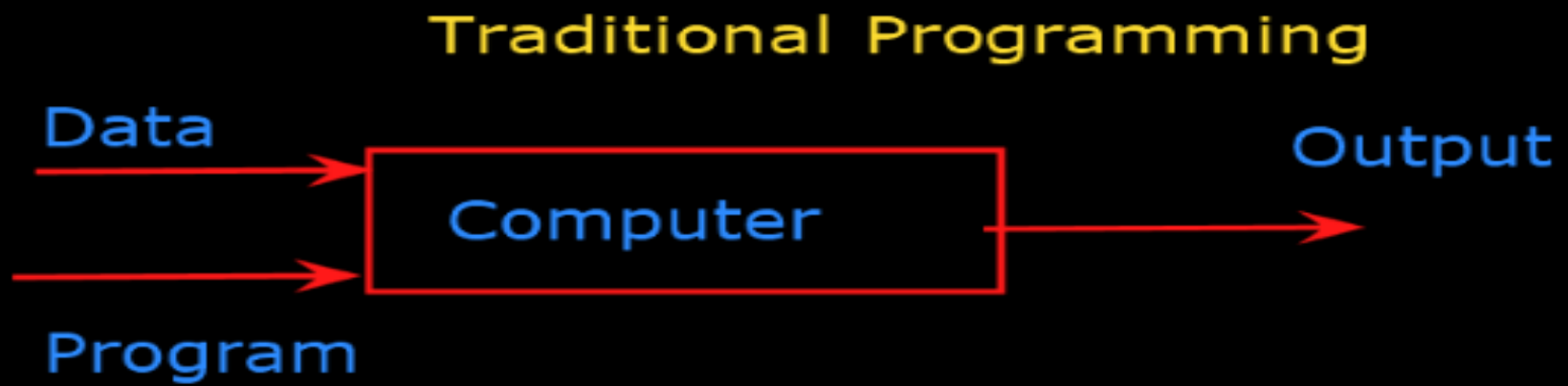
- [Tom Mitchell] algorithms that
  - improve their performance (P)
  - at some task (T)
  - with experience (E)

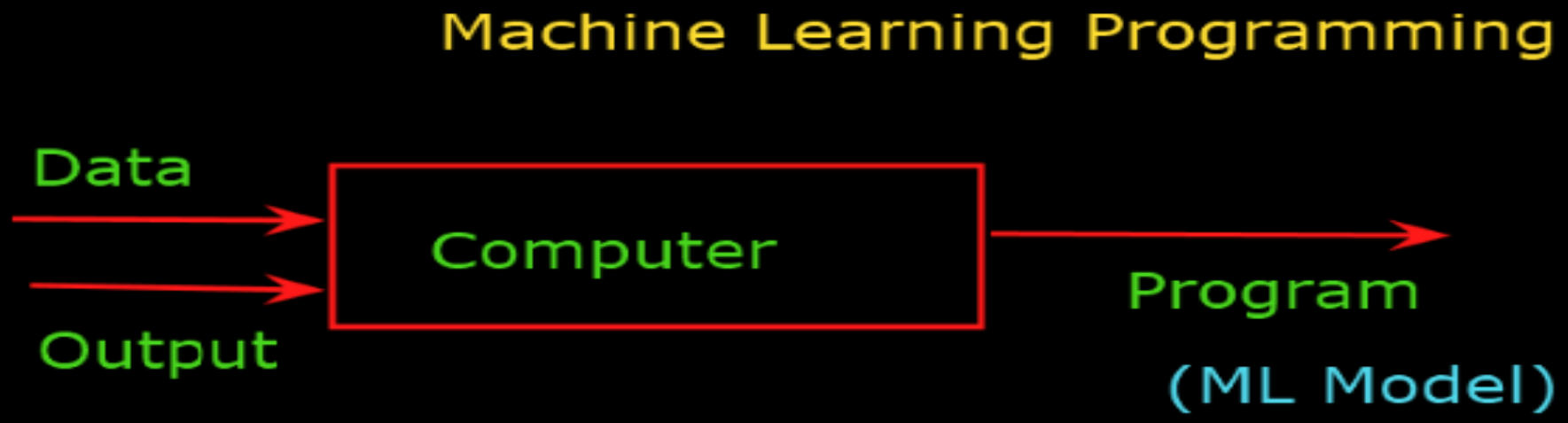
Definition

A computer program which learns from experience is called a *machine learning program* or simply a *learning program*. Such a program is sometimes also referred to as a *learner*.

What is Machine Learning?

- If you are a Scientist


Traditional Programming
Data
Program
Computer
Output


Machine Learning Programming
Data
Output
Computer
Program
(ML Model)

What is Machine Learning?

- [Arthur Samuel, 1959]
  - Field of study that gives computers
  - the ability to learn without being explicitly programmed

- [Kevin Murphy] algorithms that
  - automatically detect patterns in data
  - use the uncovered patterns to predict future data or other outcomes of interest

- [Tom Mitchell] algorithms that
  - improve their performance (P)
  - at some task (T)
  - with experience (E)

Learning: the acquisition of skills

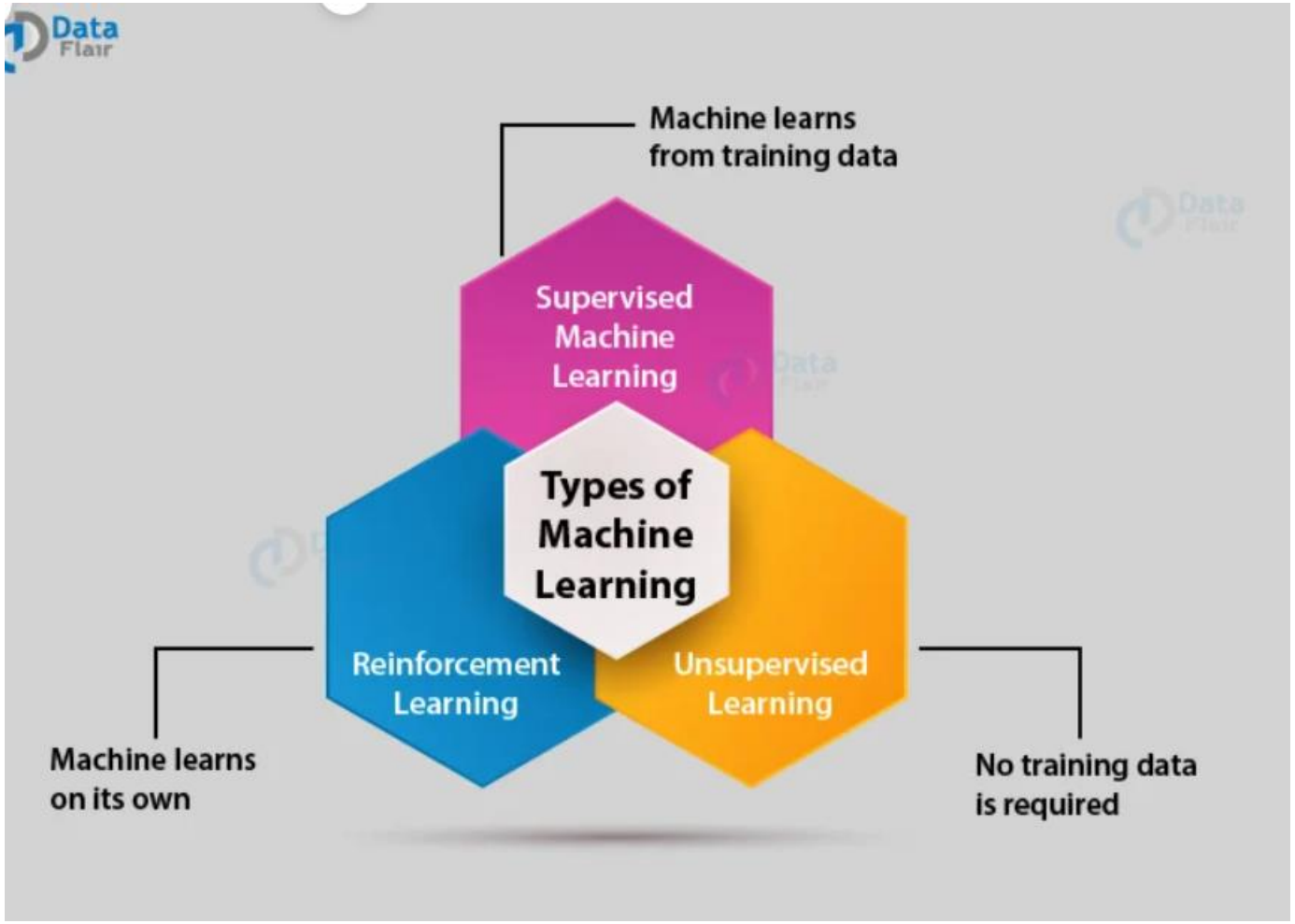
What is Machine Learning Model?

- **Definition:**
  - Machine Learning is a concept which allows the machine
  - **to learn from examples and experience**,
  - and that **too without being explicitly programmed**.

- Machine Learning algorithms are an evolution of **normal algorithms**.
- They make your **programs "smarter"**, by **allowing them to automatically learn** from the data you provide.
- The algorithm is mainly divided into:
  - **Training Phase**
  - **Testing phase**

Types of Machine Learning

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

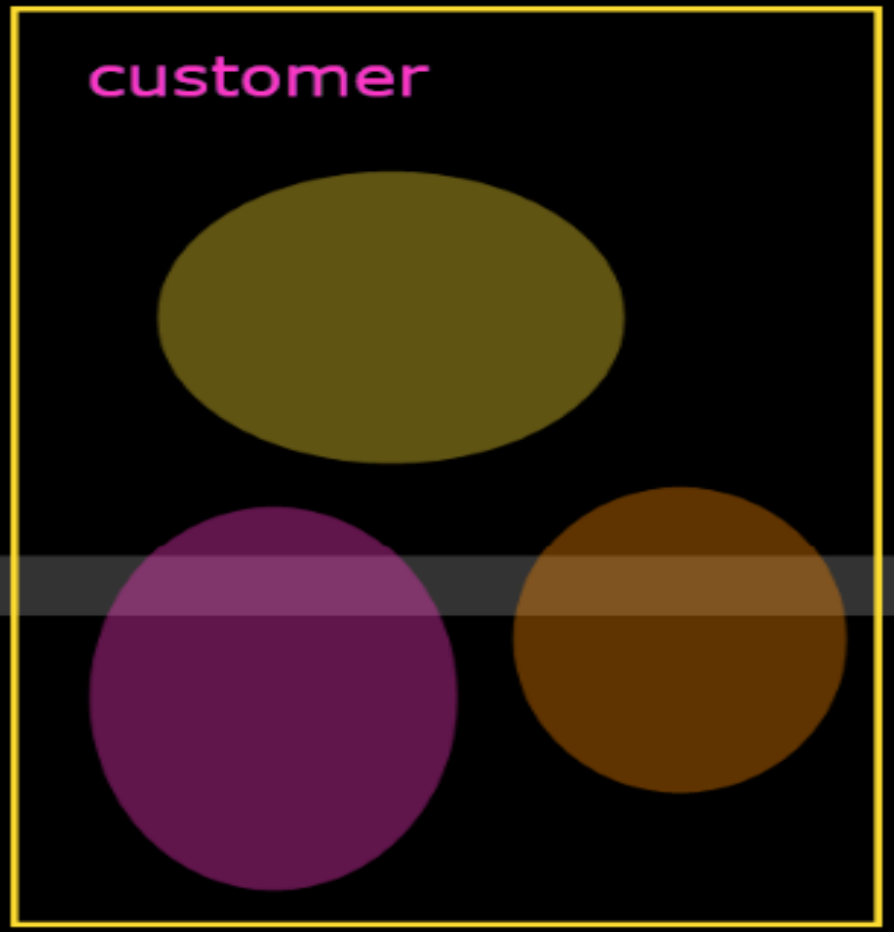


Who can see what you share here? Recording On

Types of learning:

Supervised:

-Labelled
-guidance will be provided
-explicitly learning model
-predicts the future outcomes
-train the algorithm on a labelled dataset
and then we will set it to prediction
-Input/ Output
-classification, prediction



Unsupervised:

-Not labelled
-No guidance (self learning)
-identify the pattens, trends.
-can not predict
-train an algorithms to find patterns
(similarities/ abnormalities)
-Input
-Cluster,anomaly detection

Machine Learning End Product

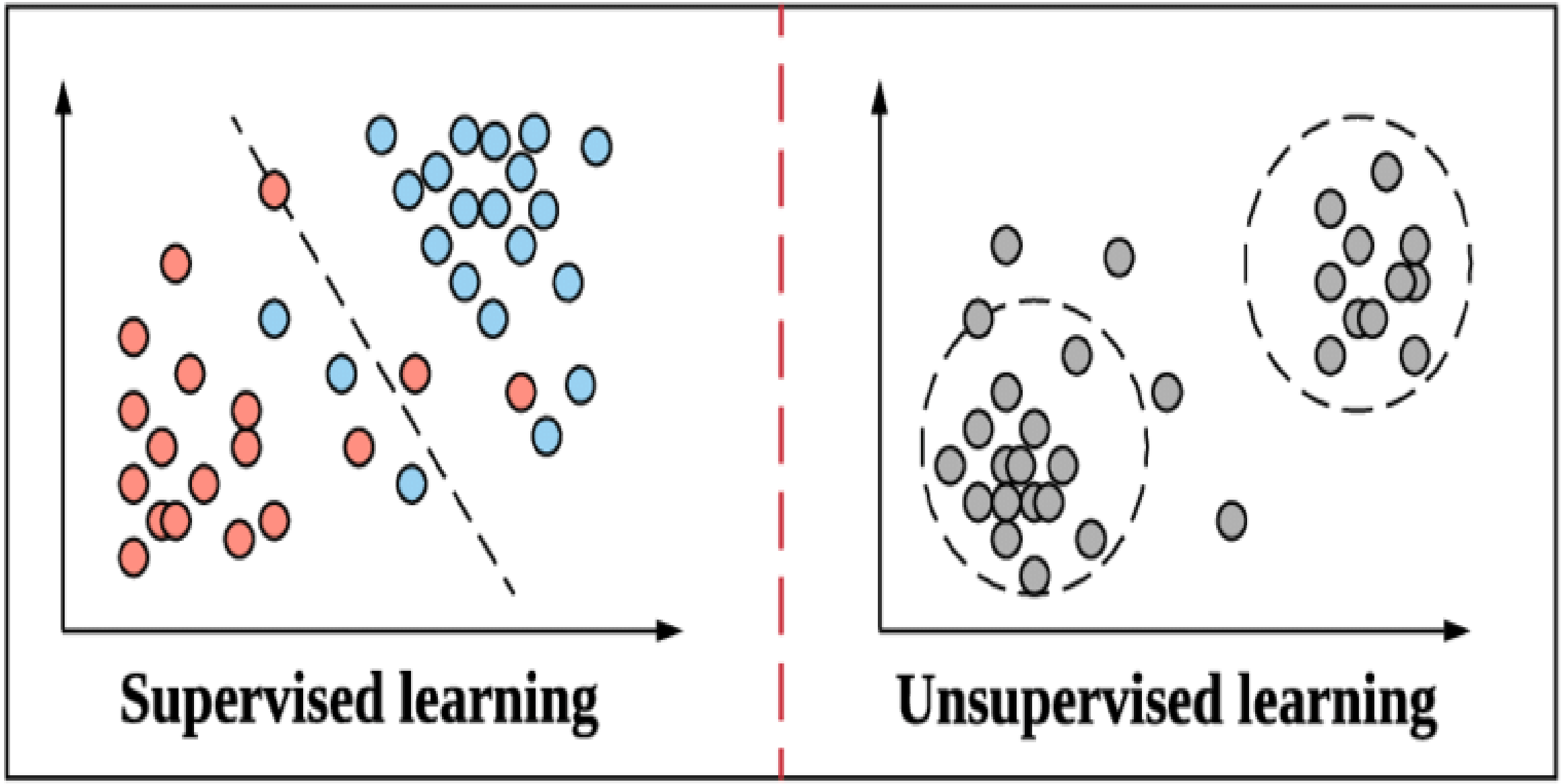
Ordinary System

With AI

Machine Learning

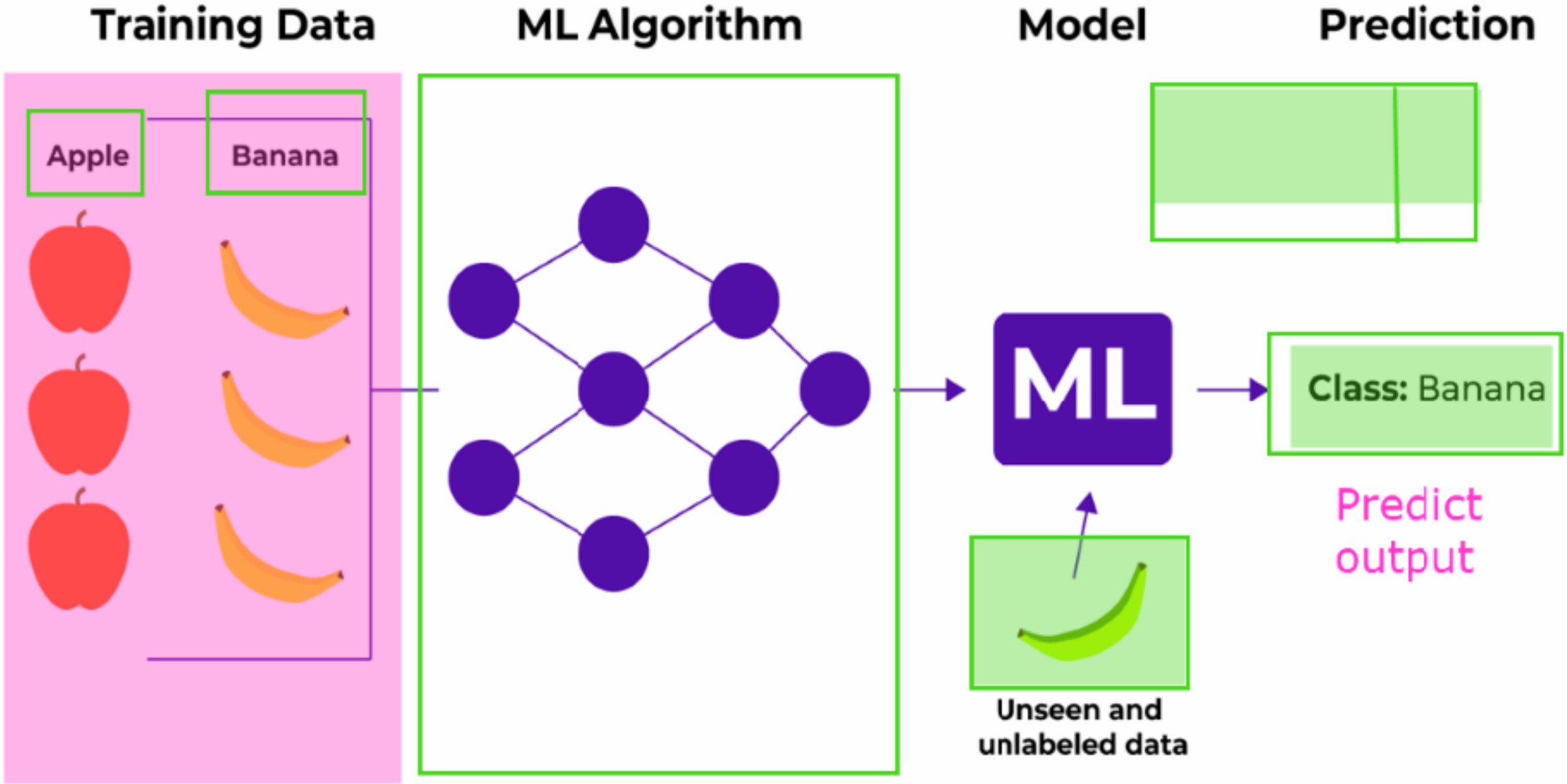
Learns

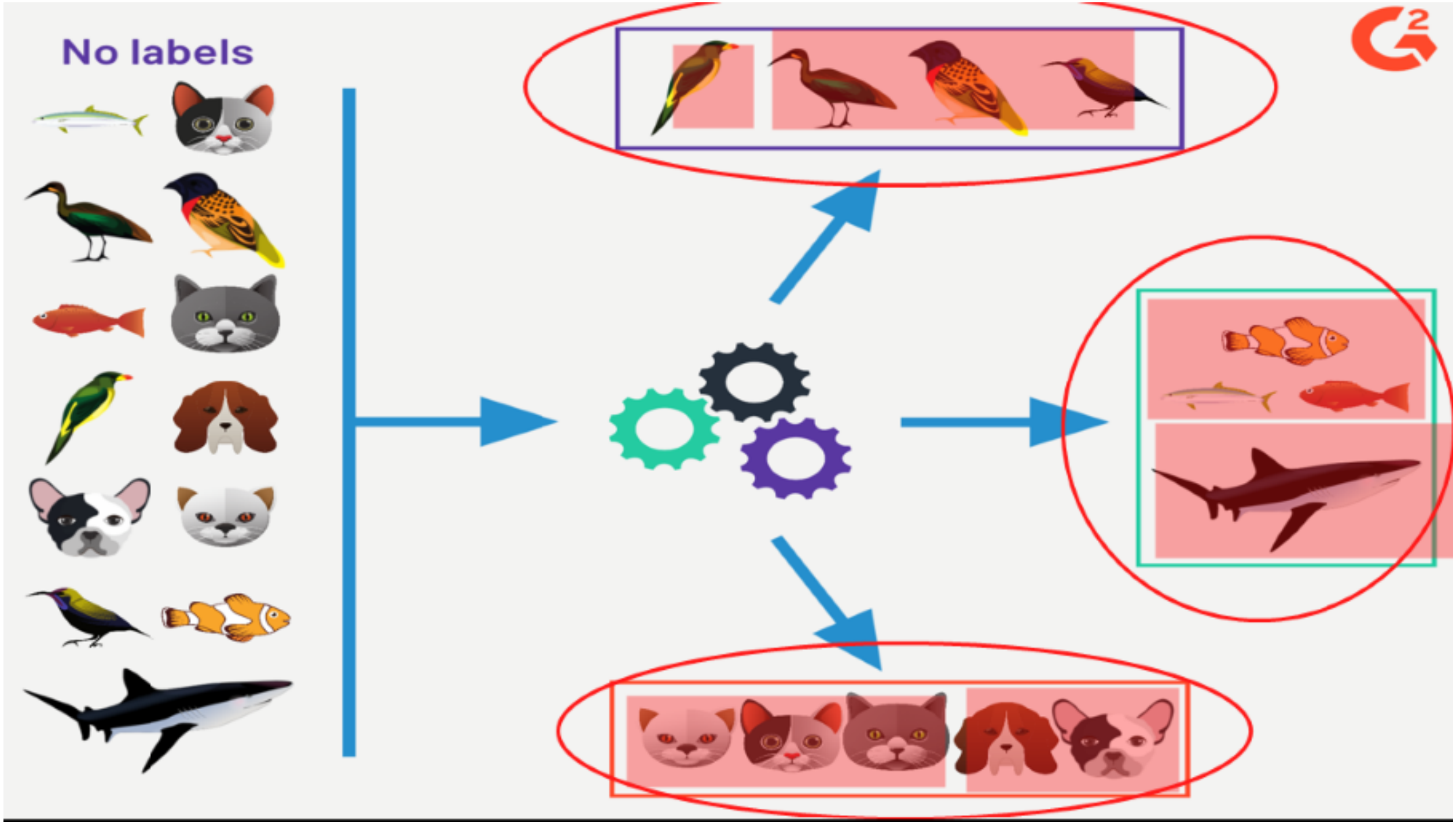
Predicts


Supervised learning
Unsupervised learning

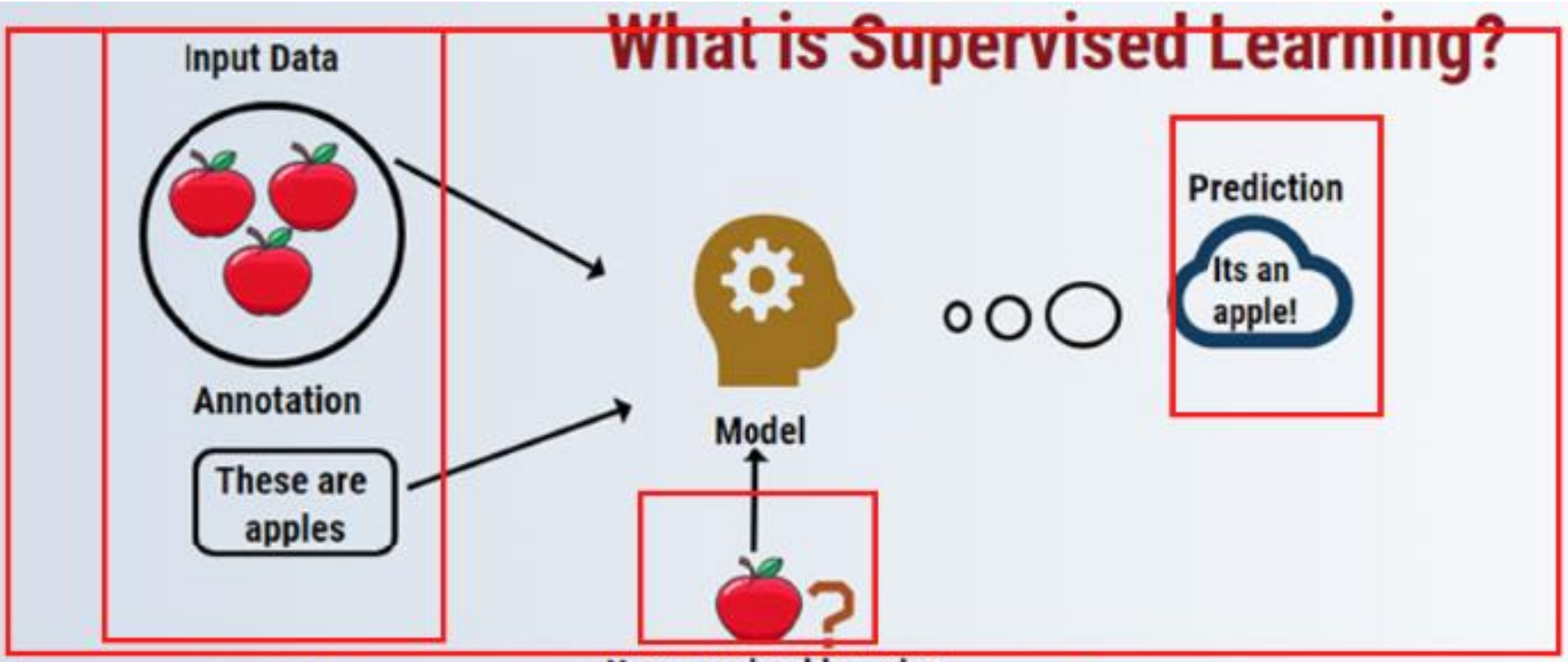
Supervised learning

- 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**.

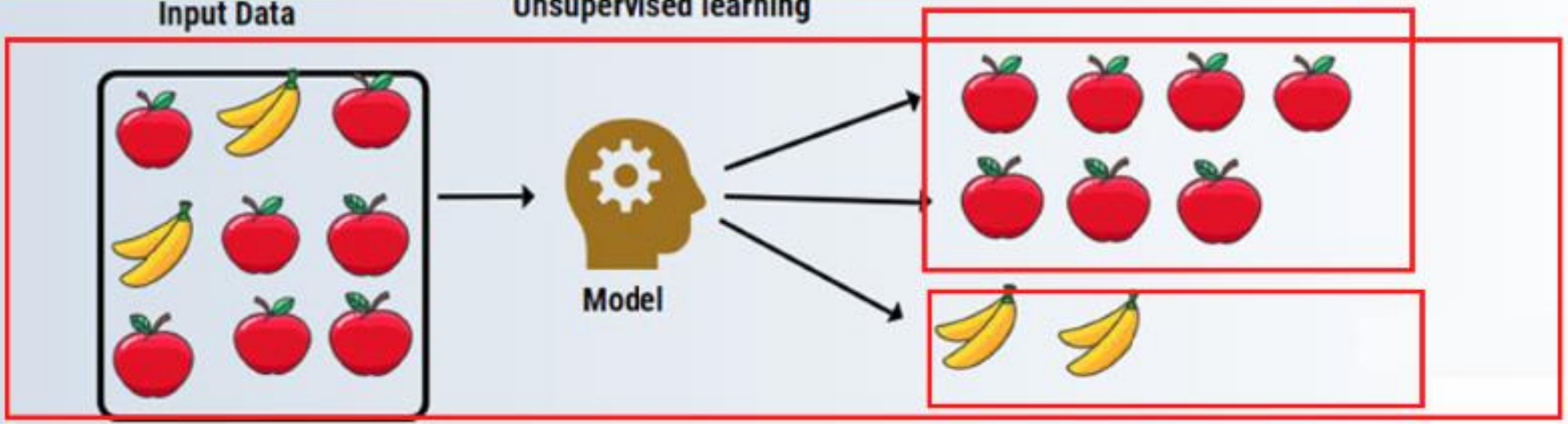

Training Data
ML Algorithm
Model
Prediction
Apple
Banana
ML
Class: Banana
Predict output
Unseen and unlabeled data


No labels

What is Supervised Learning?



Unsupervised learning



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

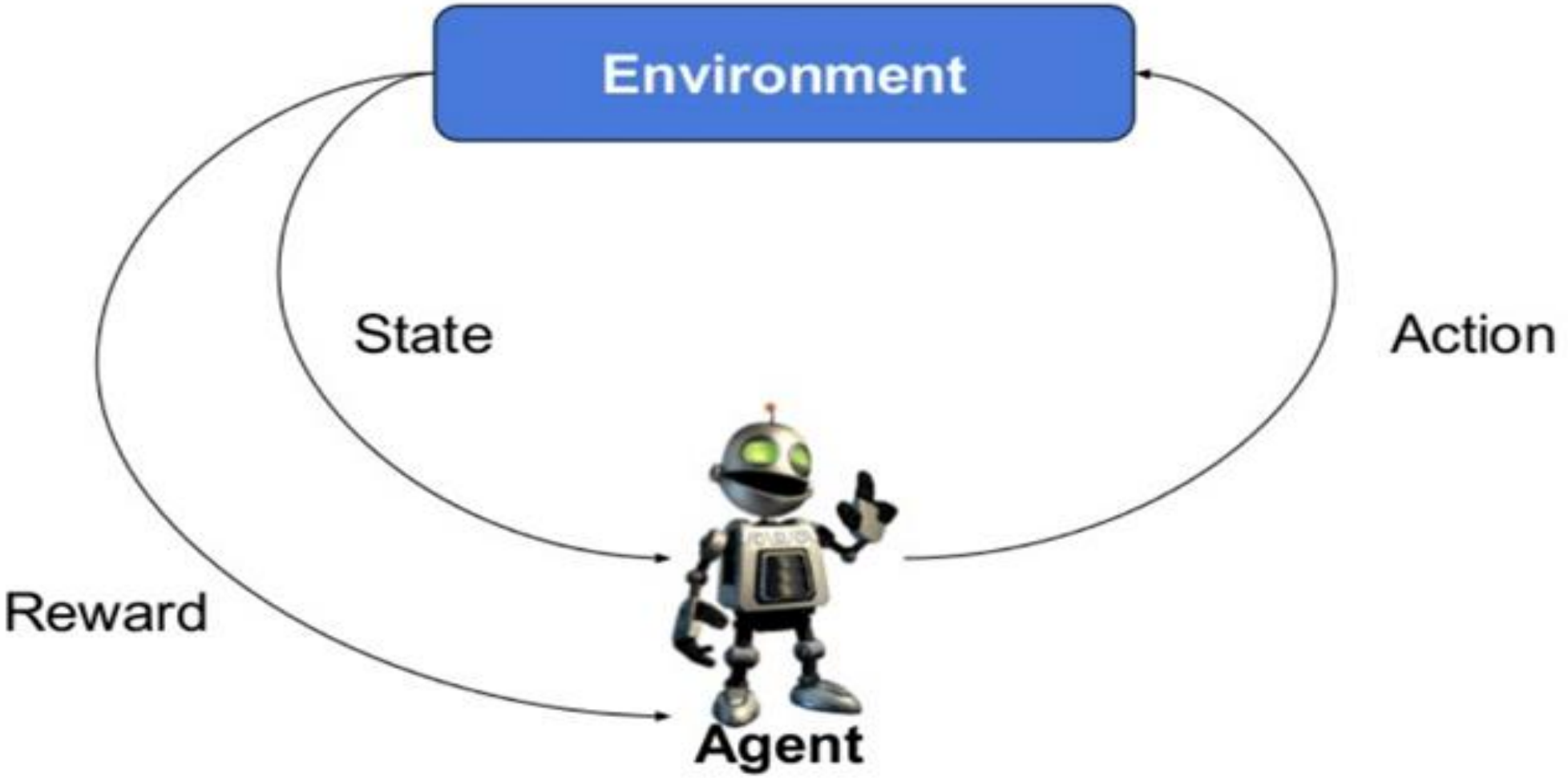

No label

Reinforcement Learning

- Reinforcement learning is one among three basic machine learning paradigms, alongside supervised learning and unsupervised learning.
- 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.

Reinforcement Learning

Typical RL scenario




environment
agent
actions
rewards
observations

Yes → **Reward +ve**

(Strong Network)

No → **Reward -ve**

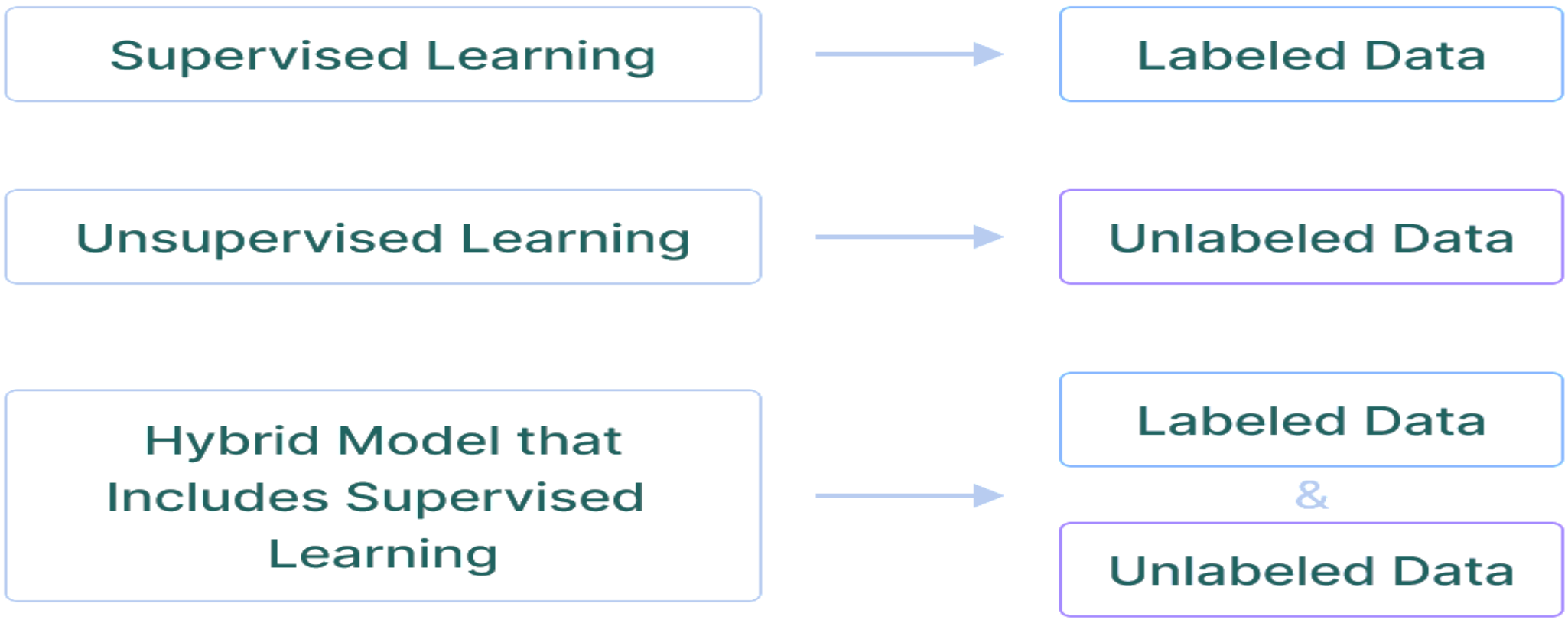
(Poor Network)

Environment

(Network Testing Zone)

Keep searching until you get strong network

Data in Supervised vs. Unsupervised Learning




Data Mining
Machine Learning
Statistics
Artificial Intelligence
Modeling
Future Predictions

Artificial Intelligence

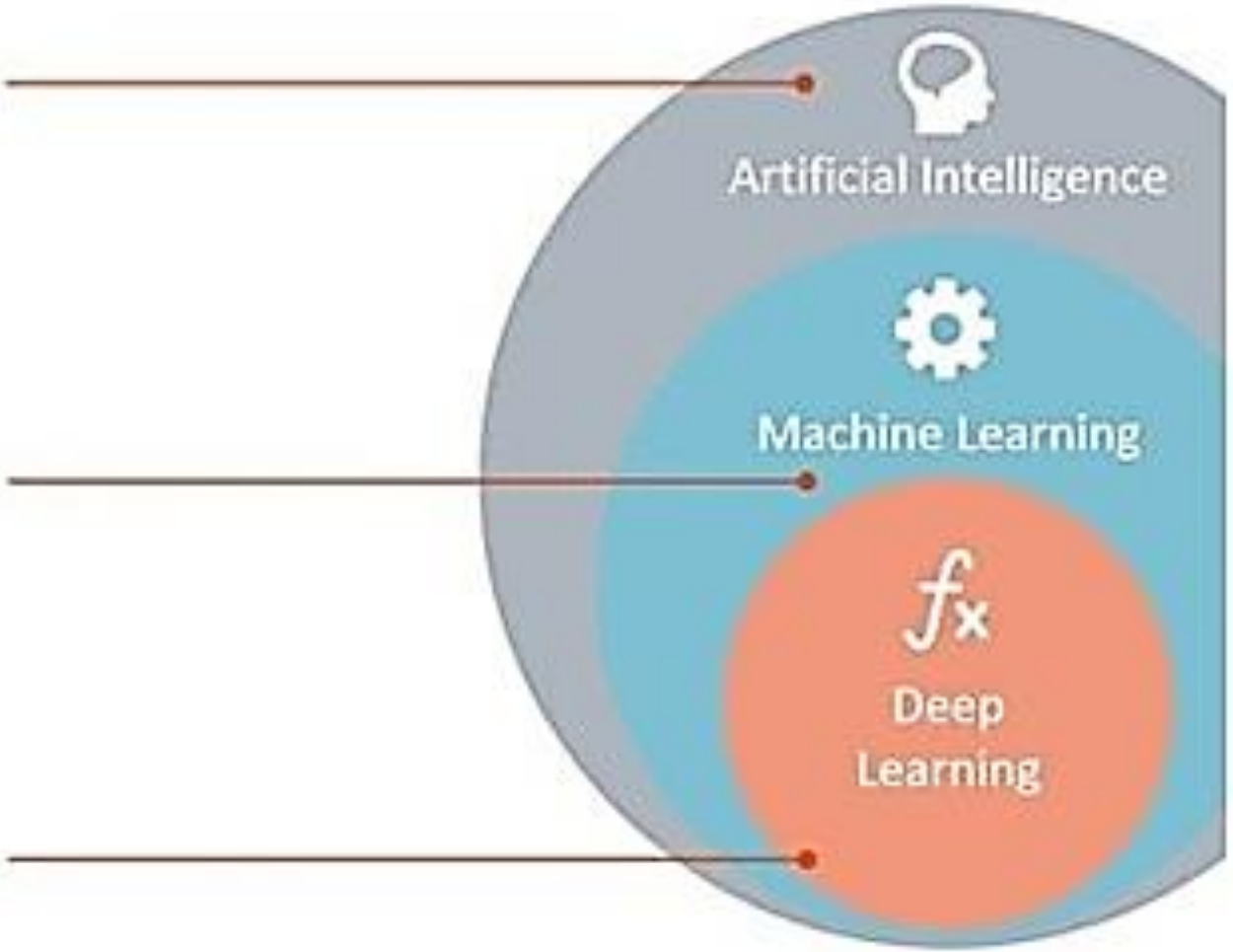
Any technique which enables computers to mimic human behavior.

Machine Learning

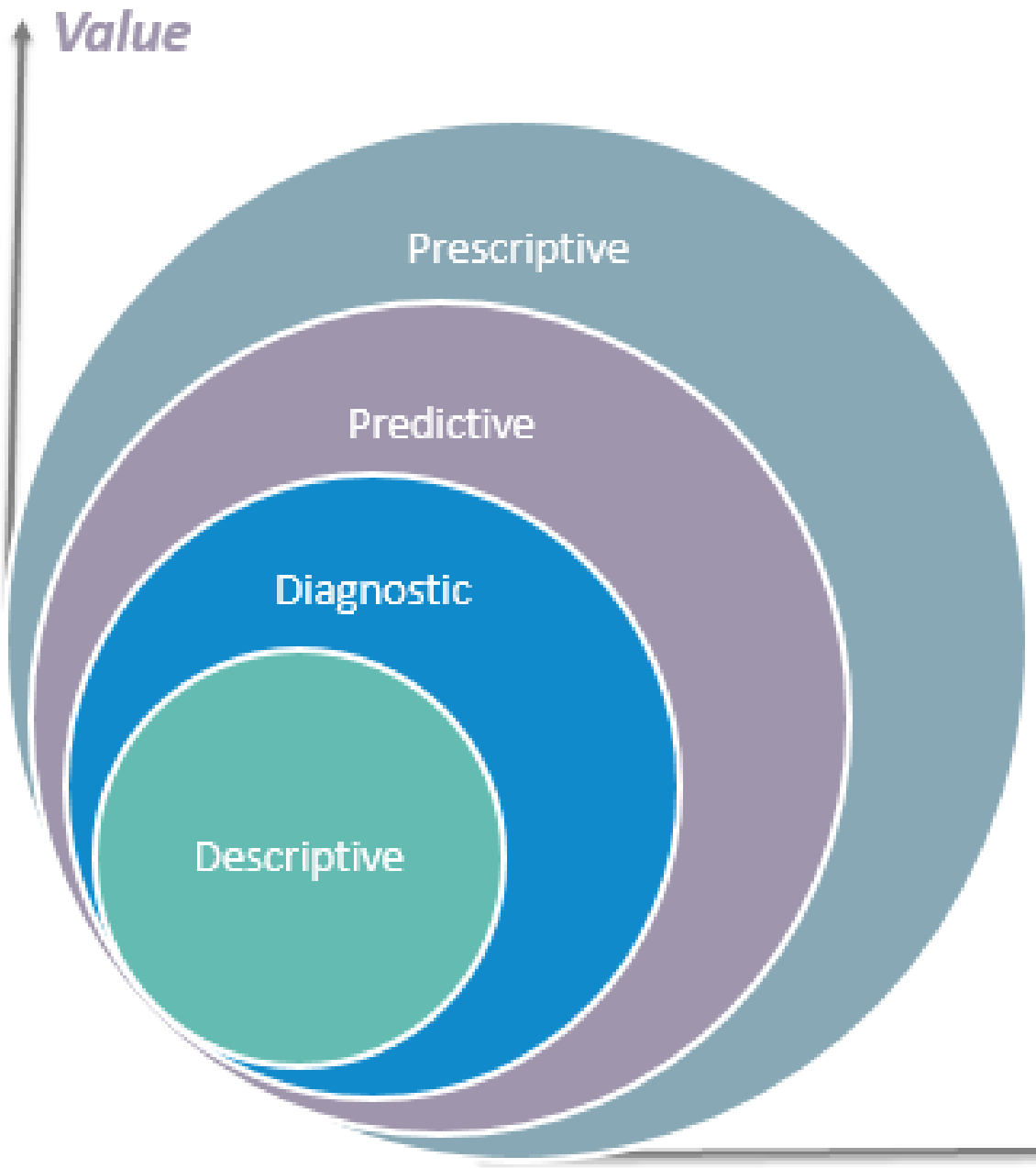
Subset of AI techniques which use statistical methods to enable machines to improve with experiences.

Deep Learning

Subset of ML which make the computation of multi-layer neural networks feasible.



4 types of Data Analytics



What is the data telling you?

Descriptive: *What's happening in my business?*

- Comprehensive, accurate and live data
- Effective visualisation

Diagnostic: *Why is it happening?*

- Ability to drill down to the root-cause
- Ability to isolate all confounding information

Predictive: *What's likely to happen?*

- Business strategies have remained fairly consistent over time
- Historical patterns being used to predict specific outcomes using algorithms
- Decisions are automated using algorithms and technology

Prescriptive: *What do I need to do?*

- Recommended actions and strategies based on champion / challenger testing strategy outcomes
- Applying advanced analytical techniques to make specific recommendations

Challenges and Limitations of Machine Learning

- The primary challenge of machine learning is the lack of data or the **diversity in the dataset.**
- A machine cannot learn if there **is no data available**.
- Besides, a dataset with a lack of diversity gives the machine a hard time.
- A machine **needs to have heterogeneity** to learn meaningful insight.
- It is rare that an algorithm can extract information when there are no or few variations.
- It is recommended to have at least 20 observations per group to help the machine learn.
- This constraint leads to poor evaluation and prediction.

Interview Questions

- 1. What is meant by “learning” in the context of machine learning?
- 2. What is Machine Learning? Give an example.
- 3. List out the types of machine learning.
- 4. What are the differences between supervised and unsupervised learning?
- 5. What is meant by supervised classification?
- 6. Explain supervised learning with an example.
- 7. What do you mean by reinforcement learning?