

# EE2211 Introduction to Machine Learning

## Lecture 1

Wang Xinchao  
[xinchao@nus.edu.sg](mailto:xinchao@nus.edu.sg)

**Office Hour: Monday 9:30 – 10:30 AM  
(Week 2-4, Week 10-12)**

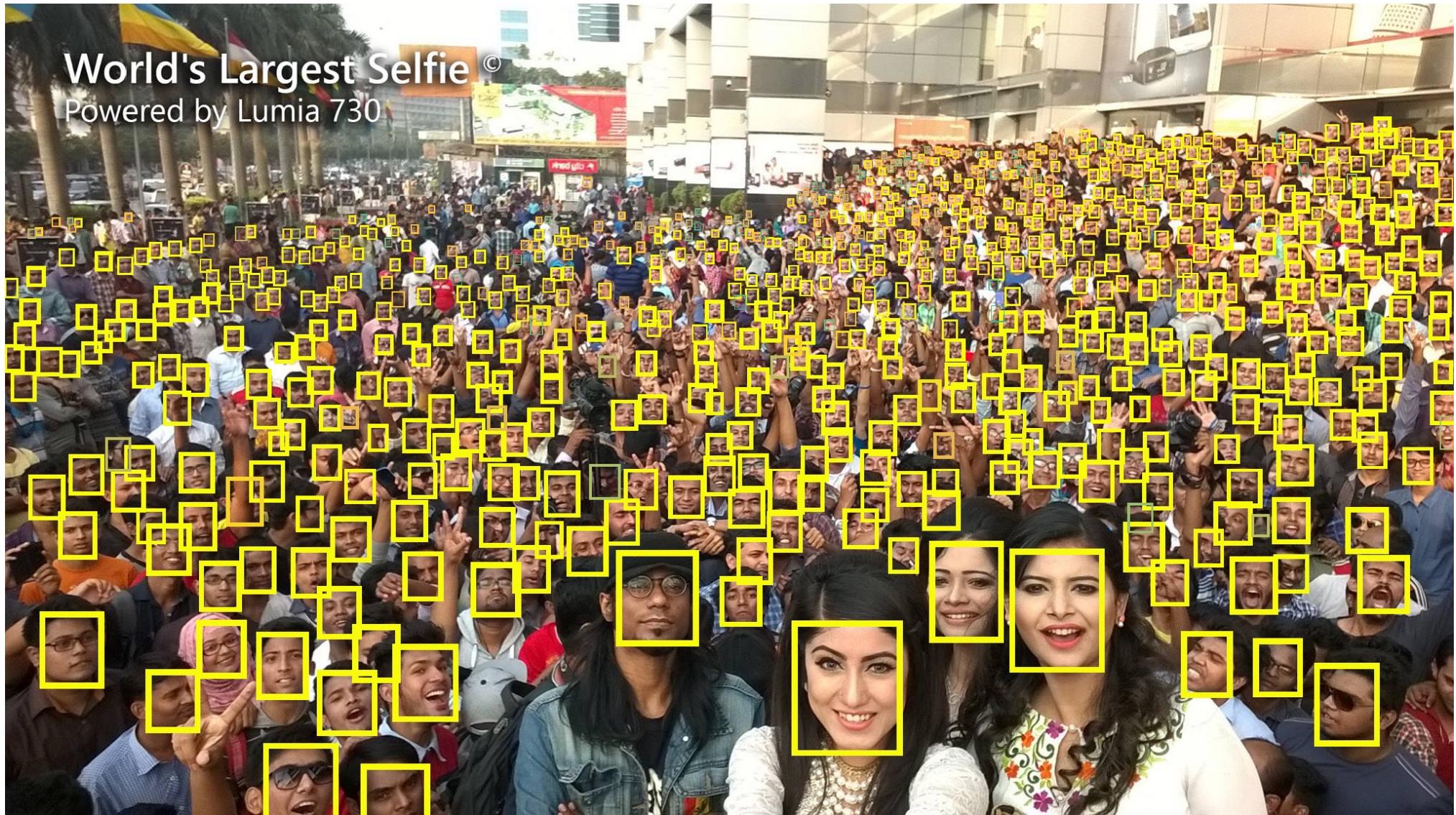
# Course Contents

- Introduction and Preliminaries (Xinchao)
  - Introduction
  - Data Engineering
  - Introduction to Probability and Statistics
- Fundamental Machine Learning Algorithms I (Yueming)
  - Systems of linear equations
  - Least squares, Linear regression
  - Ridge regression, Polynomial regression
- Fundamental Machine Learning Algorithms II (Yueming)
  - Over-fitting, bias/variance trade-off
  - Optimization, Gradient descent
  - Decision Trees, Random Forest
- Performance and More Algorithms (Xinchao)
  - Performance Issues
  - K-means Clustering
  - Neural Networks

# World's Largest Selfie



# World's Largest Selfie



# Recent Advances



ChatGPT 3.5 ▾

Sign up



You

write a poem about Singapore



ChatGPT

In a city-state where dreams ignite,  
 Singapore gleams with vibrant light,  
 Skyscrapers soar, a modern maze,  
 Where history meets tomorrow's craze.

Gardens bloom in meticulous grace,  
 Supertrees whisper, a futuristic embrace,  
 Marina Bay sparkles, a jewel aglow,  
 Reflecting a nation's spirit, in ebb and flow.

From hawker stalls to haute cuisine,  
 A melting pot where cultures convene,  
 Peranakan hues and Tamil beats,  
 Chinatown's bustle, each story repeats.



**Sora**

Prompt: A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.

# Outline

- What is machine learning?
  - Three Definition(s)
- When do we need machine learning?
  - Sometimes we need, sometimes we don't
- Applications of machine learning
- Types of machine learning
  - Supervised, Unsupervised, Reinforcement Learning
- Walking through a toy example on classification
- Inductive vs. Deductive Reasoning

# What is machine learning?

Learning is any process by which a system improves performance from experience.

- Herbert Simon

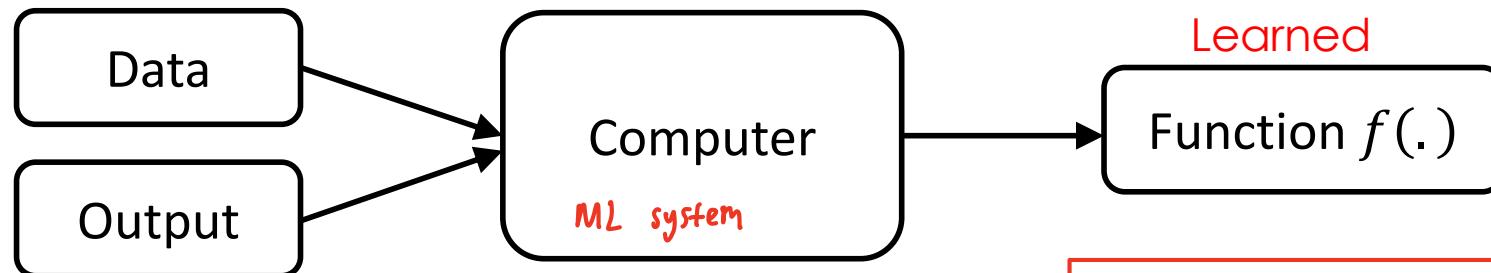
A computer program is said to learn

- from **experience E**
- with respect to some class of **tasks T**
- and **performance measure P**,

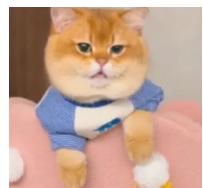
if its performance at tasks in T, as measured by P, improves with experience E.

- Tom Mitchell

# Machine Learning (Supervised Learning)



Data      Output



Cat



Dog

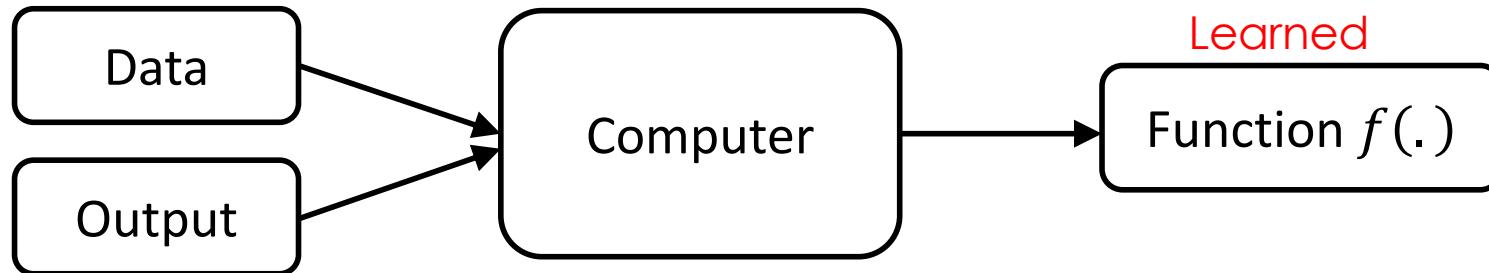
→  $f(\cdot)$  such that  
 $\downarrow$   
training

$$f(\text{Data}) = \text{output}$$

A photograph of the same orange cat from the first image.
$$f(\text{Cat}) = \text{'cat'}$$

A photograph of the same brown dog from the first image.
$$f(\text{Dog}) = \text{'dog'}$$

# Machine Learning (Supervised Learning)



Data      Output

classification



Cat

:



Dog

$\longrightarrow f(\cdot)$

When applied



$f($       )  $\longrightarrow$  Cat !

New image  $\rightarrow$  feeding new image  
by testing/inference

Machine Learning: field of study that gives computers the ability to learn without being explicitly programmed

- Arthur Samuel

# AI, Machine Learning, and Deep Learning

## ARTIFICIAL INTELLIGENCE

Any technique which enables computers to mimic human behavior



1950's    1960's    1970's    1980's    1990's    2000's    2010s

## MACHINE LEARNING

AI techniques that give computers the ability to learn without being explicitly programmed to do so



## DEEP LEARNING

A subset of ML which make the computation of multi-layer neural networks feasible

↓  
type of  
ML



Example of AI but not ML: Deductive Reasoning

↳ we are not "learning" anything so not considered ML

NUS is in Singapore, Singapore is in Asia -> NUS is in Asia

# When do we need machine learning?

Lack of human expertise  
(Navigating on Mars)



Involves huge amount of data  
(Genomics)



**Learning is not always useful:**

No need to “learn” to calculate payroll!

My Salary = Days\_of\_work \* Daily Salary + Bonus

# Application of Machine Learning

\*\*\*\*\*

Task T, Performance P, Experience E

T: Digit Recognition  
 P: Classification Accuracy  
 E: Labelled Images

(Given) Data	Output (Label)
4	“four”
3	“three”

Labels -> Supervision!

Digits 0~9



# Application of Machine Learning

(Goal)

Task T, Performance P, Experience E

T: Email Categorization

P: Classification Accuracy

E: Email Data, Some Labelled → supervised learning  
(refer to pg 26)



# Application of Machine Learning

(Goal)

Task T, Performance P, Experience E

T: Playing Go Game

P: Chances of Winning

E: Records of Past Games



# Application of Machine Learning

Task T, Performance P, Experience E

No supervision because we do not know how Covid-19 spreads

T: Identifying Covid-19 Clusters

P: Small Internal Distances

Larger External Distances

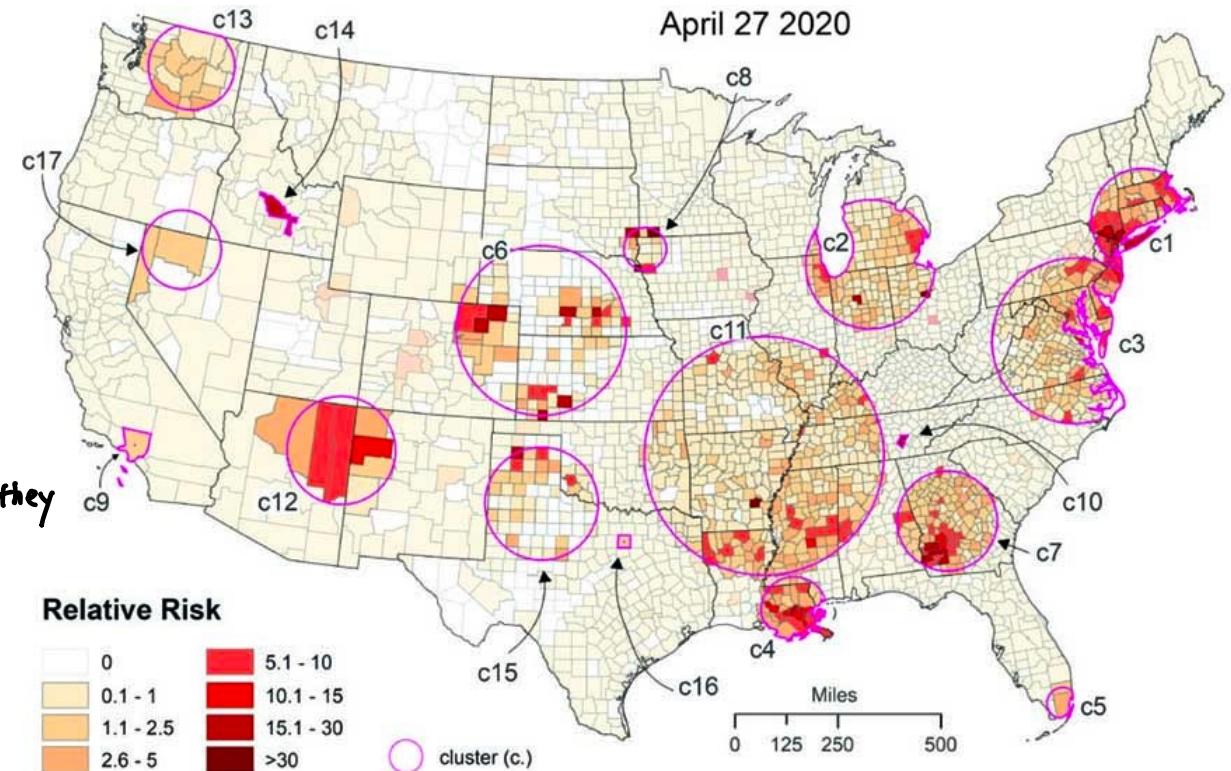
E: Records of Patients

To measure if cluster good (P) ie, C<sub>6</sub>

1) If pick any two counties from one cluster, they should be close

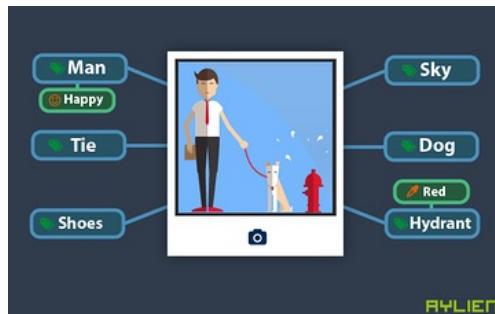
2) If pick any two counties from different clusters, they should be far away

ie, C<sub>6</sub> and C<sub>12</sub>





*Web Search Engine*



*Photo Tagging*



*Product Recommendation*



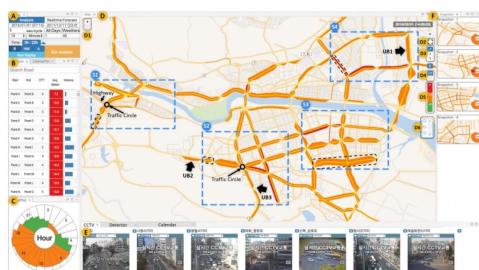
*Virtual Personal Assistant*



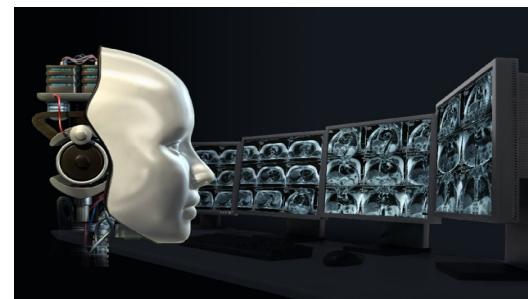
*Language Translation*



*Portfolio Management*



*Traffic Prediction*



*Medical Diagnosis*



*Algorithmic Trading*

# Types of Machine Learning

## Supervised Learning

Input:

- 1) Training Samples,
- 2) Desired Output  
(Teacher/Supervision)

Output:

A rule that maps input to output

## Unsupervised Learning

Input:

Samples

Output:

Underlying patterns in data

## Reinforcement Learning

Input:

Sequence of States,  
Actions, and  
Delayed Rewards

Output:

Action Strategy: a rule  
that maps the  
environment to action

# Types of Machine Learning

## Supervised Learning

Input:

- 1) Training Samples,
- 2) Desired Output  
(Teacher/Supervision)

Output:

A rule that maps input to output

Data      Output



Cat

:

$$\longrightarrow f(\cdot) \text{ such that}$$



Dog

## Unsupervised Learning

Input:

Samples

Output:

Underlying patterns in data

## Reinforcement Learning

Input:

Sequence of States,  
Actions, and  
Delayed Rewards

Output:

Action Strategy: a rule  
that maps the  
environment to action

$$f(\text{cat}) = \text{'cat'}$$

$$f(\text{dog}) = \text{'dog'}$$

# Supervised Learning

Data

$x$

Regression

$x$

Classification

Output

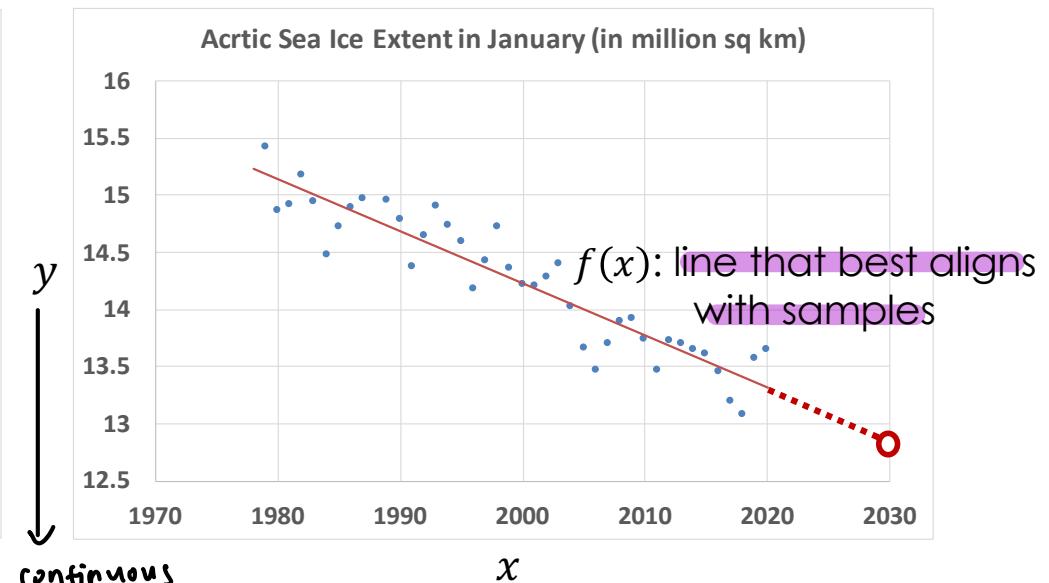
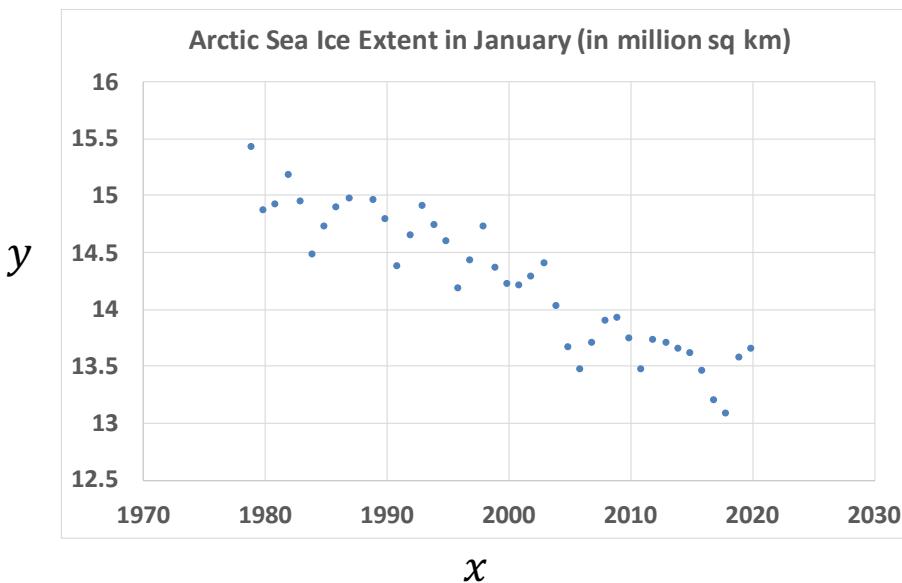
$y$  (Continuous)

$y$  (Categorical)

↓  
label is used for  
categorical, not for  
continuous

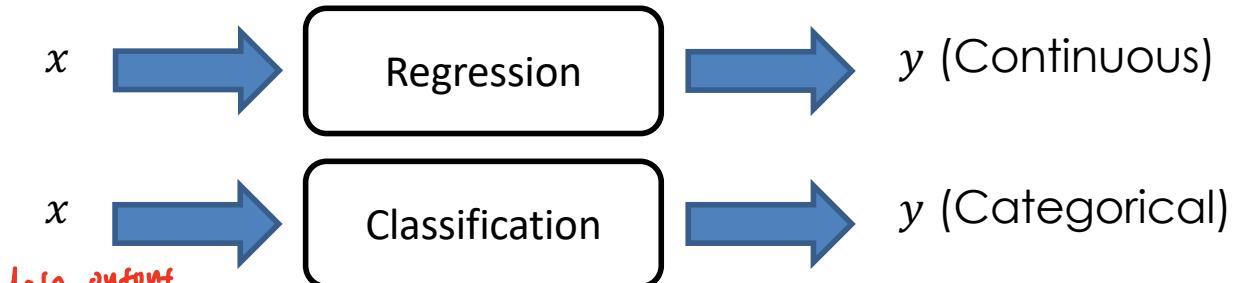
## Regression

- Given  $(\underline{x}_1, \underline{y}_1), (\underline{x}_2, \underline{y}_2), \dots, (\underline{x}_N, \underline{y}_N)$
- Learn a function  $f(x)$  to predict real-valued  $y$  given  $x$



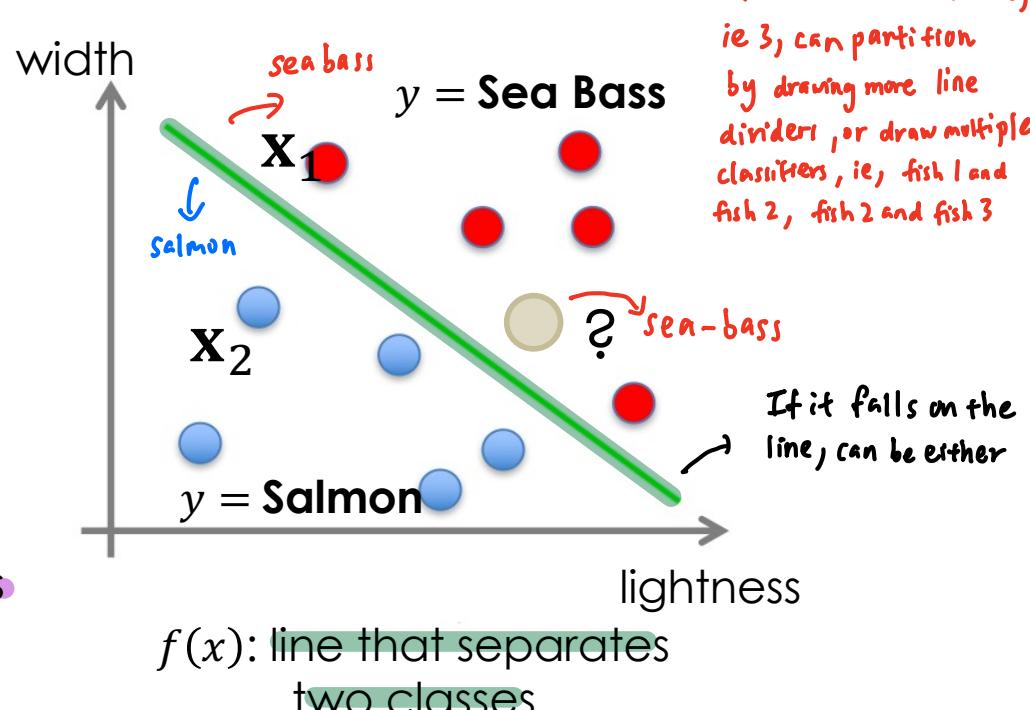
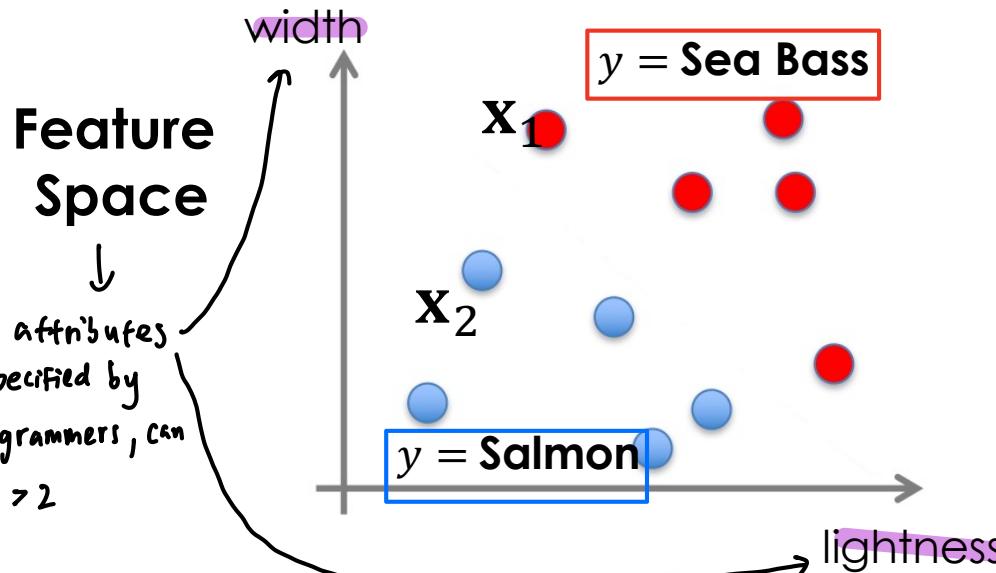
# Supervised Learning

## Data



## Classification

- Given  $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)$
- Learn a function  $f(\mathbf{x})$  to predict categorical  $y$  given  $\mathbf{x}$



# Types of Machine Learning

## Supervised Learning

Input:

- 1) Training Samples,
- 2) Desired Output  
(Teacher/Supervision)

Output:

A rule that maps input to output

## Unsupervised Learning

Input:

Samples

Output:

Underlying patterns in data

## Reinforcement Learning

Input:

Sequence of States,  
Actions, and  
Delayed Rewards

Output:

Action Strategy: a rule  
that maps the  
environment to action

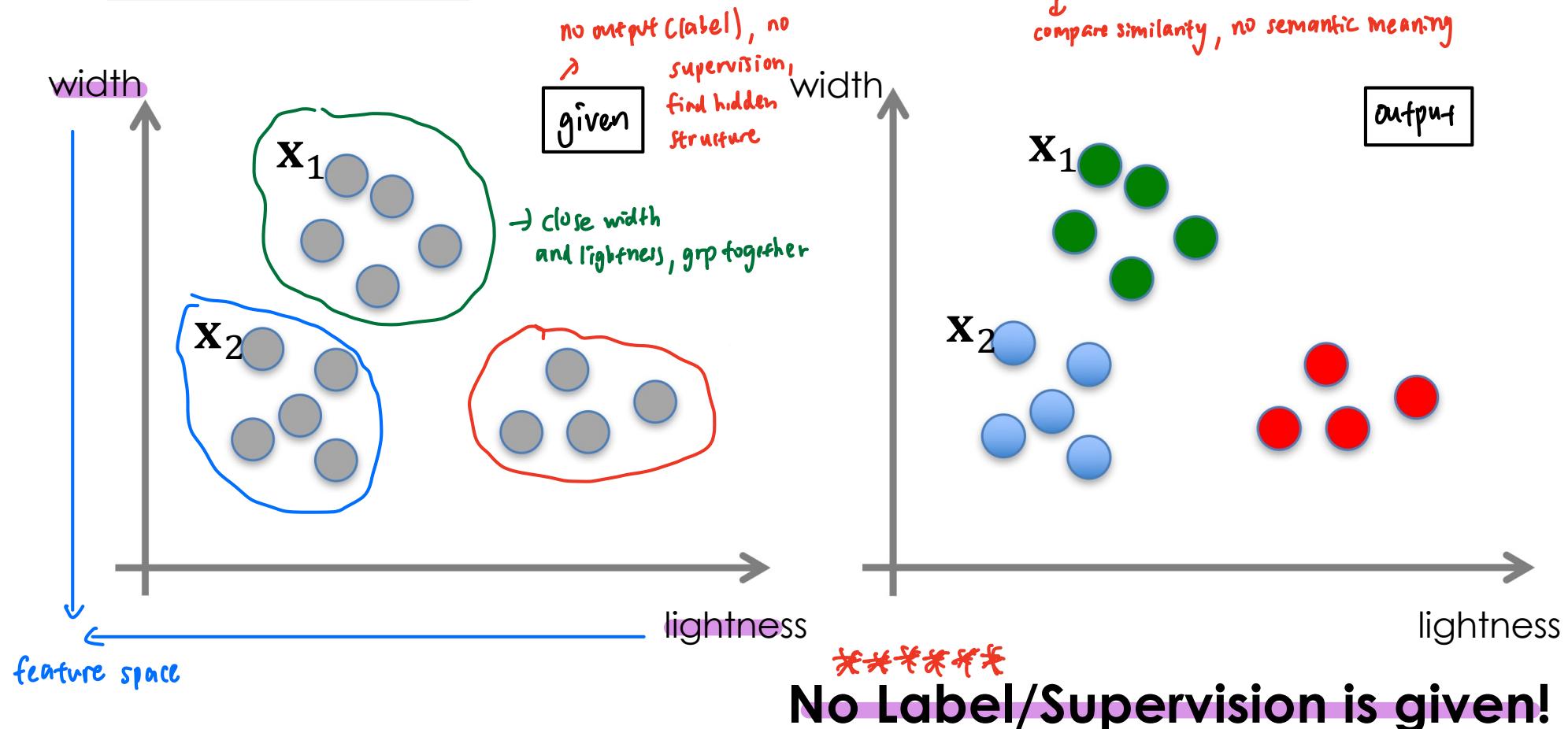
**Key different w.r.t. supervised learning:  
No Label/Supervision is given!**

# Unsupervised Learning

(grouping / clustering)

## Clustering

- Given  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ , without labels
- Output Hidden Structure Behind



# Types of Machine Learning

## Supervised Learning

Input:

- 1) Training Samples,
- 2) Desired Output  
(Teacher/Supervision)

Output:

A rule that maps input to output

## Unsupervised Learning

Input:

Samples

Output:

Underlying patterns in data

## Reinforcement Learning

Input:

Sequence of States,  
Actions, and  
Delayed Rewards

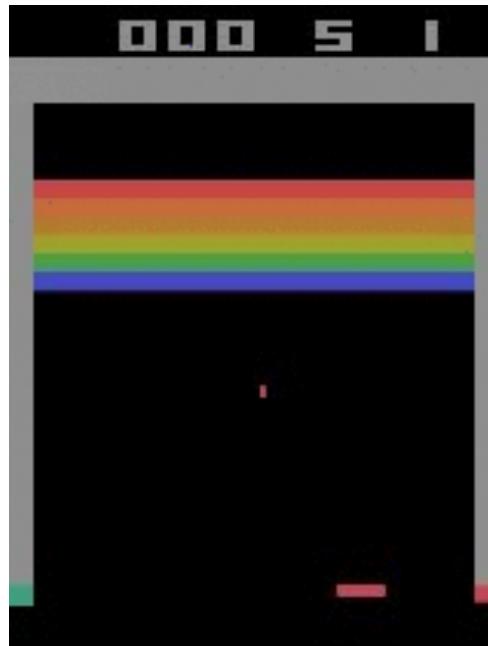
Output:

Action Strategy: a rule  
that maps the  
environment to action

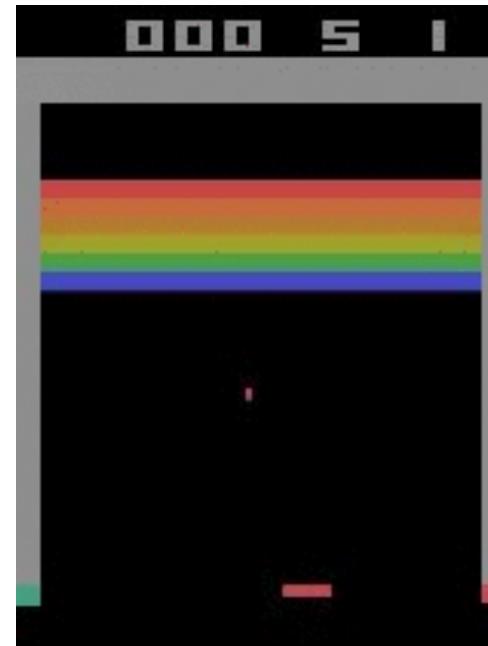
every step, it knows what  
action to take

# Reinforcement Learning

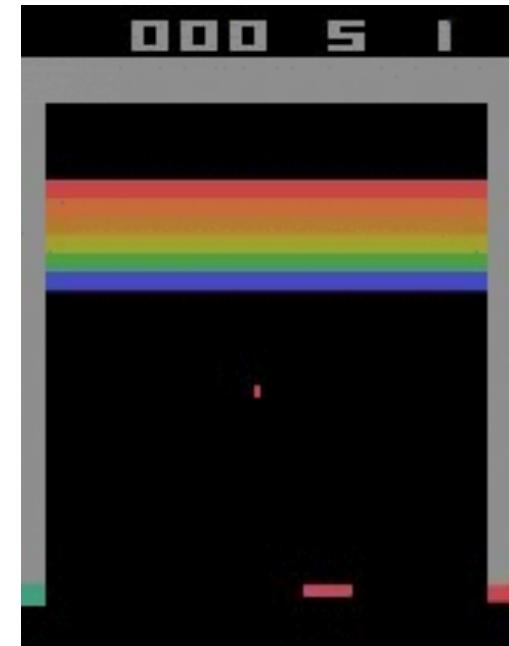
## Breakout Game



Initial Performance



Training 15 minutes

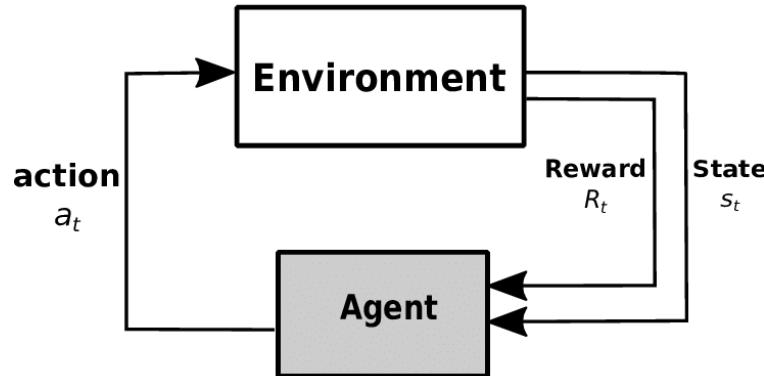


Training 30 minutes

# Reinforcement Learning

↳ check for sequence of states to identify RL

- Given sequence of states  $S$  and actions  $A$  with (delayed) rewards  $R$
- Output a policy  $\pi(a, s)$ , to guide us what action  $a$  to take in state  $s$



**S:** Ball Location,  
Paddle Location, Bricks

**A:** left, right

**R:**

- positive reward  
Knocking a brick, clearing all bricks
- negative reward  
Missing the ball
- zero reward  
Cases in between

# Supervised Unsupervised Reinforcement

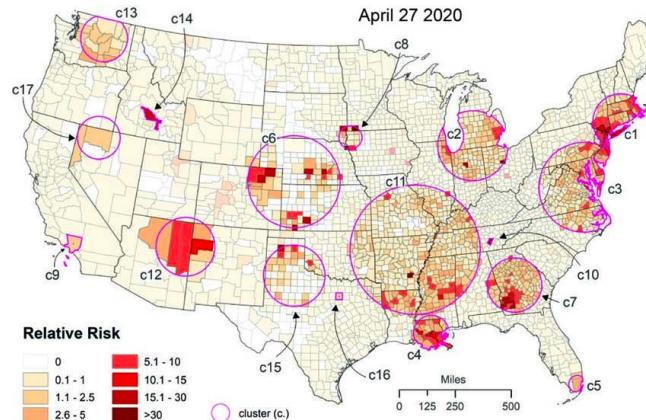
# Quiz Time!

0 0 0 1 1 1 1 1 1 2  
 2 2 2 2 2 2 3 3 3  
 3 4 4 4 4 4 5 5 5  
 6 6 7 7 7 7 8 8 8  
 8 8 8 8 9 9 9 9 9

**Supervised** → need initial input data



**Supervised** → by removing  
spam emails from  
spam, we are "supervising" ML



**Unsupervised** → only given Covid records and  
we want to find out clusters

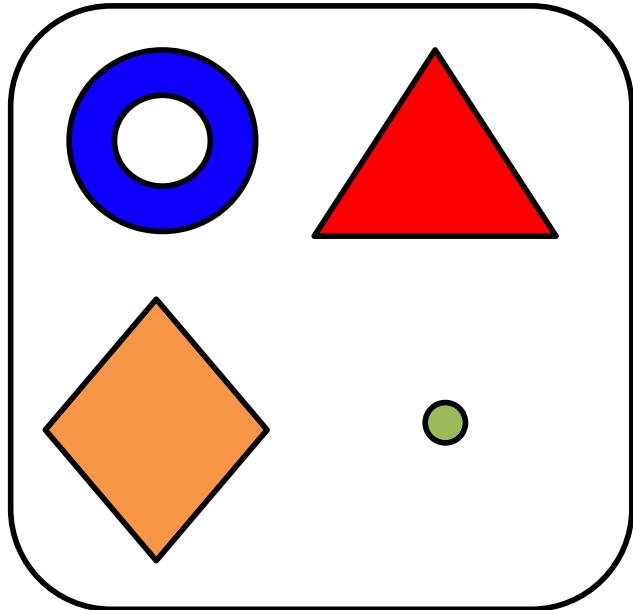


**Reinforcement**

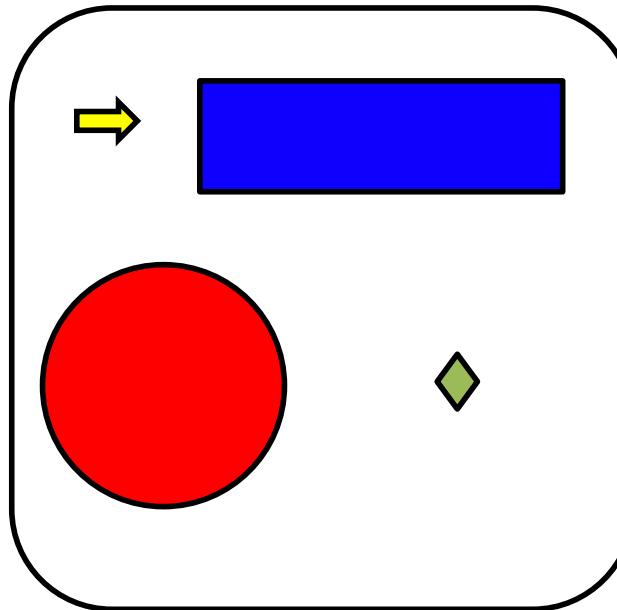
(sequence of states)

↳ AlphaGo needs to decide what action to take

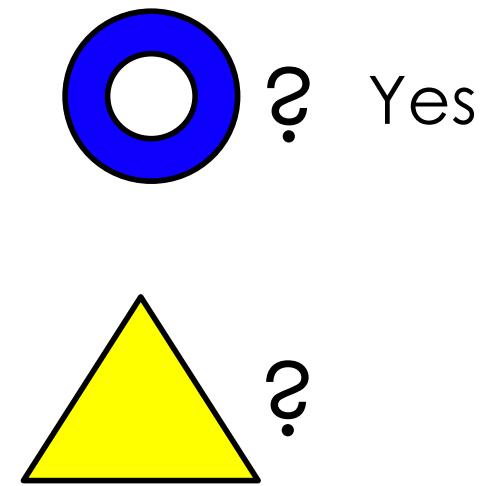
# Walking Through A Toy Example: Token Classification



Yes



No



?

?

**Step 1: Feature Extraction**  
Extract Attributes of Samples

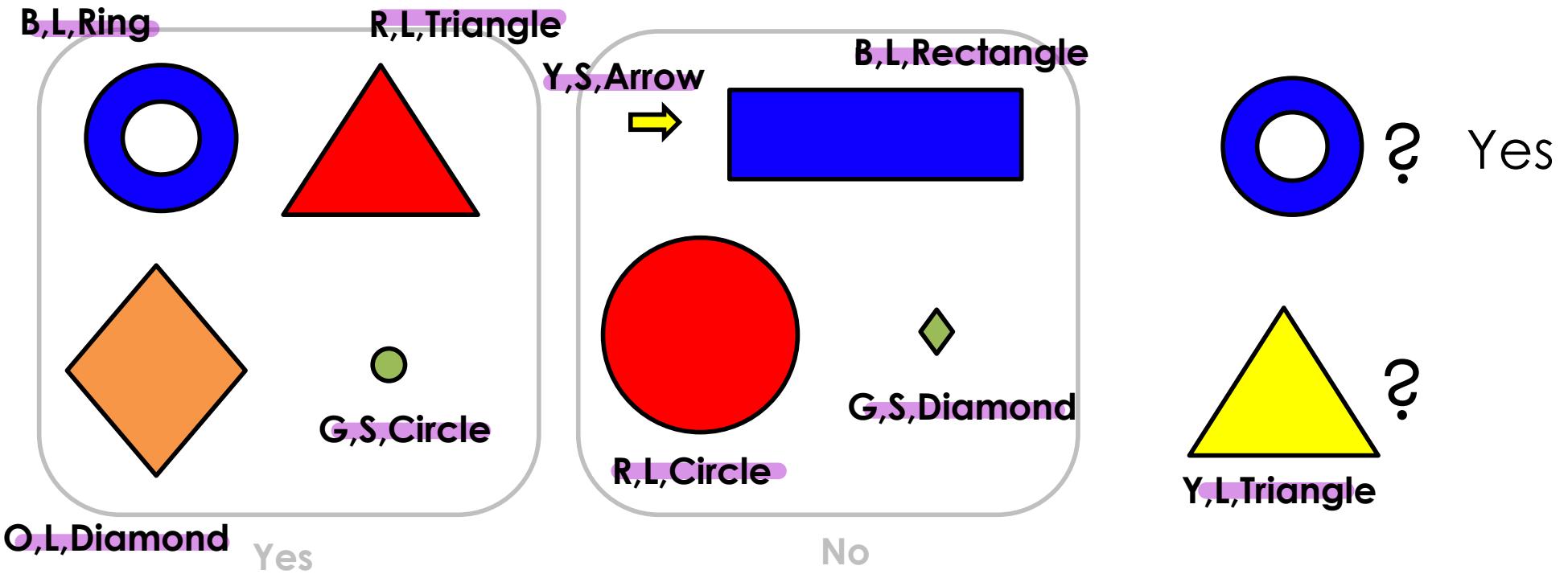
Shape, size, colour, etc



**Step 2: Sample Classification**  
Decide Label for a Sample

Decide if "yes" or "no"

## Walking Through A Toy Example: Token Classification



### Step 1: Feature Extraction

Color, Size, Shape → extract this feature

## Walking Through A Toy Example: Token Classification

### Feature Extraction

	<b>Color</b>	<b>Size</b>	<b>Shape</b>	<b>Label</b>
O	Blue	Large	Ring	Yes
▲	Red	Large	Triangle	Yes
◆	Orange	Large	Diamond	Yes
●	Green	Small	Circle	Yes
→	Yellow	Small	Arrow	No
■	Blue	Large	Rectangle	No
●	Red	Large	Circle	No
◆	Green	Small	Diamond	No
▲	Yellow	Large	Triangle	?

# Walking Through A Toy Example: Token Classification

## Feature Extraction

Color	Size	Shape	Label	
O	Blue	Large	Ring	Yes
▲	Red	Large	Triangle	Yes
◆	Orange	Large	Diamond	Yes
●	Green	Small	Circle	Yes
→	Yellow	Small	Arrow	No
■	Blue	Large	Rectangle	No
○	Red	Large	Circle	No
◆	Green	Small	Diamond	No

# Walking Through A Toy Example: Token Classification

## Feature Extraction



Color	Size	Shape	Label
Blue	Large	Ring	Yes
Red	Large	Triangle	Yes
Orange	Large	Diamond	Yes
Green	Small	Circle	Yes
Yellow	Small	Arrow	No
Blue	Large	Rectangle	No
Red	Large	Circle	No
Green	Small	Diamond	No

[compare]

## Similarity

Color	Size	Shape	Total
0	1	0	1
0	1	1	2
0	1	0	1
0	0	0	0
1	0	0	1
0	1	0	1
0	1	0	1
0	0	0	0

# Walking Through A Toy Example: Token Classification

Similarity

	Color	Size	Shape	Total
	0	1	0	1
	0	1	1	2
	0	1	0	1
	0	0	0	0
	1	0	0	1
	0	1	0	1
	0	1	0	1
	0	0	0	0

what if more than one maximum? ↗

- 1) Add more features [prevent this from occurring]
- 2) Choose either one

↗ most similar

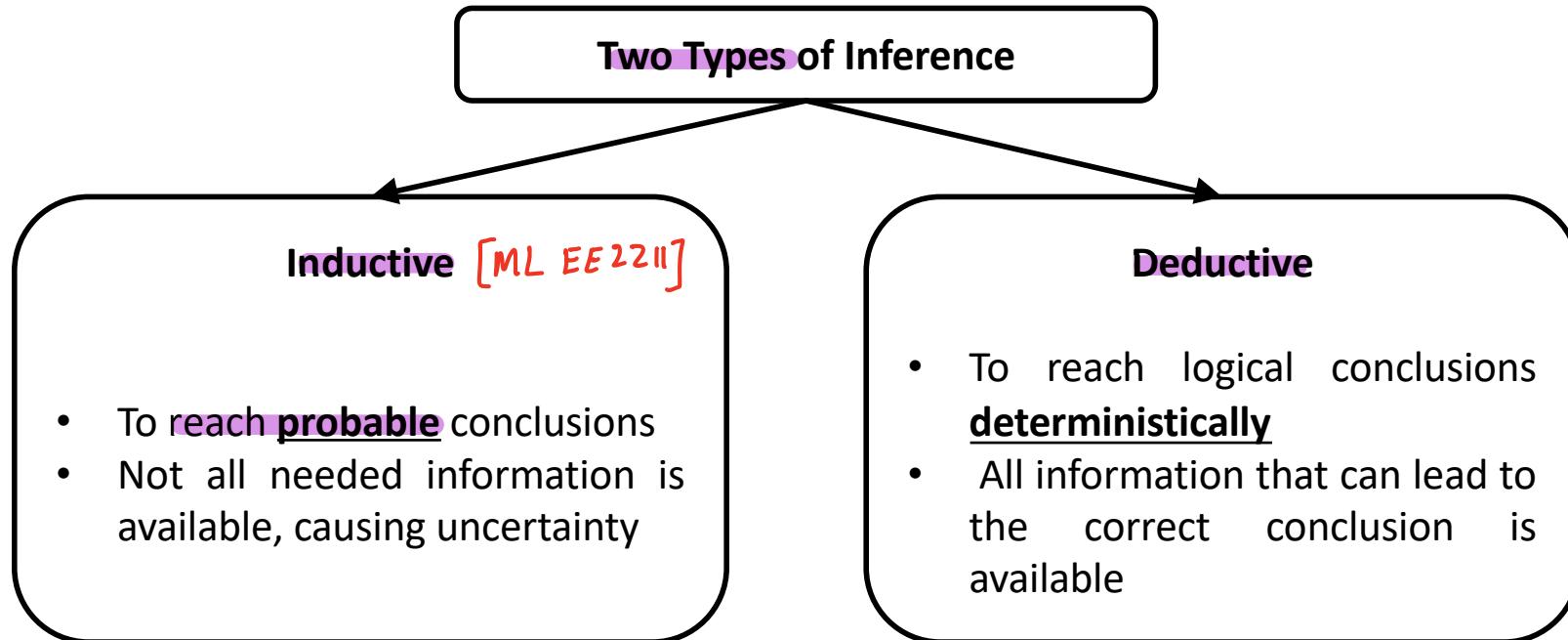
## Nearest Neighbor Classifier:

- 1) Find the “nearest neighbor” of a sample in the feature space
- 2) Assign the label of the nearest neighbor to the sample

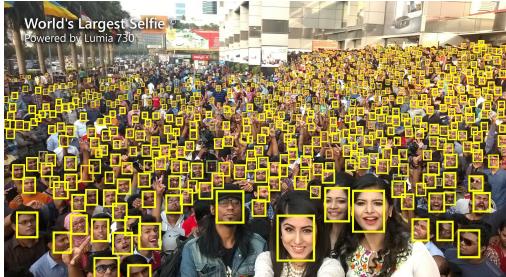
→ \* Don't put this new triangle data into the input dataset as it may do harm to dataset, if small dataset can consider

# Inductive vs. Deductive Reasoning

- Main Task of Machine Learning: to make inference



Probability and Statistics



Rule-based reasoning

NUS is in Singapore, Singapore is in Asia ->  
NUS is in Asia

# Inductive Reasoning

Note: humans use inductive reasoning all the time and not in a formal way like using probability/statistics.

B. C.

by Johnny hart



Ref: Gardner, Martin (March 1979). "[MATHEMATICAL GAMES: On the fabric of inductive logic, and some probability paradoxes](#)" (PDF). *Scientific American*. 234

# Summary by Quick Quiz

## Three Components in ML Definition

Task T, Performance P, Experience E

Both supervised and unsupervised learning involve feature extraction as the first step, second step for supervised learning can be classification or regression while for unsupervised learning can be clustering.

## Two Types of Supervised Learning

Classification, Regression

## Three Types of in ML

Supervised Learning  
Unsupervised Learning  
Reinforcement Learning

## One Type of Unsupervised Learning

Clustering

## Inductive and Deductive

Inductive: Probable  
Deductive: Rule-based

## Example of a Classifier Model

Nearest Neighbor Classifier

# Practice Question

## (Type of Question to Expect in Exams)

Which of the following statement is true?

- A. Nearest Neighbor Classifier is an example of unsupervised learning X → have input data, classification is supervised learning (pg 19)
- B. Nearest Neighbor Classifier is an example of deductive learning X inductive, we don't use deductive
- C. Nearest Neighbor Classifier is an example of feature selection X extraction separate feature extraction with classification (Step 1 separate with Step 2 pg 27)
- D. None of the above is correct. ✓ ↓ diff feature extraction likely lead to diff classification

