UNIT I: Introduction to Machine Learning,
History and Overview of machine learning, Applications, Types of Machine Learning, Basic Concepts.
Concept Learning and candidate elimination learning Algorithm.

Developed Design By Prof. Makhan Kumbhkar

Agenda

01 Introduction to Machine Learning

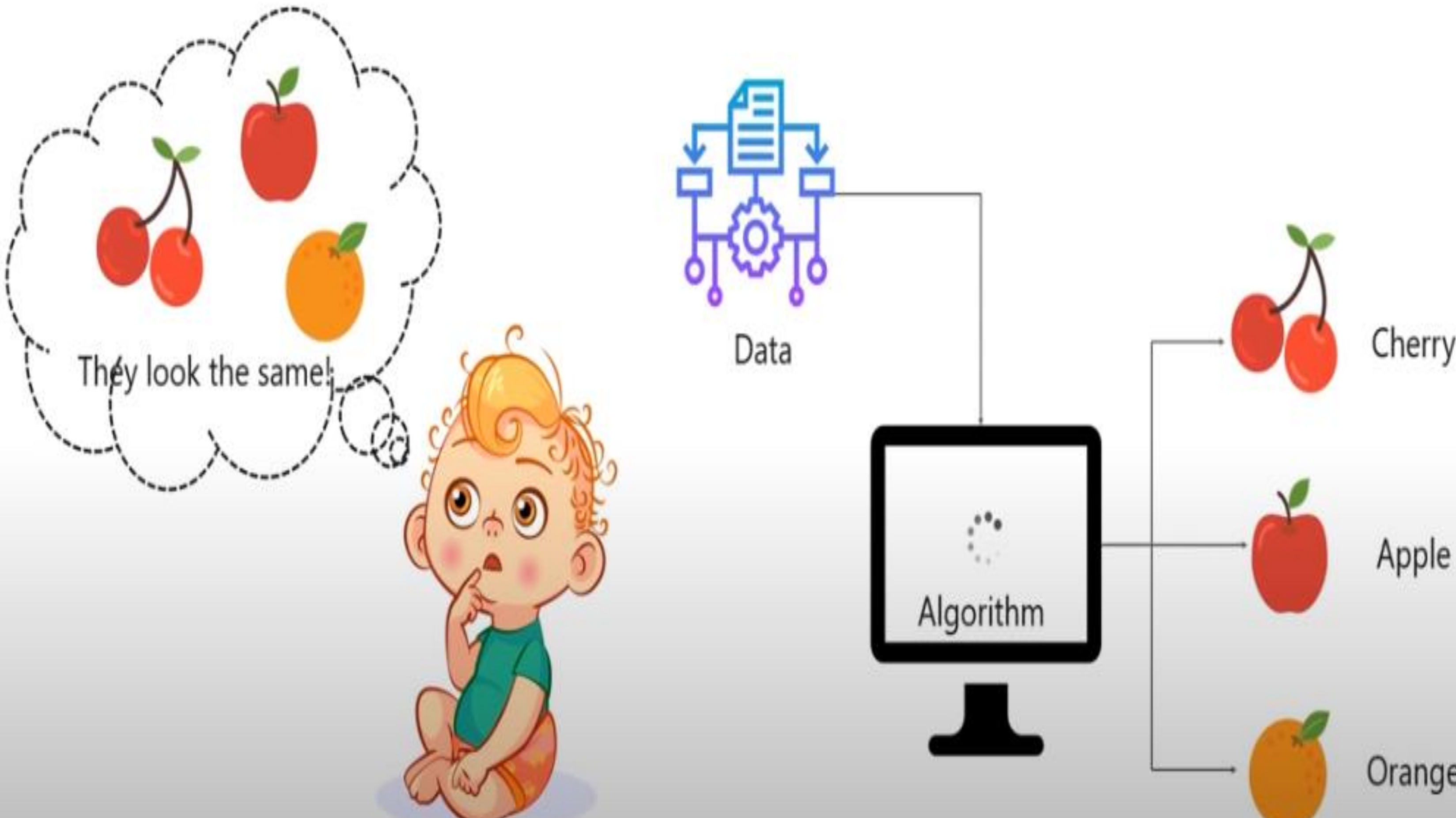
02 Types of Machine Learning

03 Supervised vs Unsupervised vs Reinforcement learning

04 Use Cases

What Is Machine Learning?

Machine learning is a subset of artificial intelligence (AI) which provides machines the ability to learn automatically & improve from experience without being explicitly programmed.



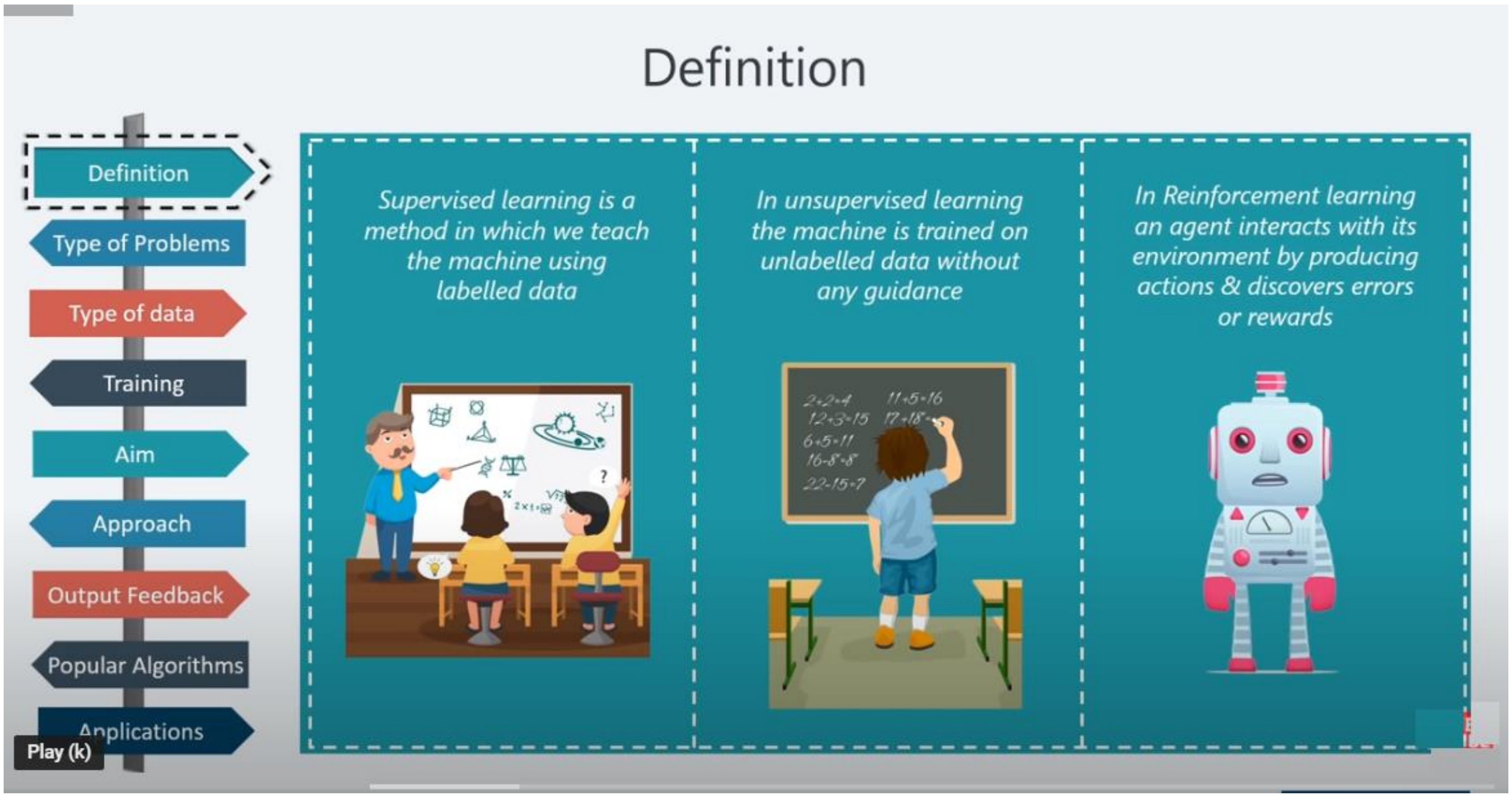
Types Of Machine Learning

Supervised Learning

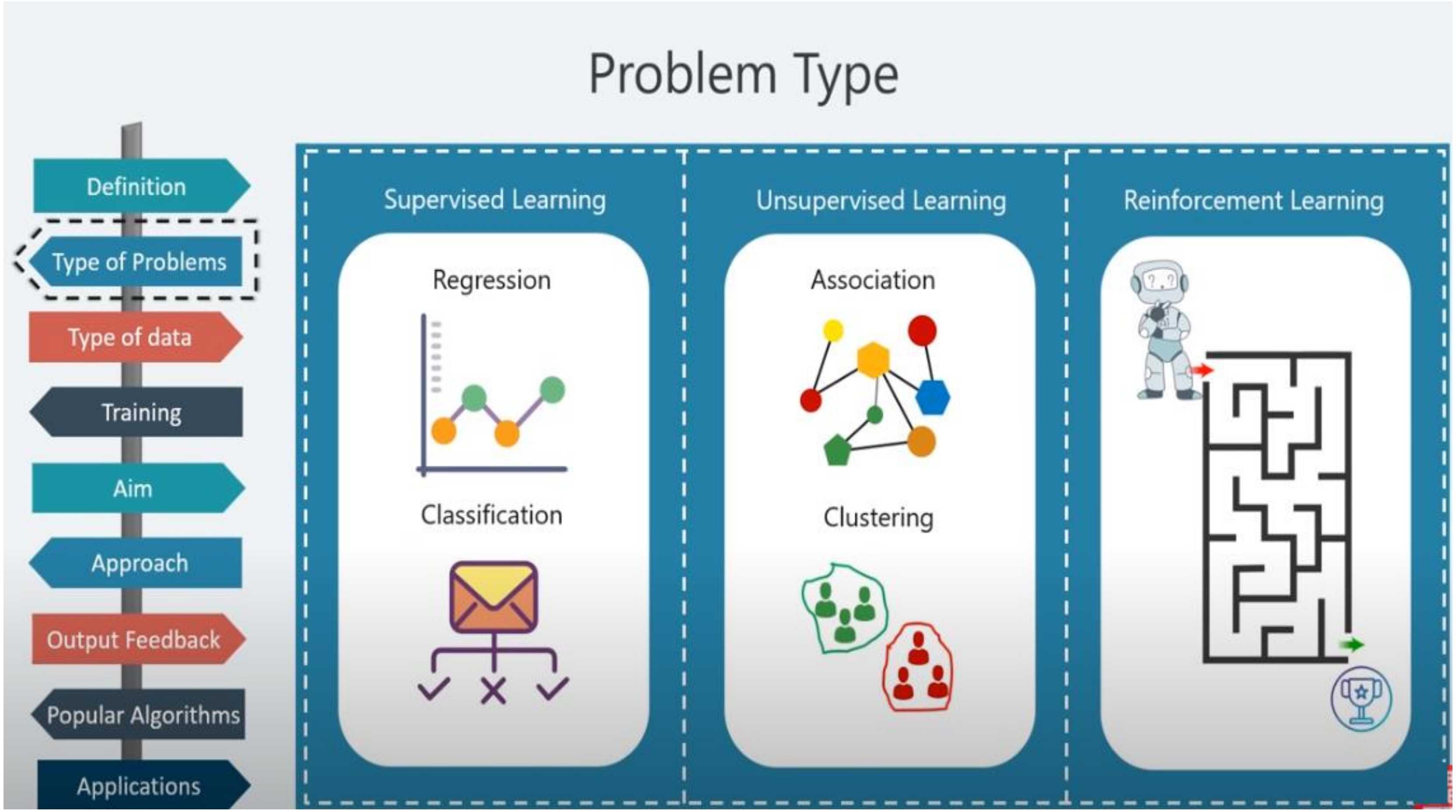
Unsupervised Learning

Reinforcement Learning

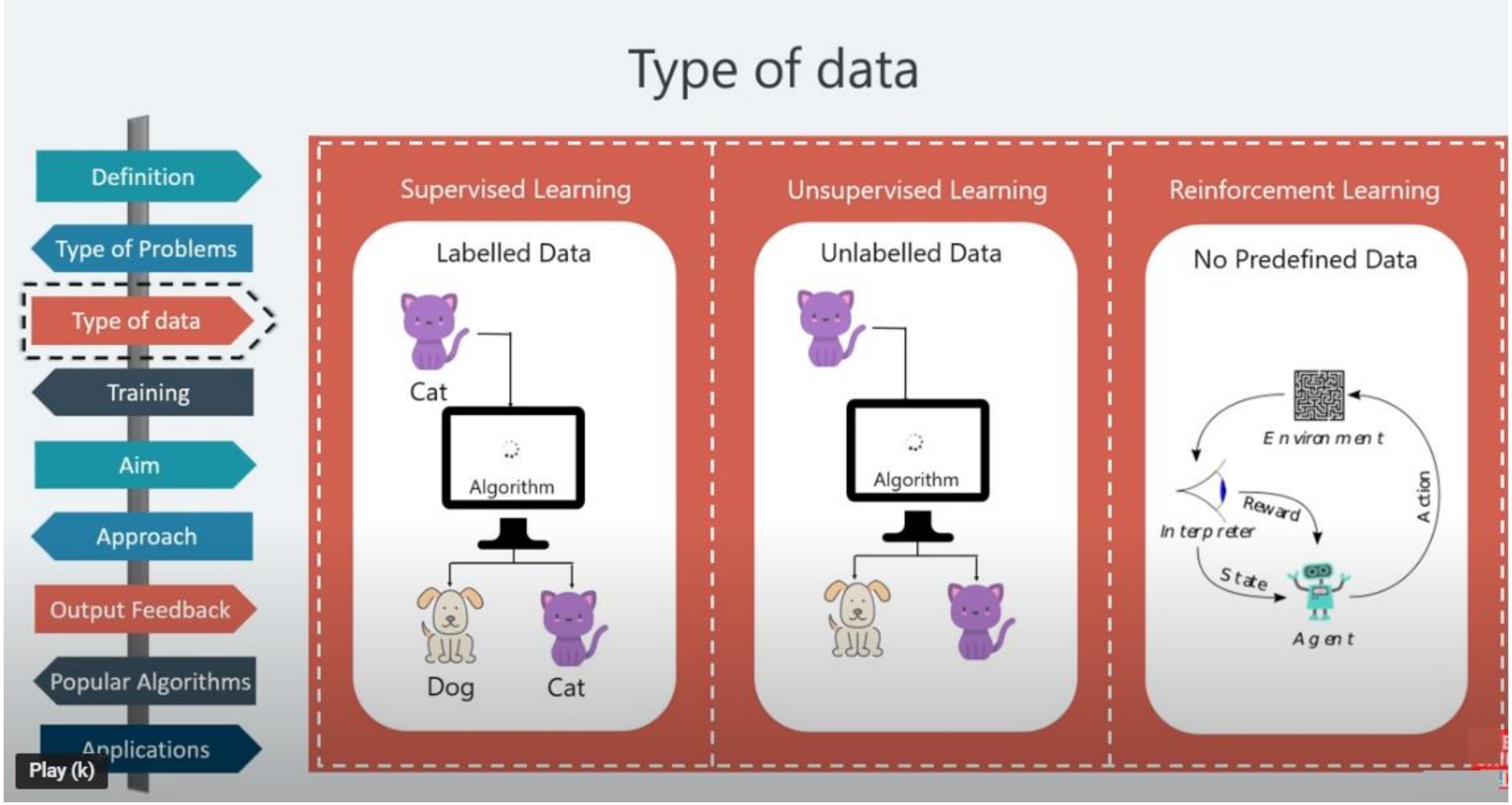
Difference Between Types of M/c Learning



Problem Types



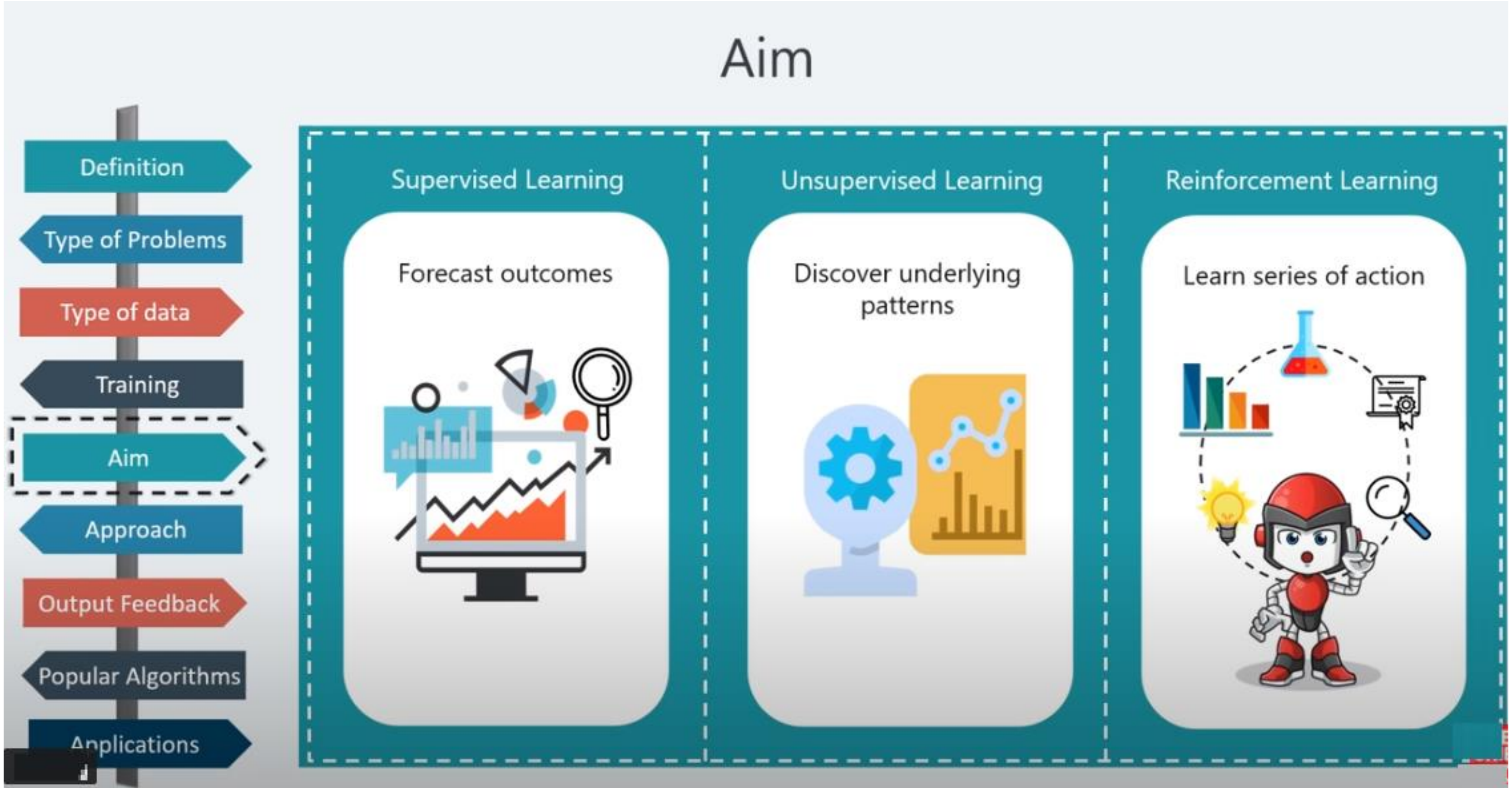
Types of Data



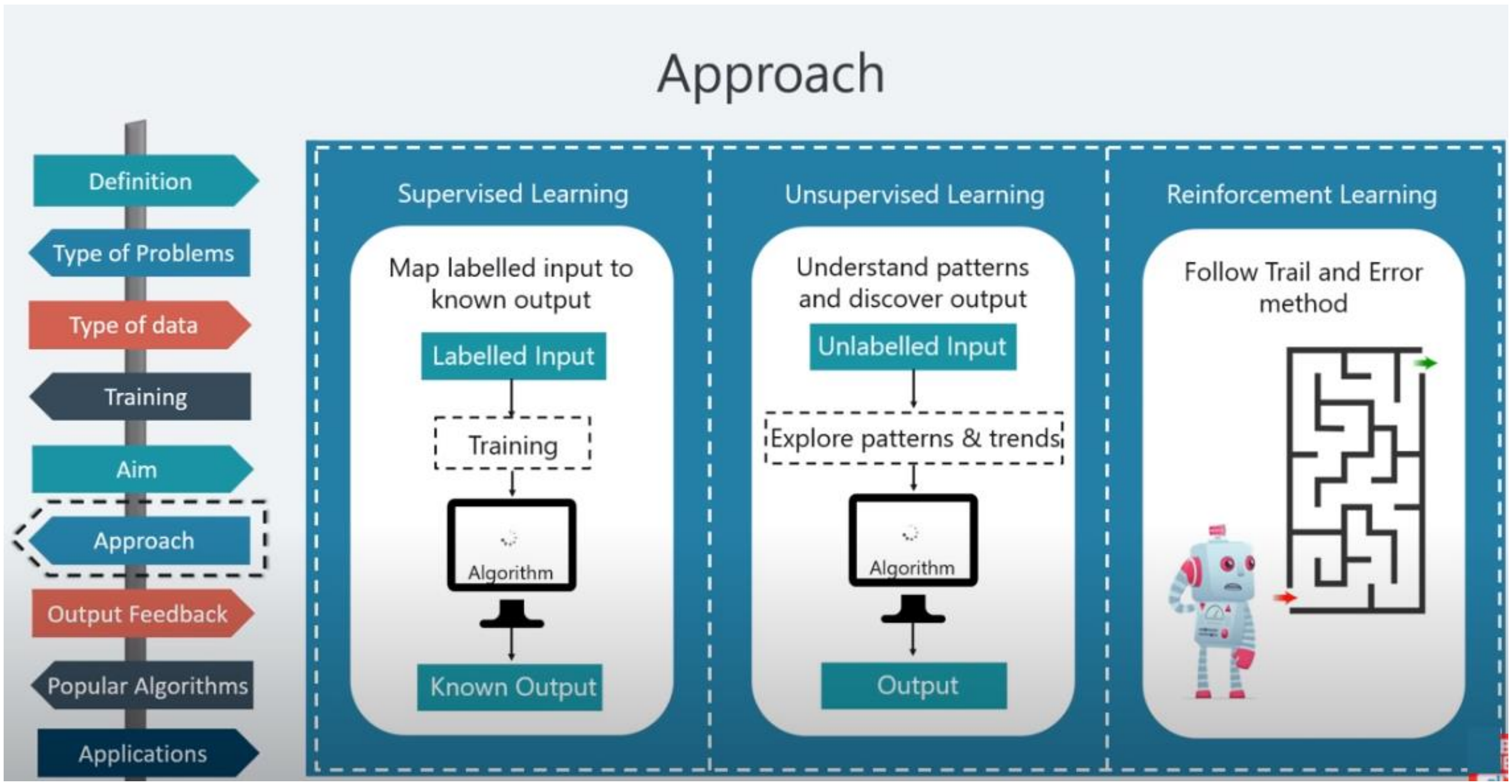
Training


Training
Definition
Type of Problems
Type of data
Training
Aim
Approach
Output Feedback
Popular Algorithms
Applications
Supervised Learning
External supervision
Unsupervised Learning
No supervision
Reinforcement Learning
No supervision

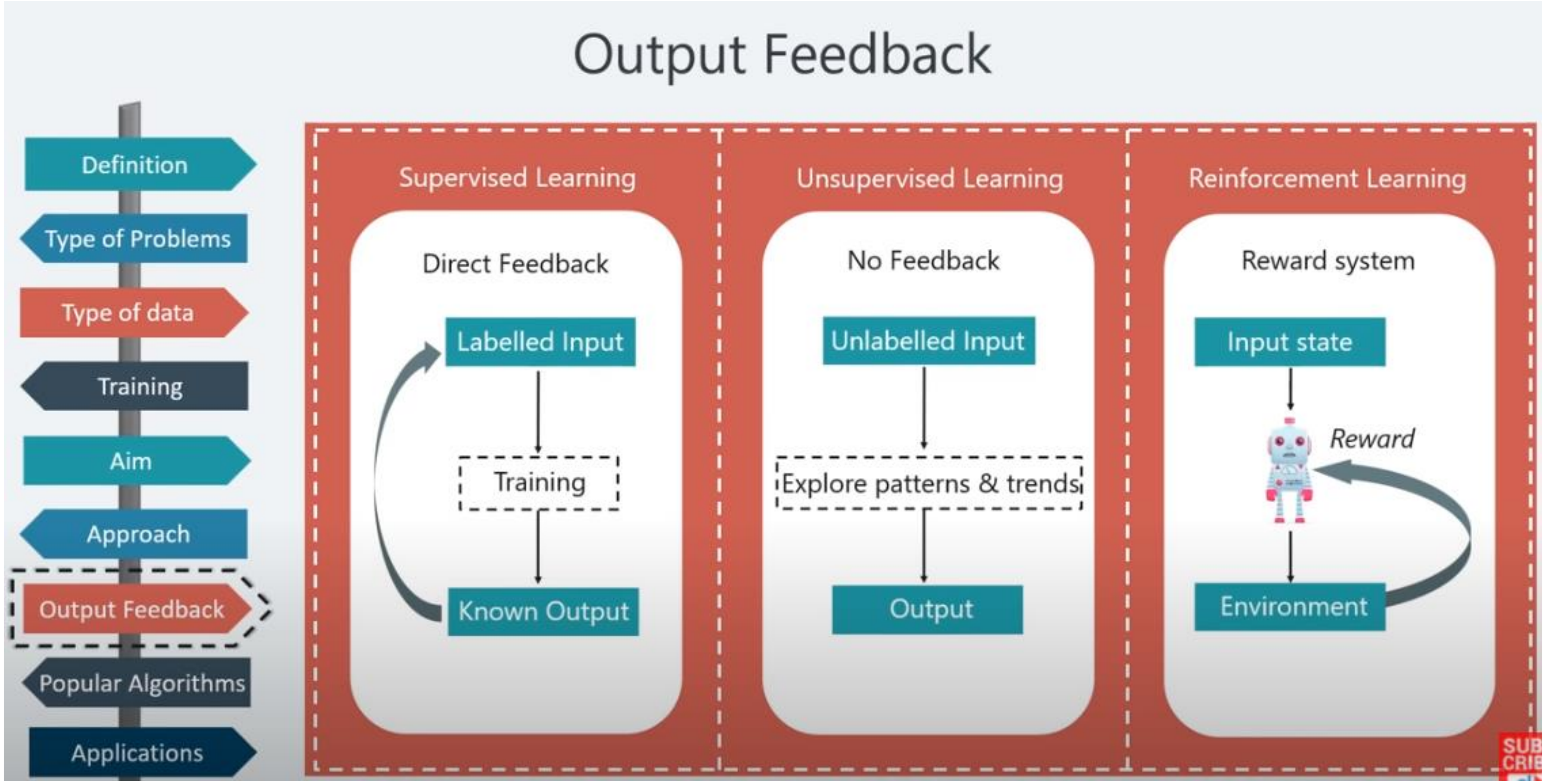
Aim


Aim
Definition
Type of Problems
Type of data
Training
Aim
Approach
Output Feedback
Popular Algorithms
Applications
Supervised Learning
Forecast outcomes
Unsupervised Learning
Discover underlying patterns
Reinforcement Learning
Learn series of action

Approach

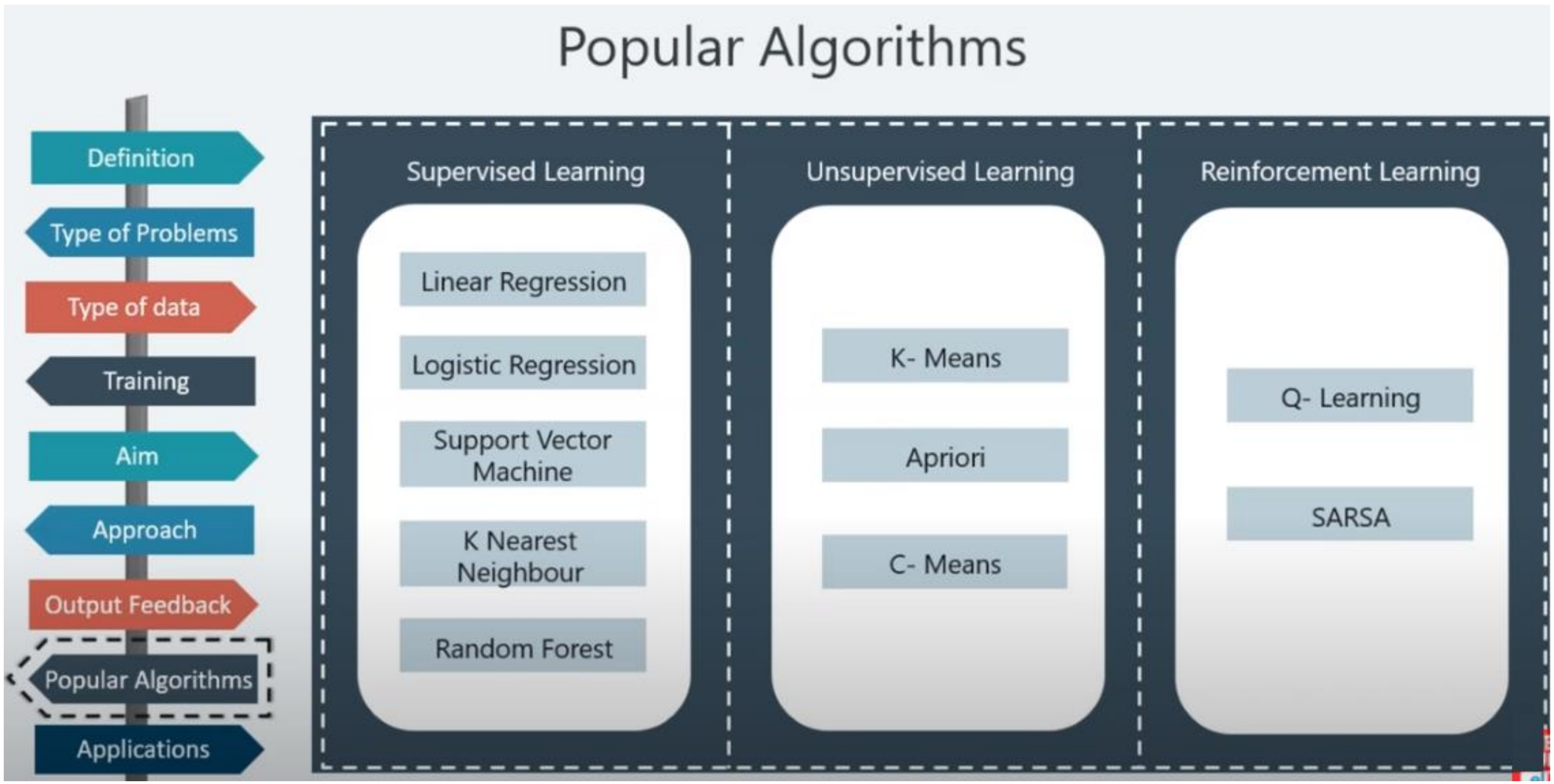

Approach
Definition
Type of Problems
Type of data
Training
Aim
Approach
Output Feedback
Popular Algorithms
Applications
Supervised Learning
Map labelled input to known output
Labelled Input
Training
Algorithm
Known Output
Unsupervised Learning
Understand patterns and discover output
Unlabelled Input
Explore patterns & trends
Algorithm
Output
Reinforcement Learning
Follow Trail and Error method

Approach

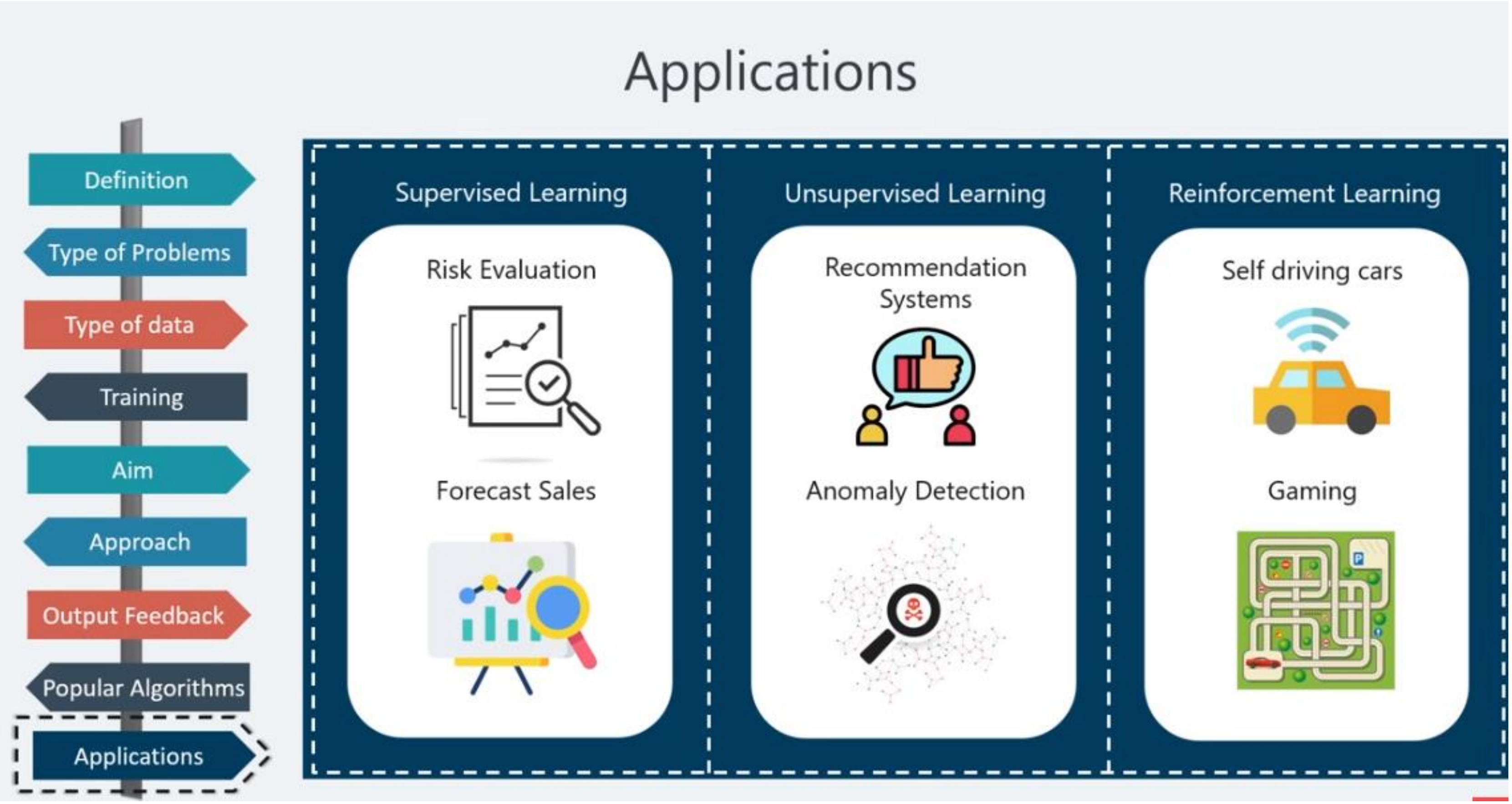


Output Feedback
Definition
Type of Problems
Type of data
Training
Aim
Approach
Output Feedback
Popular Algorithms
Applications
Supervised Learning
Direct Feedback
Labelled Input
Training
Known Output
Unsupervised Learning
No Feedback
Unlabelled Input
Explore patterns & trends
Output
Reinforcement Learning
Reward system
Input state
Reward
Environment

Algorithms


Popular Algorithms
Definition
Type of Problems
Type of data
Training
Aim
Approach
Output Feedback
Popular Algorithms
Applications
Supervised Learning
Linear Regression
Logistic Regression
Support Vector Machine
K Nearest Neighbour
Random Forest
Unsupervised Learning
K- Means
Apriori
C- Means
Reinforcement Learning
Q- Learning
SARSA

Applications

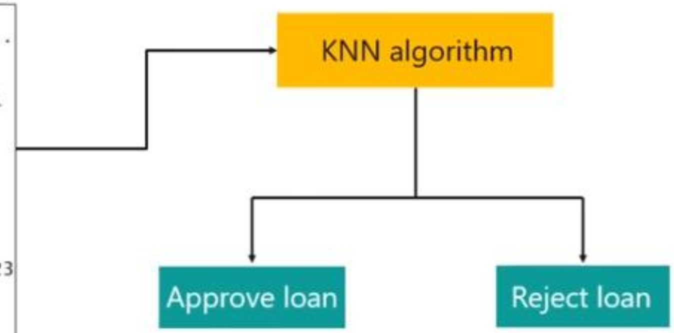


Use Cases

Use Case 1

Problem Statement: Study a bank credit dataset and make a decision about whether to approve the loan of an applicant based on his profile

```
$ Account.Balance                  : int  1 1 2 1 1 1 1 1 4 2 ...
$ Duration.of.Credit..month.       : int  18 9 12 12 12 10 8 6 18 24 ..
$ Payment.Status.of.Previous.Credit: int  4 4 2 4 4 4 4 4 4 2 ...
$ Purpose                          : int  2 0 9 0 0 0 0 0 3 3 ...
$ Credit.Amount                    : int  1049 2799 841 2122 2171 2241
$ Value.Savings.Stocks             : int  1 1 2 1 1 1 1 1 1 3 ...
$ Length.of.current.employment     : int  2 3 4 3 3 2 4 2 1 1 ...
$ Instalment.per.cent              : int  4 2 2 3 4 1 1 2 4 1 ...
$ Sex...Marital.Status             : int  2 3 2 3 3 3 3 3 2 2 ...
$ Guarantors                       : int  1 1 1 1 1 1 1 1 1 1 ...
$ Duration.in.Current.address      : int  4 2 4 2 4 3 4 4 4 4 ...
$ Most.valuable.available.asset    : int  2 1 1 1 2 1 1 1 3 4 ...
$ Age..years.                      : int  21 36 23 39 38 48 39 40 65 23
$ Concurrent.Credits               : int  3 3 3 3 1 3 3 3 3 3 ...
$ Type.of.apartment                : int  1 1 1 1 2 1 2 2 2 1 ...
$ No.of.Credits.at.this.Bank       : int  1 2 1 2 2 2 2 1 2 1 ...
$ Occupation                       : int  3 3 2 2 2 2 2 2 1 1 ...
$ No.of.dependents                 : int  1 2 1 2 1 2 1 2 1 1 ...
$ Telephone                        : int  1 1 1 1 1 1 1 1 1 1 ...
$ Foreign.Worker                   : int  1 1 1 2 2 2 2 2 1 1 ...
```



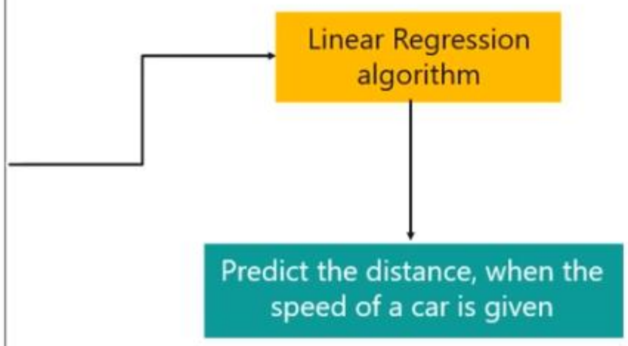
Use Cases

Use Case 2

Problem Statement: To establish a mathematical equation for **distance** as a function of **speed**, so you can use it to predict **distance** when only the **speed** of the car is known.

```
> cars
```

	speed	dist
1	4	2
2	4	10
3	7	4
4	7	22
5	8	16
6	9	10
7	10	18
8	10	26
9	10	34
10	11	17
11	11	28
12	12	14
13	12	20
14	12	24
15	12	28

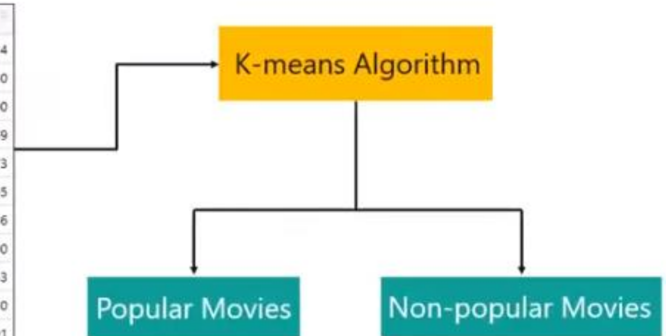


Use Cases

Use Case 3

Problem Statement: To cluster a set of movies as either good or average based on their social media out reach

	director_facebook_likes	actor_3_facebook_likes	actor_1_facebook_likes	cast_total_facebook_likes
AvatarÂ	0	855	1000	4834
Pirates of the C...	563	1000	40000	48350
SpectreÂ	0	161	11000	11700
The Dark Knigh...	22000	23000	27000	106759
John CarterÂ	475	530	640	1873
Spider-Man 3Â	0	4000	24000	46055
TangledÂ	15	284	799	2036
Avengers: Age ...	0	19000	26000	92000
Harry Potter an...	282	10000	25000	58753
Batman v Super...	0	2000	15000	24450
Superman Retur...	0	903	18000	29991
Quantum of Sol...	395	393	451	2023
Pirates of the C...	563	1000	40000	48486

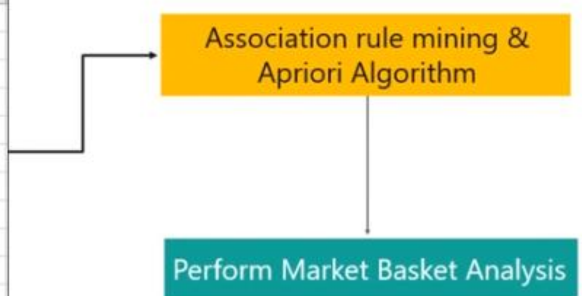


Use Cases

Use Case 4

Problem Statement: To perform Market Basket Analysis by finding association between items bought at the grocery store

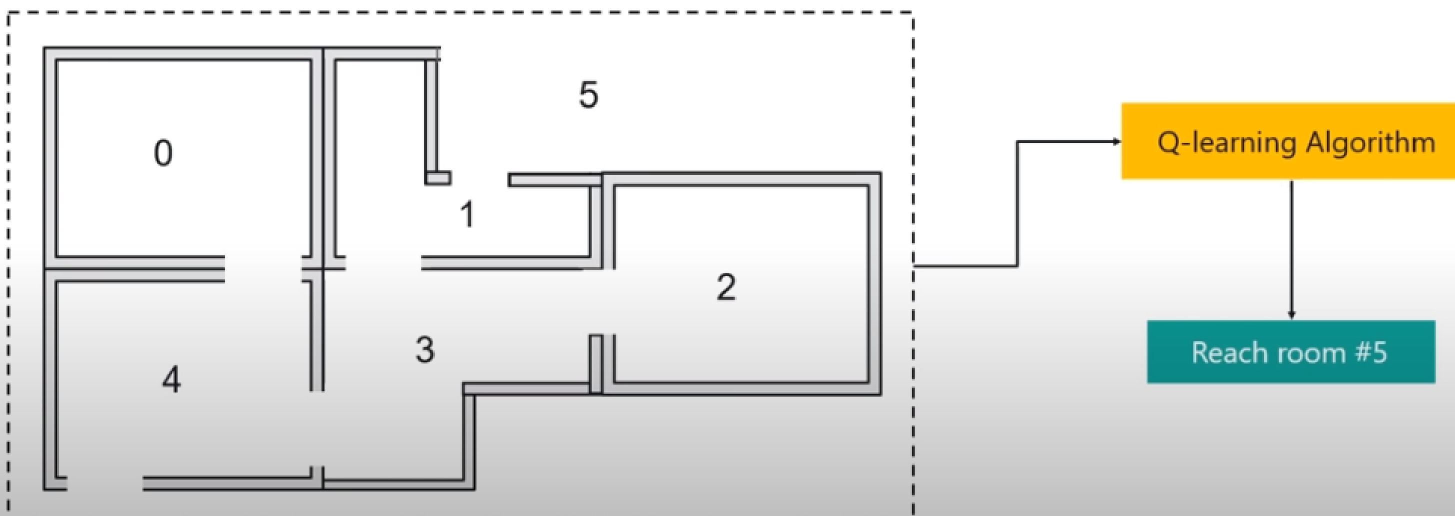
	A	B	C	D	E	F	G	H
1	citrus fruit	semi-finisl	margarine	ready soups				
2	tropical fr	yogurt	coffee					
3	whole milk							
4	pip fruit	yogurt	cream che	meat spreads				
5	other vege	whole mil	condense	long life bakery product				
6	whole mil	butter	yogurt	rice	abrasive cleaner			
7	rolls/buns							
8	other vege	UHT-milk	rolls/buns	bottled be	liquor (appetizer)			
9	pot plants							
10	whole mil	cereals						
11	tropical fr	other vege	white brea	bottled wa	chocolate			
12	citrus fruit	tropical fr	whole mil	butter	curd	yogurt	flour	bottled wa
13	beef							
14	frankfurte	rolls/buns	soda					
15	chicken	tropical fruit						



Use Cases

Use Case 5

Problem Statement: Place an agent in any one of the rooms (0,1,2,3,4) and the goal is to reach outside the building (room 5)



Concept Learning

- Concept learning also refers to a learning task in which a human or machine learner is trained to classify objects by being shown a set of example objects along with their class labels. The learner will simplify what has been observed in an example. This simplified version of what has been learned will then be applied to future examples

- Concept learning can be viewed as the task of searching through a large space of hypotheses implicitly defined by the hypothesis representation.

- The goal of this search is to find the hypothesis that best fits the training

examples.

- By selecting a hypothesis representation, the designer of the learning algorithm implicitly defines the space of all hypotheses that the program can ever represent and therefore can ever learn.

Concept Learning

Consistent Hypothesis and Version Space

An hypothesis h is **consistent** with a set of training examples D iff $h(x) = c(x)$ for each example in D

$$Consistent(h, D) \equiv (\forall \langle x, c(x) \rangle \in D)\ h(x) = c(x))$$

Example	Citations	Size	InLibrary	Price	Editions	Buy
1	Some	Small	No	Affordable	One	No
2	Many	Big	No	Expensive	Many	Yes

h1 = (?, ?, No, ?, Many) **– Consistent**

h2 = (?, ?, No, ?, ?) **– Not Consistent**

Concept Learning

Consistent Hypothesis and Version Space

- The version space $VS_{H,D}$ is the subset of the hypothesis from H *consistent* with the training example in D

$$VS_{H,D} \equiv \{h \in H \mid Consistent(h, D)\}$$

Concept Learning

***List-Then-Eliminate* algorithm**

Version space as list of hypotheses

1. *VersionSpace* $\leftarrow$ a list containing every hypothesis in *H*
2. For each training example, $\langle x, c(x) \rangle$ Remove from *VersionSpace* any hypothesis *h* for which $h(x) \neq c(x)$
3. Output the list of hypotheses in *VersionSpace*

Concept Learning

Consistent Hypothesis and Version Space

- F1 - > A, B
- F2 - > X, Y
- **Instance Space:** (A, X), (A, Y), (B, X), (B, Y) – **4 Examples**
- **Hypothesis Space:** (A, X), (A, Y), (A, **ø**), (A, ?), (B, X), (B, Y), (B, **ø**), (B, ?), (**ø**, X), (**ø**, Y), (**ø**, **ø**), (**ø**, ?), (**?**, X), (**?**, Y), (**?**, **ø**), (**?**, ?) - **16 Hypothesis**
- **Semantically Distinct Hypothesis :** (A, X), (A, Y), (A, ?), (B, X), (B, Y), (B, ?), (**?**, X), (**?**, Y (**?**, ?), (ø, **ø**) – **10**

Concept Learning

Consistent Hypothesis and Version Space

- Version Space: (A, X), (A, Y), (A, ?), (B, X), (B, Y), (B, ?), (?, X), (?, Y) (?, ?), (ø, ø),
- Training Instances

F1	F2	Target
A	X	Yes
A	Y	Yes

- Consistent Hypothesis are: (A, ?), (?, ?)

Concept Learning

***List-Then-Eliminate* algorithm**

Problems

- The hypothesis space must be finite
- Enumeration of all the hypothesis, rather inefficient