



Data
Schools

Module 1:

Introduction to Machine Learning

Overview

Faculty Introduction

Ekpe Okorafor, PhD



AFFILIATIONS:

- Accenture - Applied Artificial Intelligence
- Nigerian British University
- African University of Science & Technology

RESEARCH INTERESTS:

- Generative AI
- Machine Learning
- Big Data & Advanced Analytics
- Knowledge Graph
- Information Assurance and Cybersecurity

FUN & INTERESTS

- Drawing & Painting
- Hiking
- Soccer
- DJ & Dancing

Agenda Topics

1. **Overview: What is Machine learning**
 - *Building Classification Model – Lab Exercises*
2. **Recommender Systems**
 - *Building a Recommender Engine – Lab Exercises*
3. **From ML to Deep Learning**
 - The Rise of Gen AI – Discussion Topic

Learning Objectives

Upon successful completion of this topic, you will be able to:

- Define machine learning
- Describe the categories of machine learning
- Decide when to leverage Machine learning
- Build a simple classifier model
- Discuss approaches to ML application development
- Differentiate between the ML approaches and motivations
- Build a simple recommender engine
- Good insight to Deep Learning & Gen AI

Introduction to Machine Learning

Overview

Sub - Topics

- Overview: What is Machine learning
- Categories of Machine Learning
- Machine Learning Application Development Approach
- Building Classification Model

Overview: What is Machine Learning?

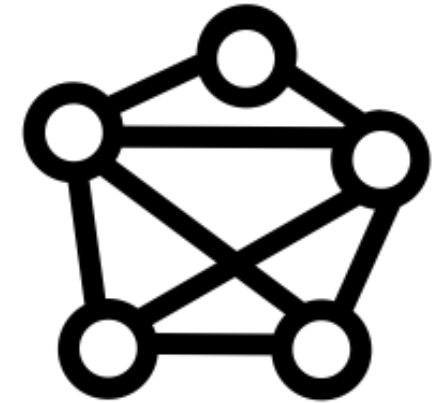
What is Machine Learning?

```
100100011101000000101000110111010110
10010011110111000000111100110100100
100001101101111101010011100001101001
111111010000110111001010111100001011
1100111110111111100100001110110110
010000110100110110000110000100010000
010101110011001111011001110100010111
001000010101100101000001000010011110
0111010011111100101110101010111100
100010000101100010101101010111000101
010010000100101011110011100001010000
010110000010011101010010101110110001
011011111010111100010100010100010000
011010011011011010001000101111001101
000101000001100110001100100010010110
10010101010001001110010101010111101
```

Data

$$f_x$$

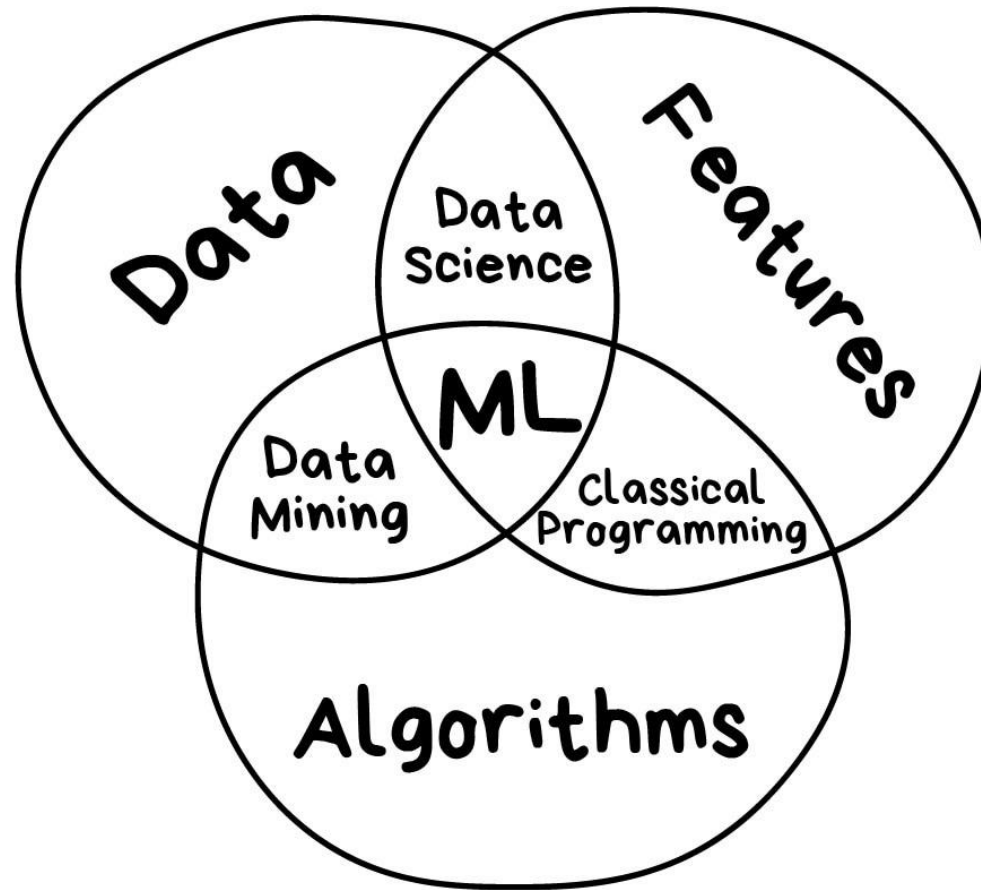
Features



Algorithms

“Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed. – Arthur L. Samuel, AI Pioneer, 1959

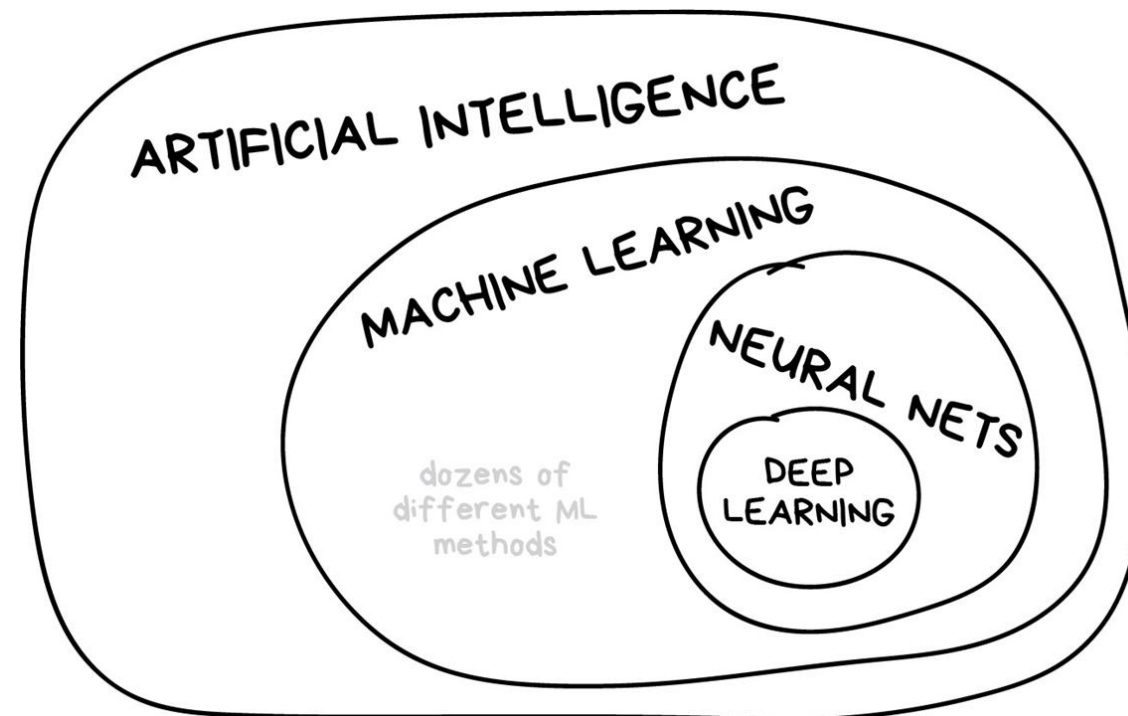
What is Machine Learning?



“Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed. – Arthur L. Samuel, AI Pioneer, 1959

What is Machine Learning?

- Artificial Intelligence (AI) is a branch of Computer Science that uses algorithms and techniques to mimic human intelligence
- Machine Learning (ML) is one of several AI techniques for sophisticated cognitive tasks



Machine Learning

Machine Learning is a particularly interesting technique because it represents a paradigm shift within AI

Traditional AI techniques



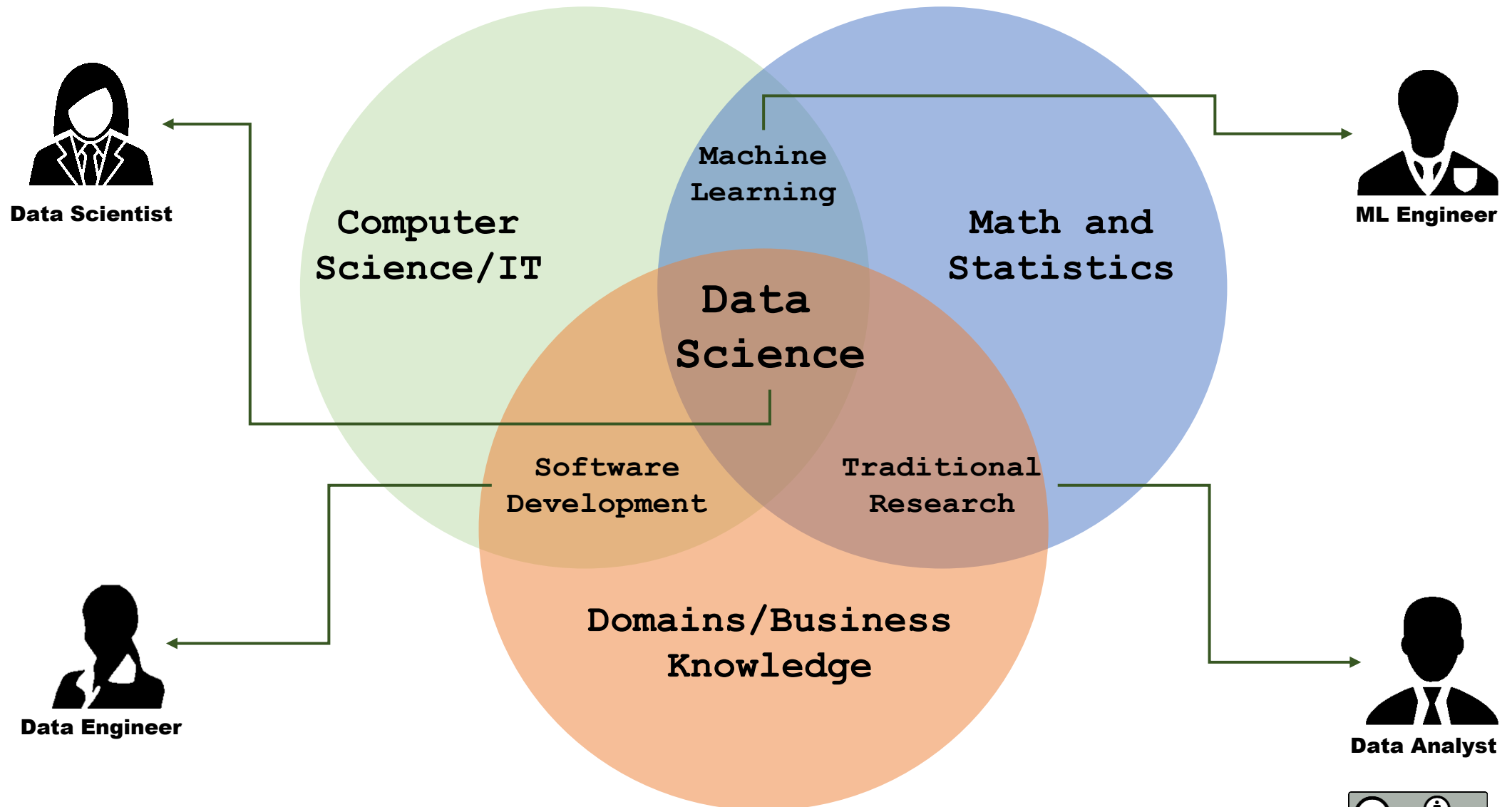
- **Static** – hard-coded set of steps and scenarios
- **Rule Based** – expert knowledge
- **No generalization** – handling special cases is difficult

Machine Learning



- **Dynamic** – evolves with data, finds new patterns
- **Data driven** – discovers knowledge
- **Generalization** – adapts to new situations and special cases

The World of Data



Categories of Machine Learning

The Main Categories of Machine Learning

1

CLASSICAL ML

Simple data with
clear features

2

NEURAL NETWORKS & DEEP LEARNING

Complicated data
with unclear
features

3

REINFORCEMENT LEARNING

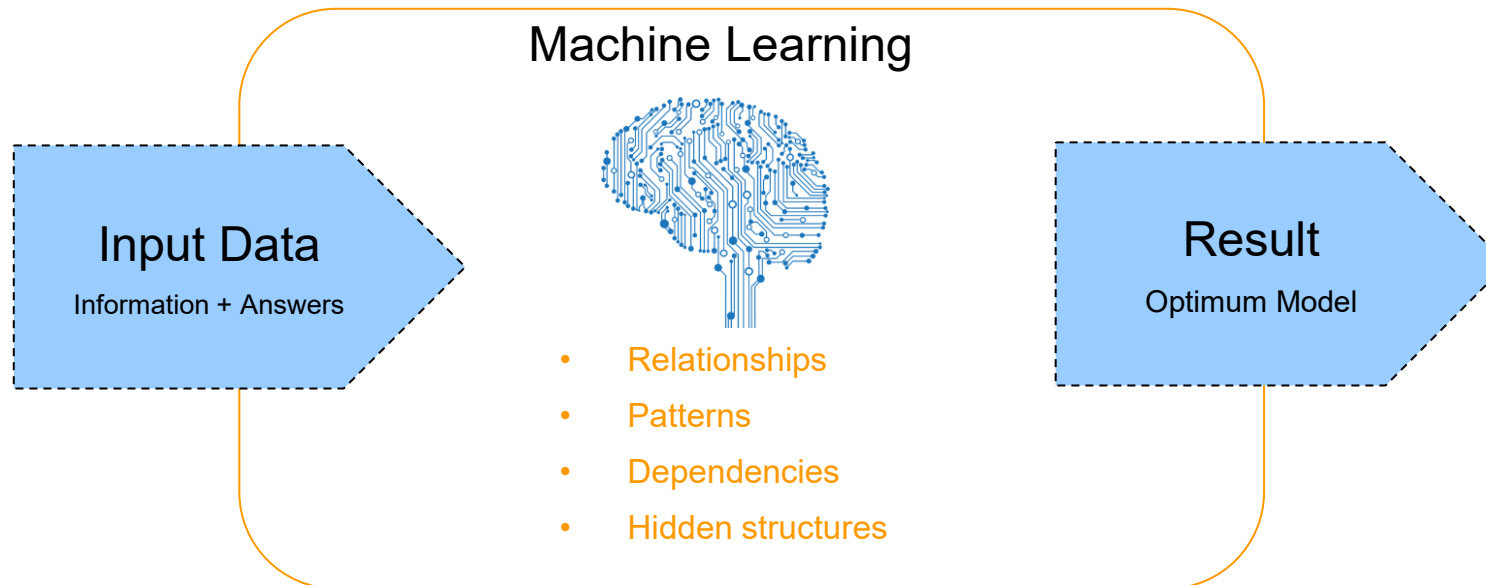
No data, but we
have an
environment to
interact with

Our focus today

Classical Machine Learning

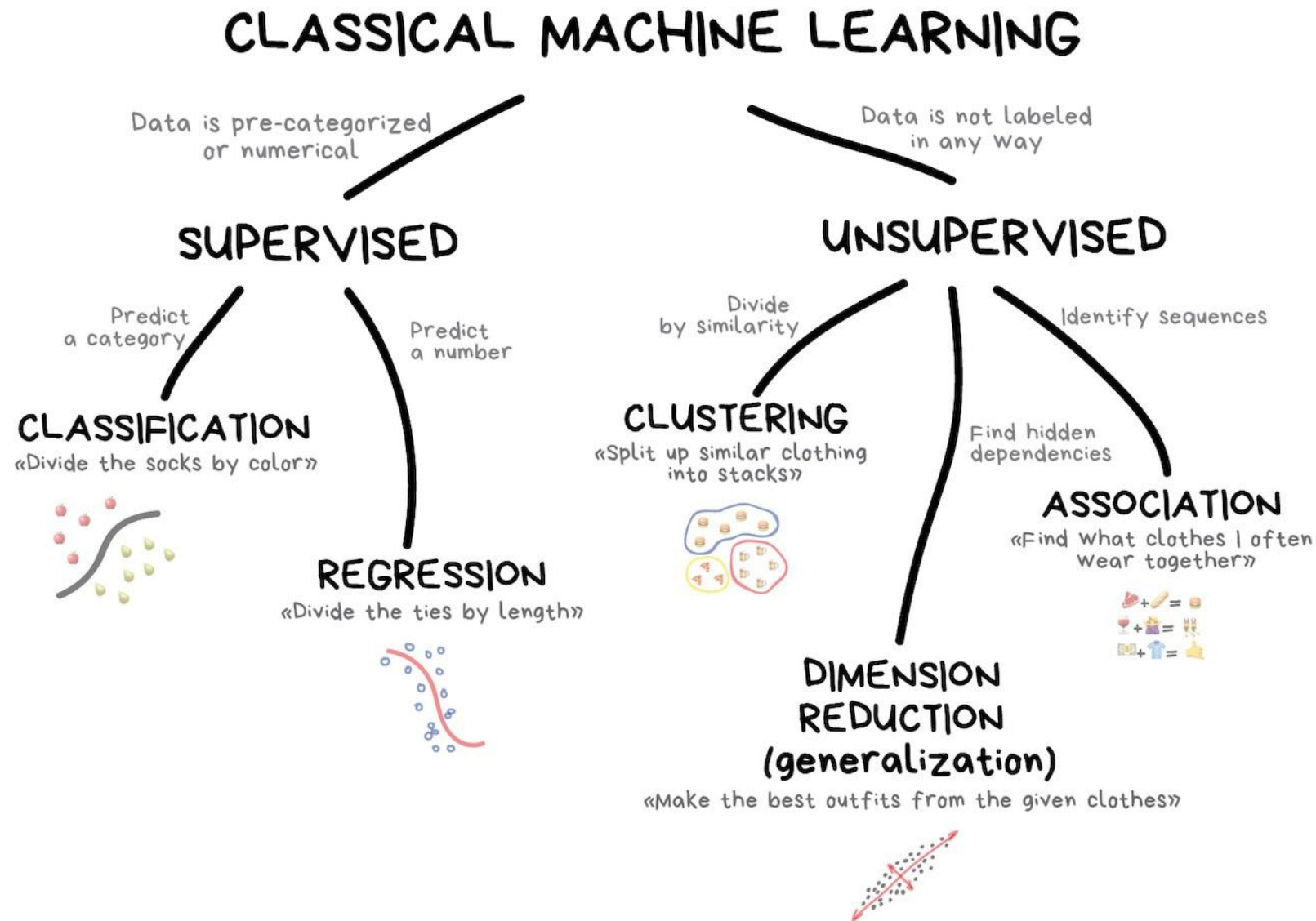
Supervised and Unsupervised Learning

1. **Supervised learning** - we already know the answers we want (found in past or completed data).



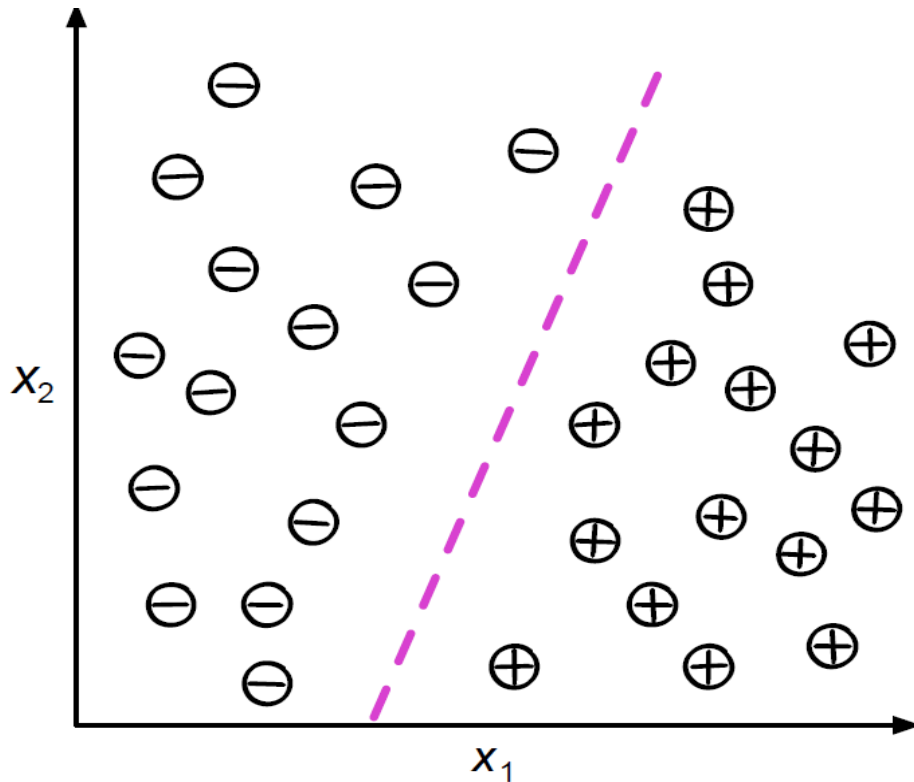
2. **Unsupervised learning** - we want to find unknown structures or trends.

Breaking It Down



Supervised Learning: Classification

Splits objects based at one of the attributes known beforehand.

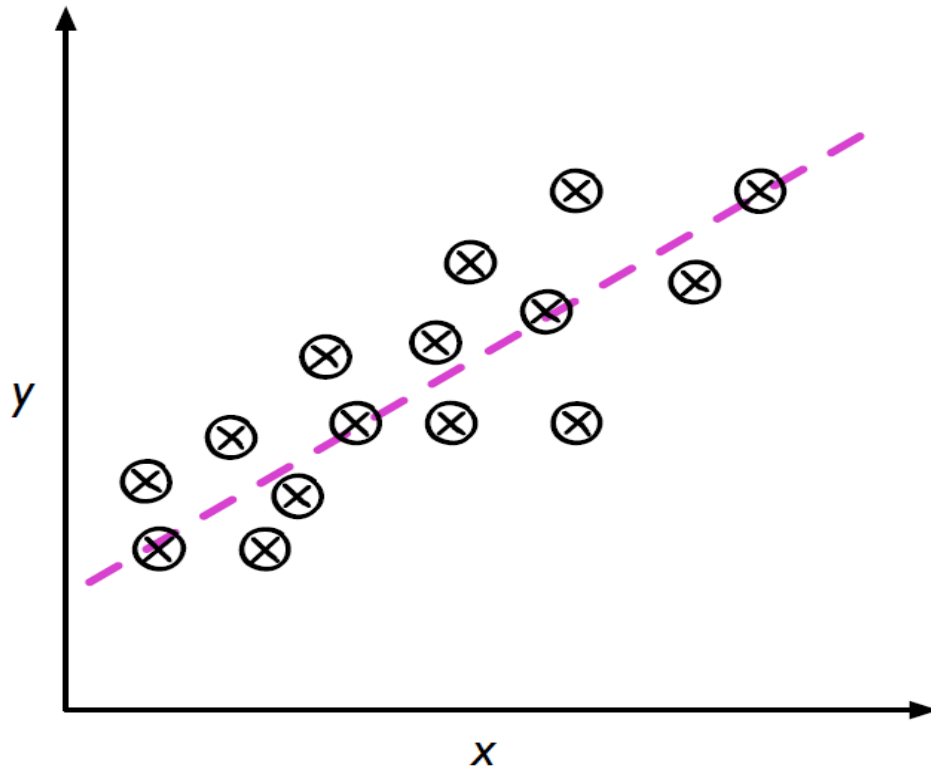


Today used for:

- Spam filtering
- Language detection
- A search of similar documents
- Sentiment analysis
- Recognition of handwritten characters and numbers
- Fraud detection

Supervised Learning: Regression

Regression is basically classification where we forecast a number instead of category. (Linear and Polynomial)

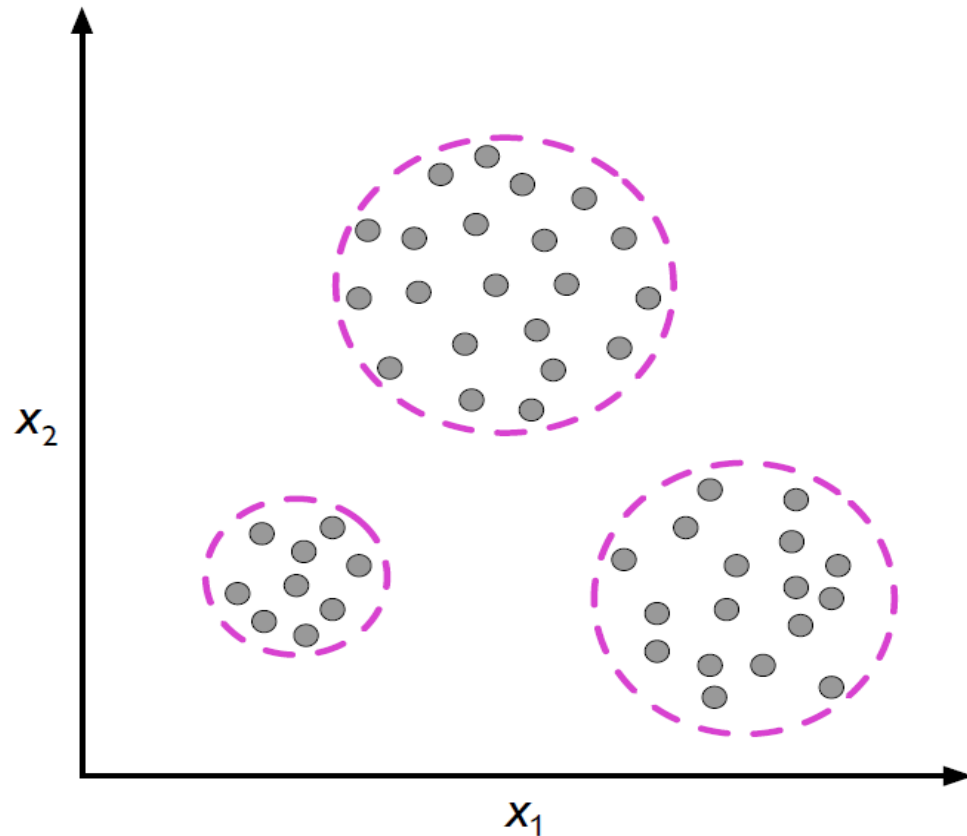


Today used for:

- Stock price forecasts
- Demand and sales volume analysis
- Medical diagnosis
- Any number-time correlations

Unsupervised Learning: Clustering

Divides objects based on unknown features. Machine chooses the best way.

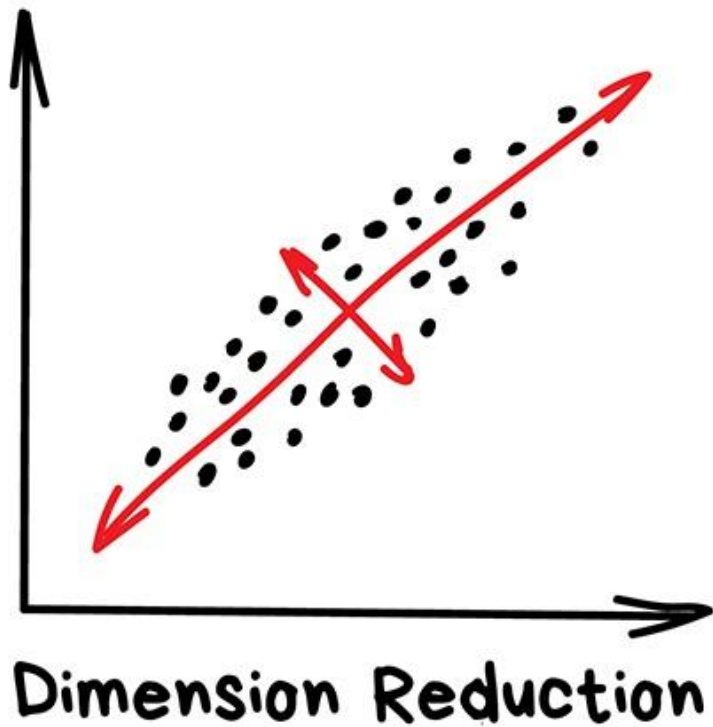


Today used for:

- For market segmentation (types of customers, loyalty)
- To merge close points on a map
- For image compression
- To analyze and label new data
- To detect abnormal behavior

Unsupervised Learning: Dimensionality Reduction

Divides objects based on unknown features. Machine chooses the best way.

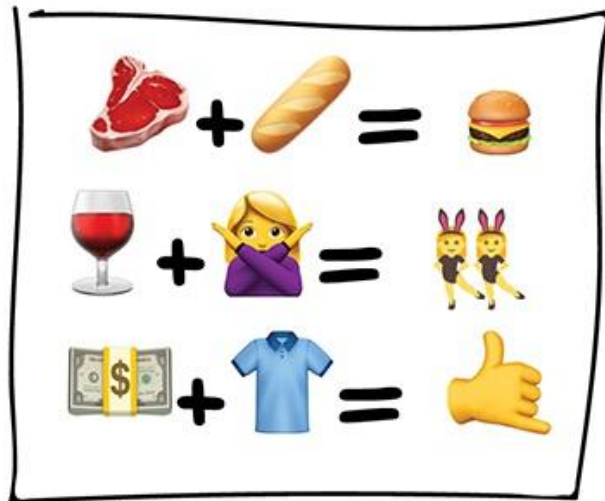


Today used for:

- Recommender systems (★)
- Beautiful visualizations
- Topic modeling and similar document search
- Fake image analysis
- Risk management

Unsupervised Learning: Association Rule Learning

Look for patterns in the orders' stream.



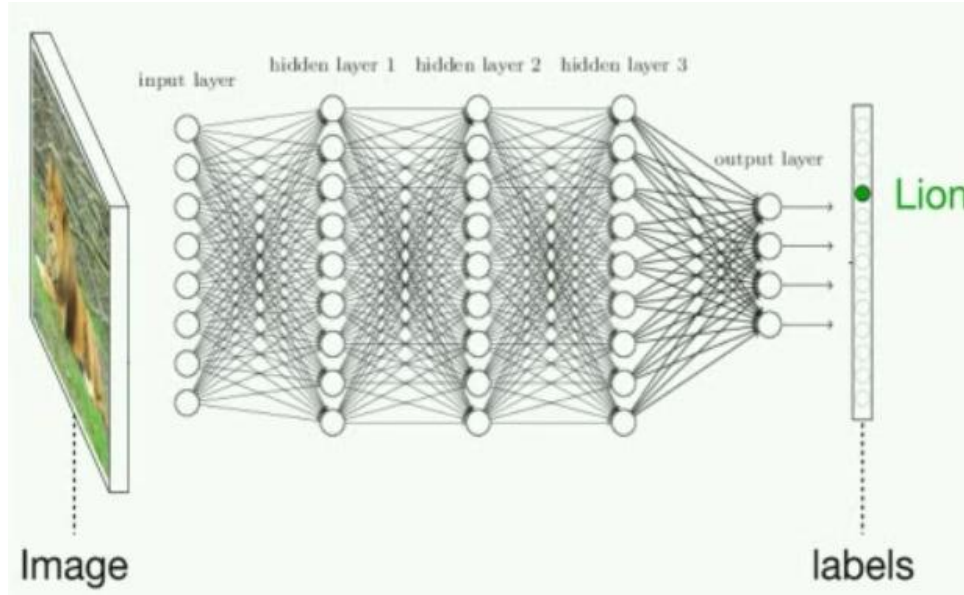
Association Rule Learning

Today used for:

- To forecast sales and discounts
- To analyze goods bought together
- To place the products on the shelves
- To analyze web surfing patterns

Neural Networks and Deep Learning

Mimic the way the brain works – collection of neurons and connections



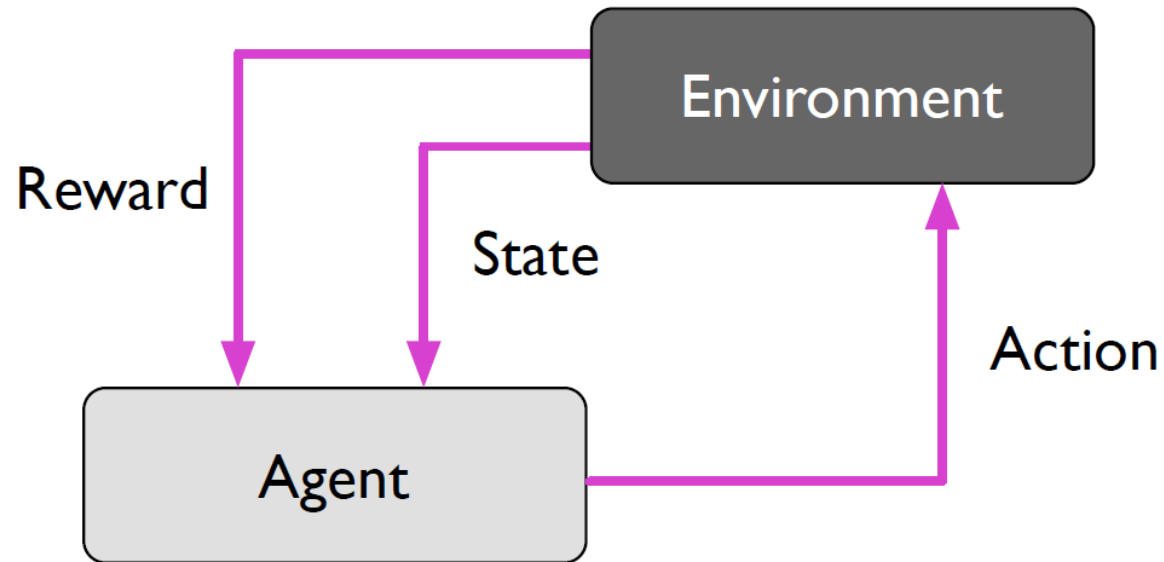
Neural Networks

Today used for:

- Replacement of all algorithms above
- Object identification on photos and videos
- Speech recognition and synthesis
- Image processing, style transfer
- Machine translation

Reinforcement Learning:

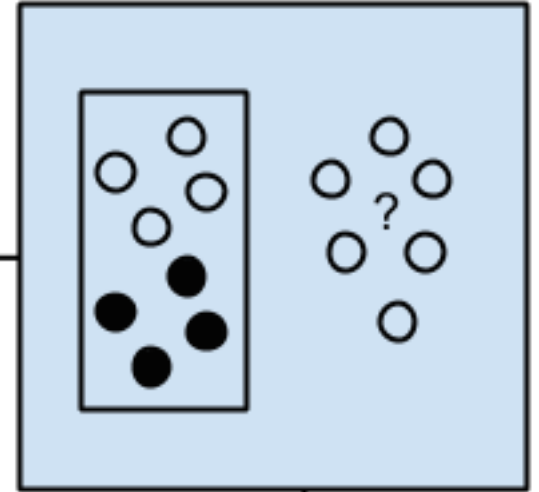
Throw a robot into a maze and let it find an exit.



Today used for:

- Self-driving cars
- Robot vacuums
- Games
- Automating trading
- Enterprise resource management

Supervised Learning – Simple Example



Math Quiz #1 - Teacher's Answer Key

$$1) \quad 2 \quad 4 \quad 5 \quad = \quad 3$$

$$2) \quad 5 \quad 2 \quad 8 \quad = \quad 2$$

$$3) \quad 2 \quad 2 \quad 1 \quad = \quad 3$$

$$4) \quad 4 \quad 2 \quad 2 \quad = \quad 6$$

$$5) \quad 6 \quad 2 \quad 2 \quad = \quad 10$$

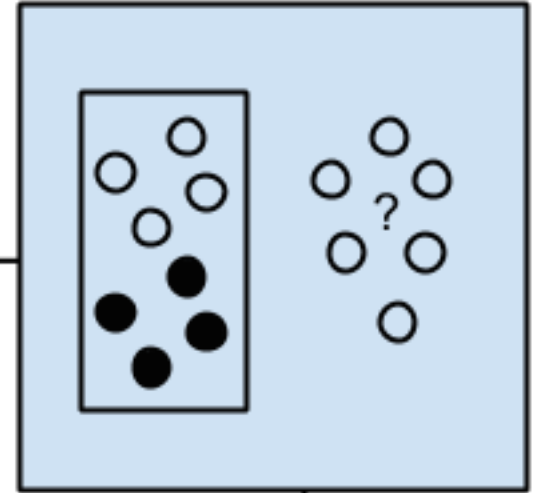
$$6) \quad 3 \quad 1 \quad 1 \quad = \quad 2$$

$$7) \quad 5 \quad 3 \quad 4 \quad = \quad 11$$

$$8) \quad 1 \quad 8 \quad 1 \quad = \quad 7$$

Supervised Learning
Algorithms

Supervised Learning – Simple Example



Math Quiz #1 - Teacher's Answer Key

1) $2 \times 4 - 5 = 3$

2) $5 \times 2 - 8 = 2$

3) $2 \times 2 - 1 = 3$

4) $4 \times 2 - 2 = 6$

5) $6 \times 2 - 2 = 10$

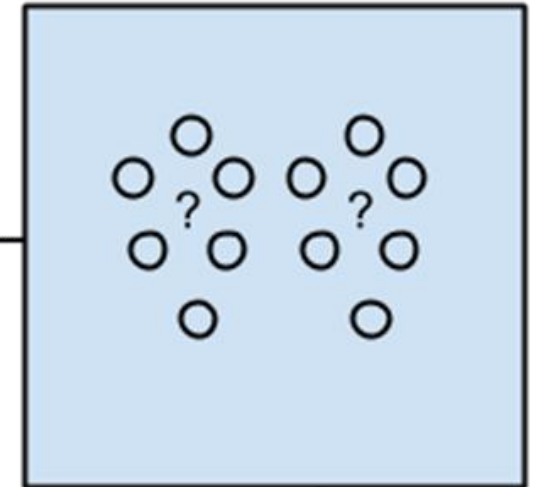
6) $3 \times 1 - 1 = 2$

7) $5 \times 3 - 4 = 11$

8) $1 \times 8 - 1 = 7$

Supervised Learning
Algorithms

Unsupervised Learning – Simple Example



Unsupervised Learning
Algorithms

Math Quiz #1 - Teacher's Answer Key

1) 2 4 5 =

2) 5 2 8 =

3) 2 2 1 =

4) 4 2 2 =

5) 6 2 2 =

6) 3 1 1 =

7) 5 3 4 =

8) 1 8 1 =

Supervised & Unsupervised Learning

Supervised Learning:

Predicting values. **Known** targets.

User inputs correct answers to learn from. Machine uses the information to guess new answers.

CLASSIFICATION:

Identify a unique class
(Discrete values, Boolean, Categories)

REGRESSION:

Estimate continuous values
(Real-valued output)

Unsupervised Learning:

Search for structure in data. **Unknown** targets.

User inputs data with undefined answers. Machine finds useful information hidden in data

CLUSTER ANALYSIS:

Group into sets

DIMENSION REDUCTION:

Find hidden dependencies

ASSOCIATION:

Identify Sequences

Supervised & Unsupervised Learning

Supervised Learning:

Classification

- Decision Trees
- K-Nearest Neighbors
- Support Vector Machine
- Logistic Regression
- Naïve Bayes
- Random Forests

Regression

- Linear Regression
- Polynomial Regression

Unsupervised Learning:

Cluster Analysis

- K-Means Clustering
- Hierarchical Clustering

Dimension Reduction

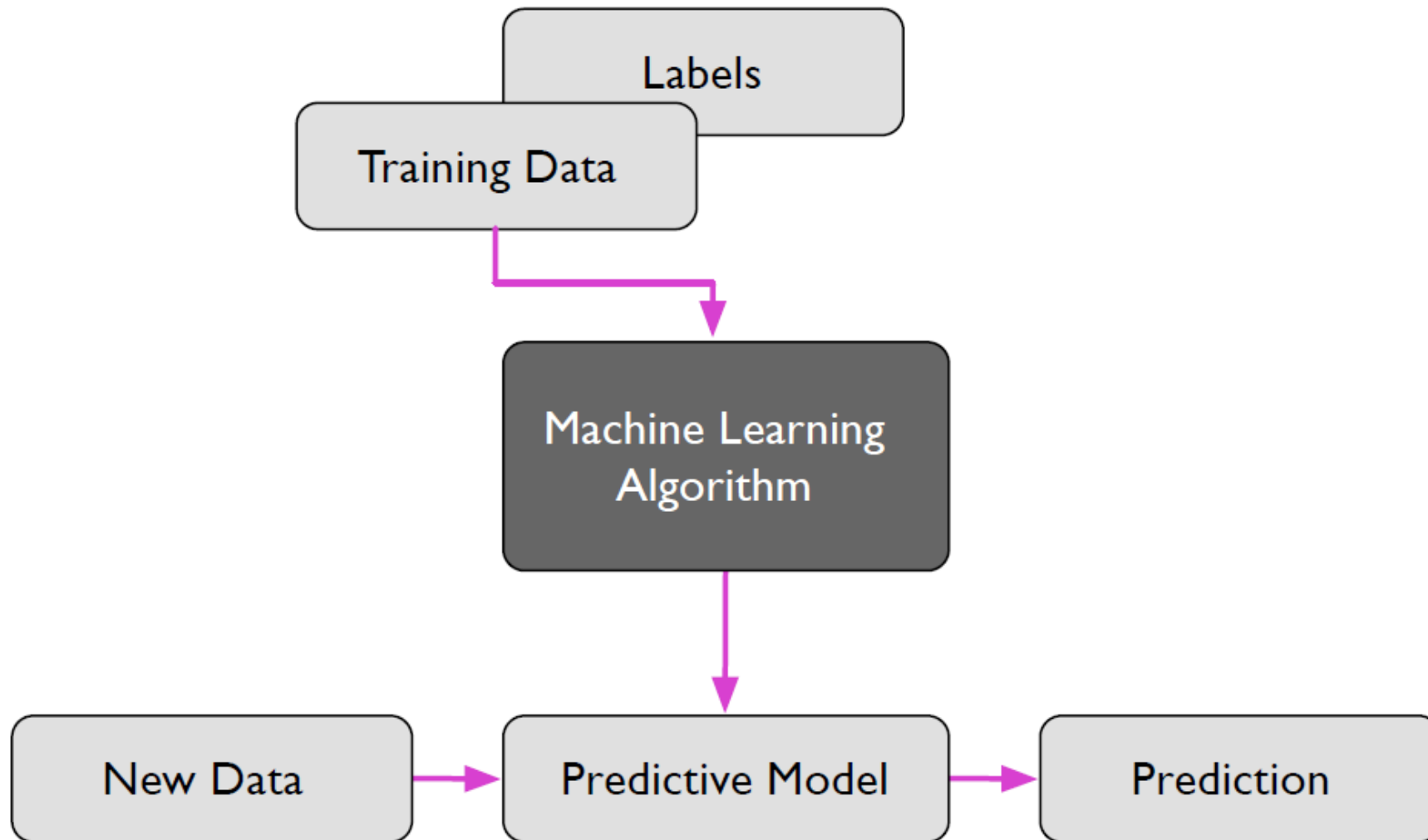
- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)

Association Rule

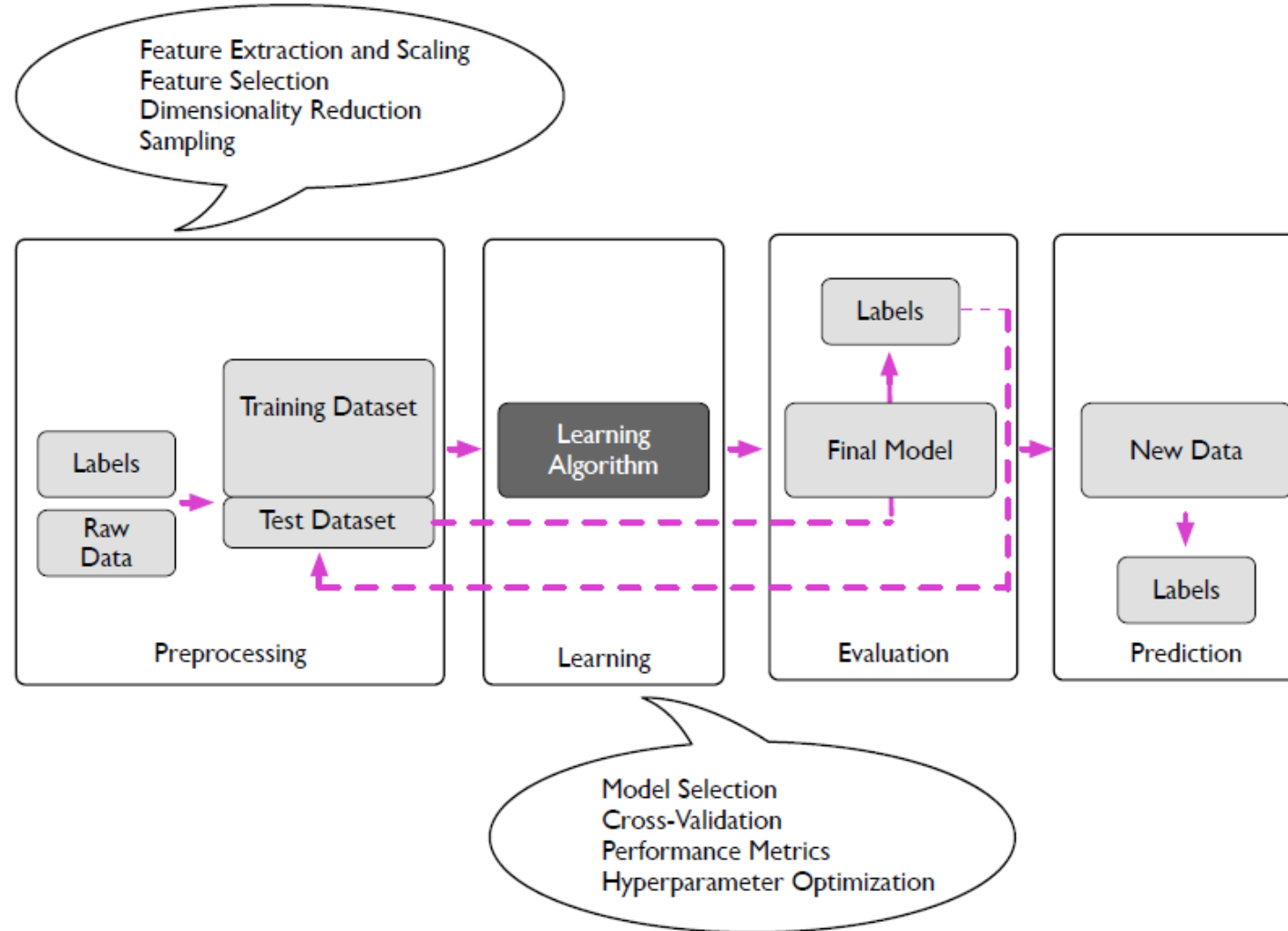
- Apriori
- Euclat
- FP-growth

Machine Learning Application Design Approach

Supervised Learning Workflow



Supervised Learning Process



Supervised Learning Process

Example steps for ML Application Design

1. Define the problem to be solved.
2. Collect (labeled) data.
3. Prepare the data.
4. Choose an algorithm class.
5. Train the model.
6. Choose a metric or measure for evaluating the model.
7. Tune parameters
8. Prediction or inference

Building Classification Model

Lab 1 Exercise

Building classification models is one of the most important data science use cases. Classification models are models that predict a categorical label. A few examples of this include predicting whether a customer will churn or whether a bank loan will default. In this guide, you will learn how to build and evaluate a classification model in Python. We will train the logistic regression algorithm, which is one of the oldest yet most powerful classification algorithms.

https://drive.google.com/drive/folders/1QUaVwd_P6_C-hXGBG5op59fH4maM1Un?usp=drive_link

Exercise: Train a Simple Image Classifier with Teachable Machine

Step 1: Open Teachable Machine

Go to:

<https://teachablemachine.withgoogle.com/>

Click “Get Started” > “Image Project” > “Standard Image Model.”

Step 2: Define Your Classes

Create two or three classes (e.g.):

- Class 1: Thumbs up
- Class 2: Thumbs down
- (Optional) Class 3: Hand open

Rename the classes for clarity.

Step 3: Train the Model

1. Click the “Webcam” button under each class.
2. Record about 50–100 images for each class (move your hand slightly to create variety).
3. After collecting data for all classes, click the “Train Model” button.

Step 4: Test Your Model

- Try showing your gestures again on camera.
- See if the model correctly identifies your input.

Conclusion

- We now have some basic understanding of the process to build an ML application.