lecture 3

Dong-Geol Choi
Hanbat Nat'l Univ.

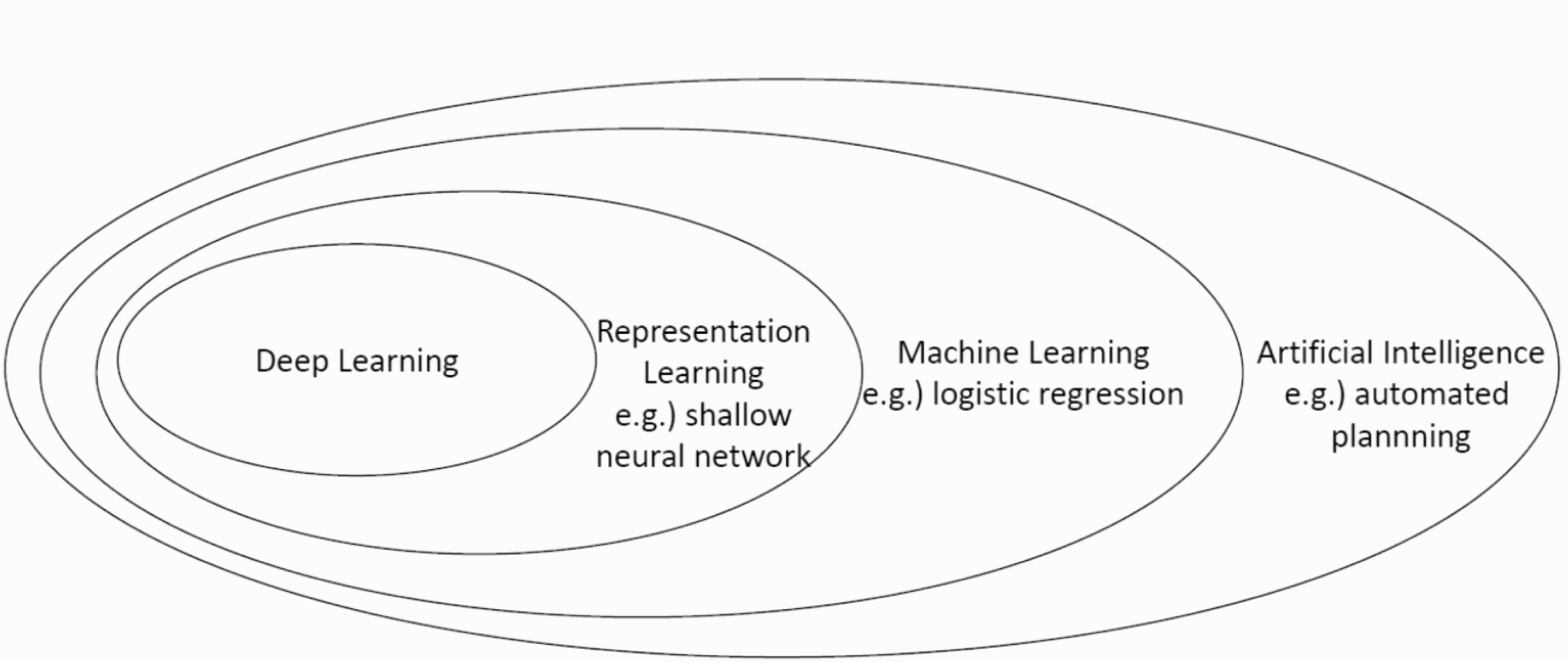
REMIND

What is Machine Learning?

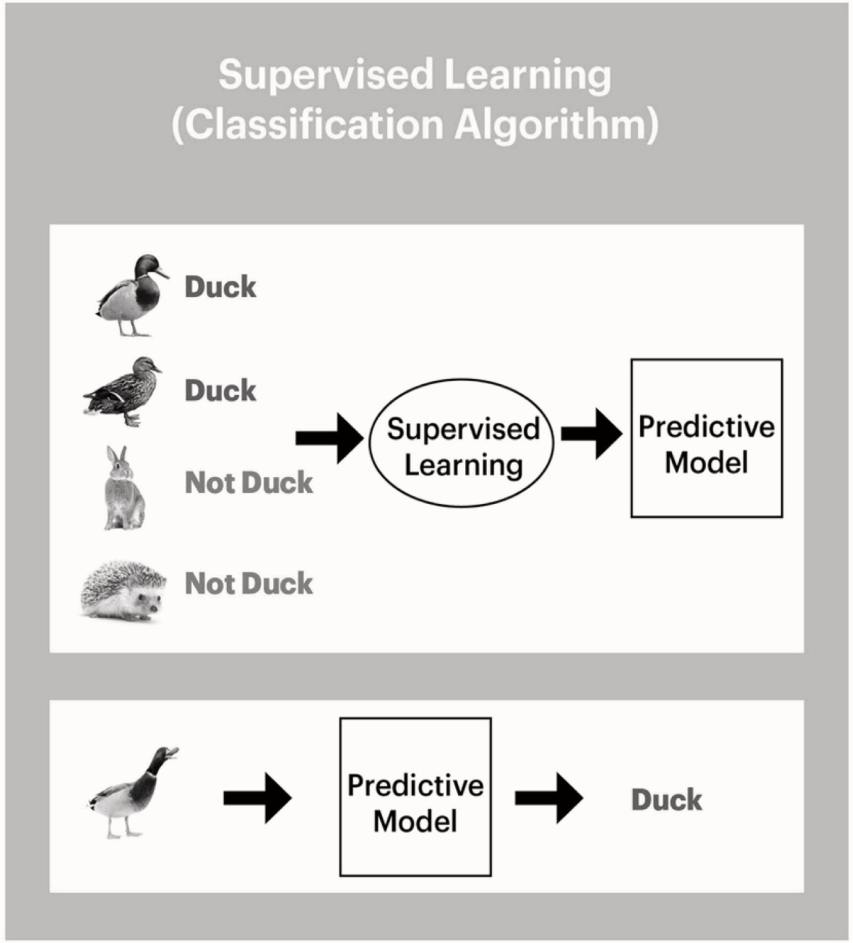
"A Field of study that gives computer the ability to learn without being explicitly programmed"

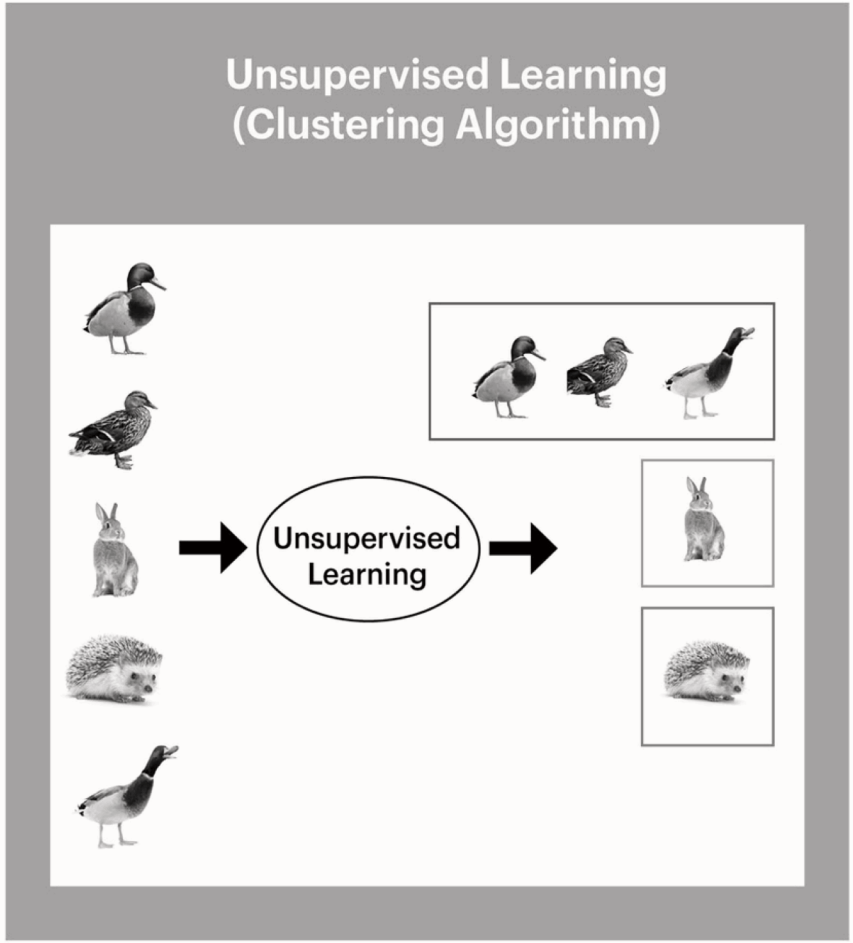
– Arthur Samuel, 1959

Deep Learning, Machine Learning, Artificial Intelligence



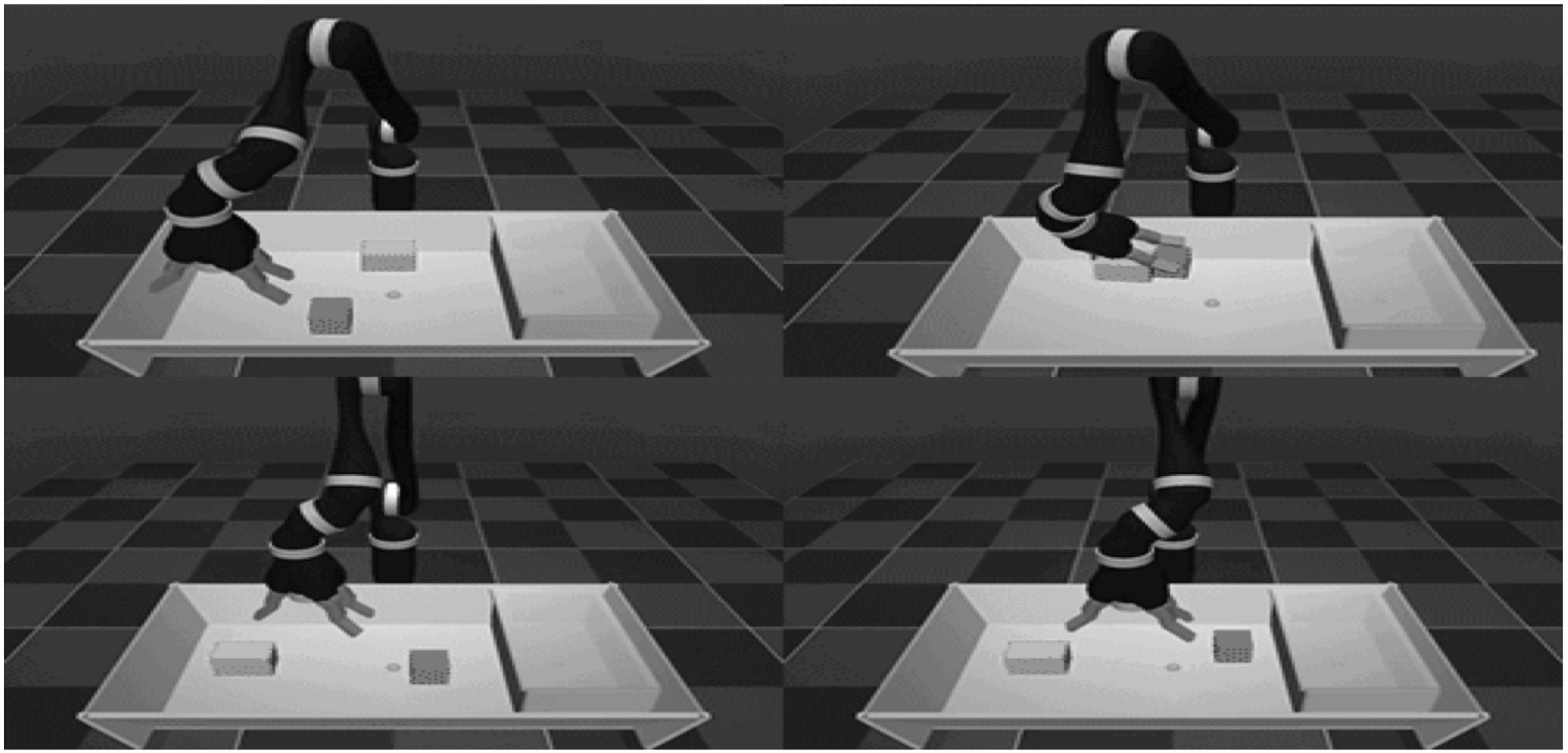
Categories of ML Problems



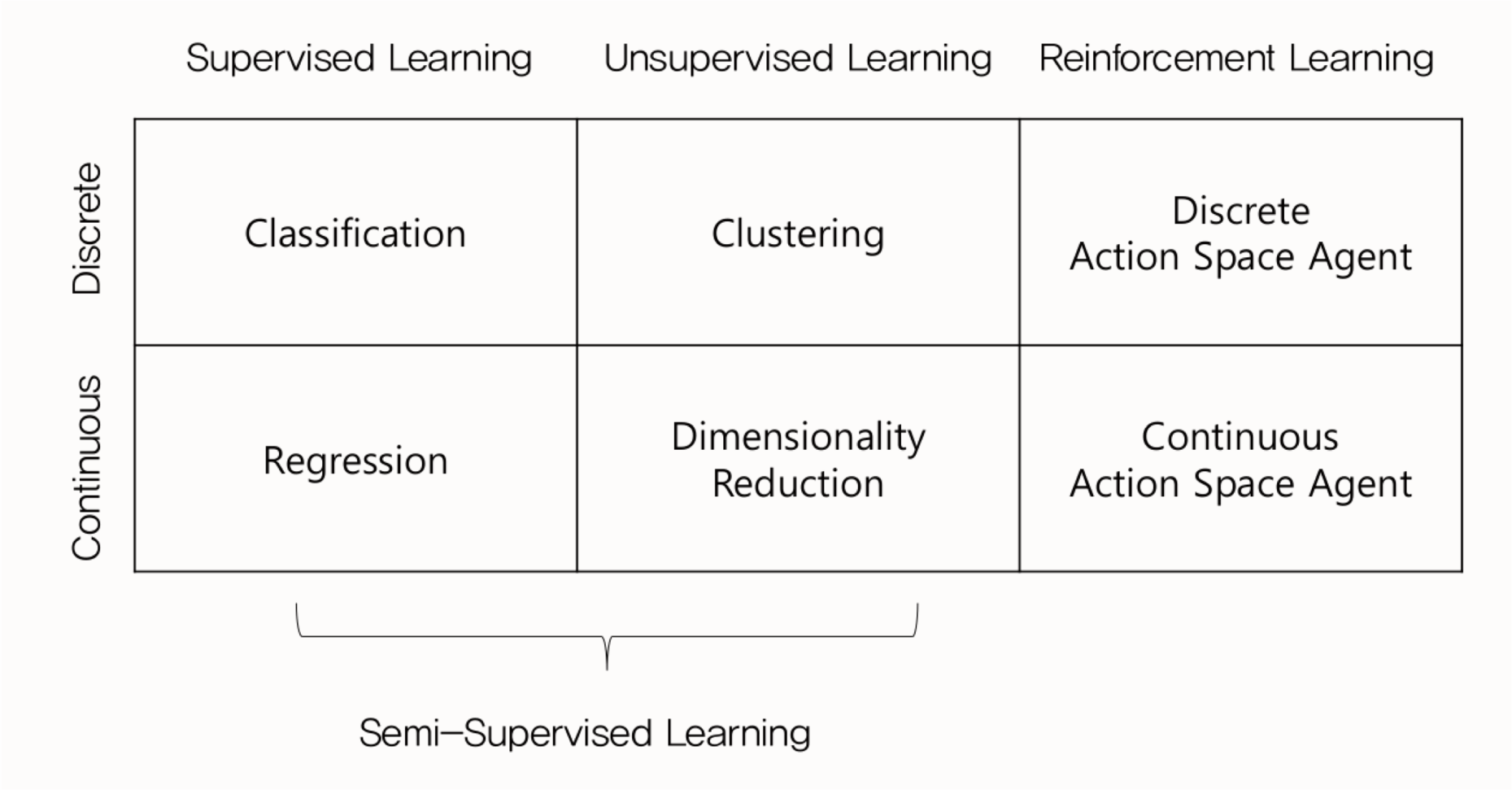


Categories of ML Problems

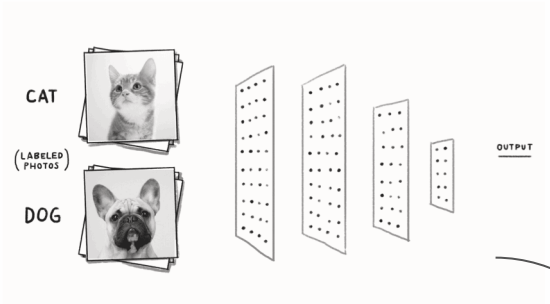
Reinforcement Learning

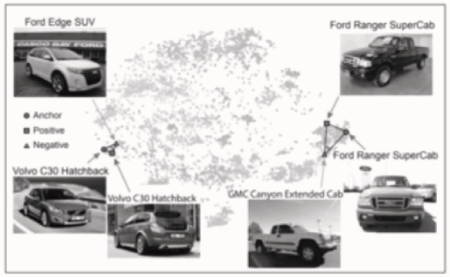


Categories of ML Problems



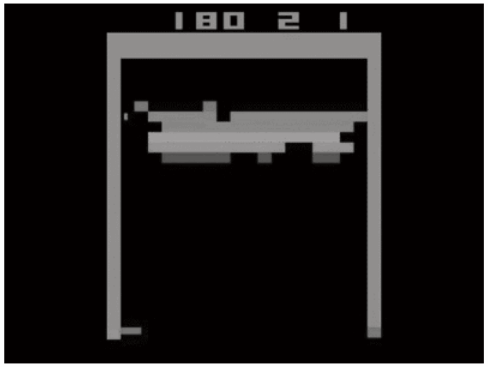
Categories of ML Problems

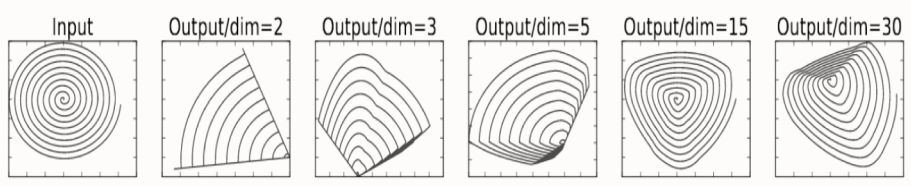




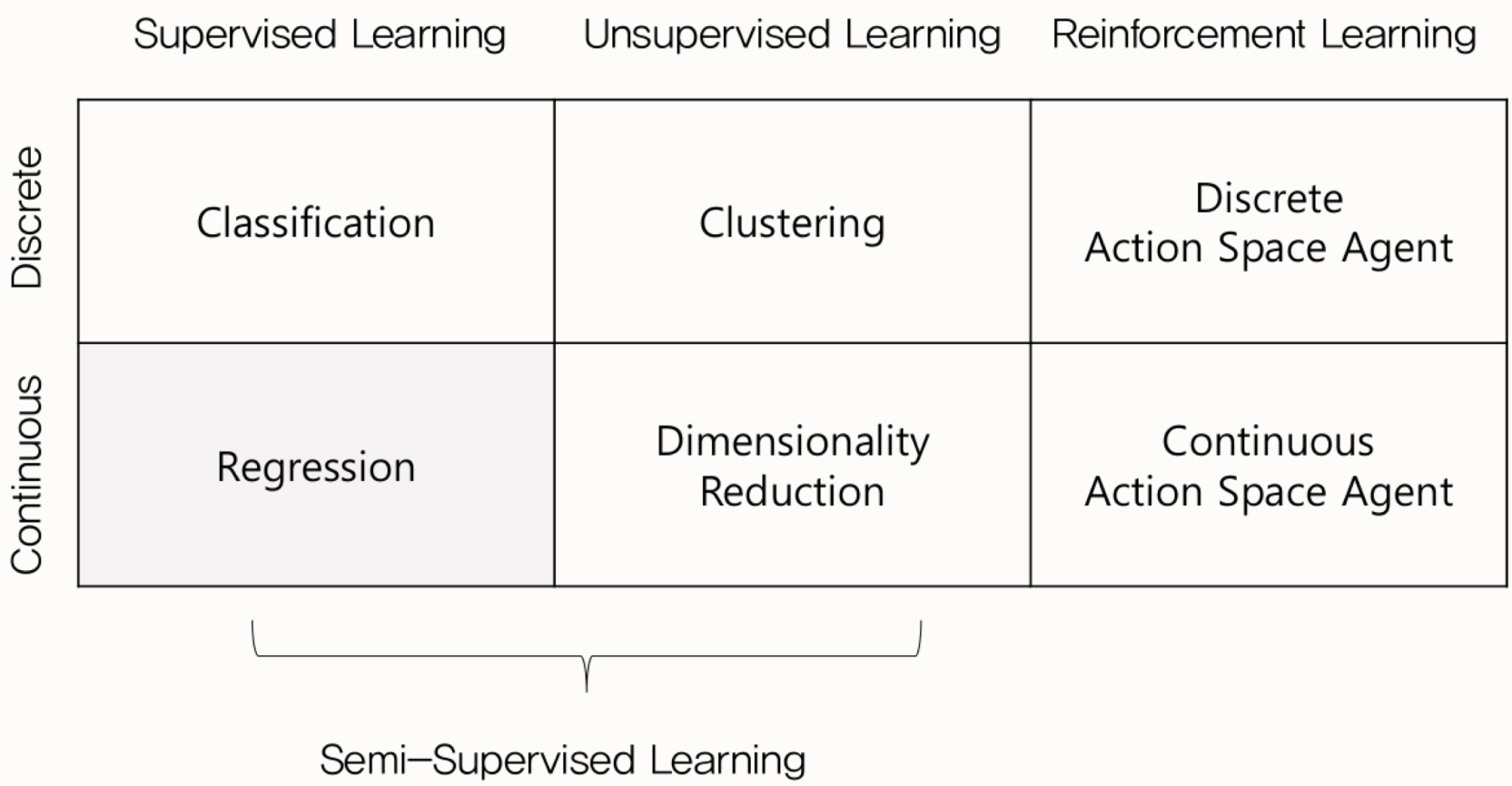
	Supervised Learning	Unsupervised Learning	Reinforcement Learning
Discrete	Classification	Clustering	Discrete Action Space Agent
Continuous	Regression	Dimensionality Reduction	Continuous Action Space Agent

Semi–Supervised Learning





Categories of ML Problems

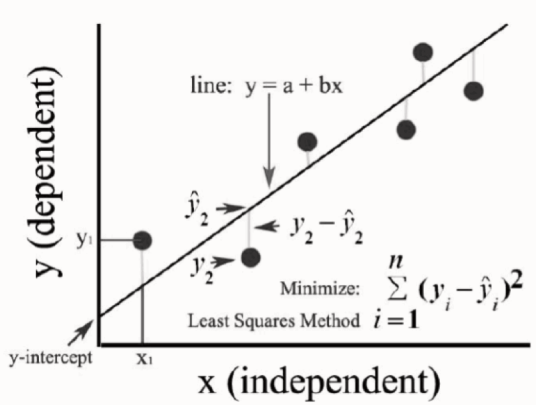


Regression Problem

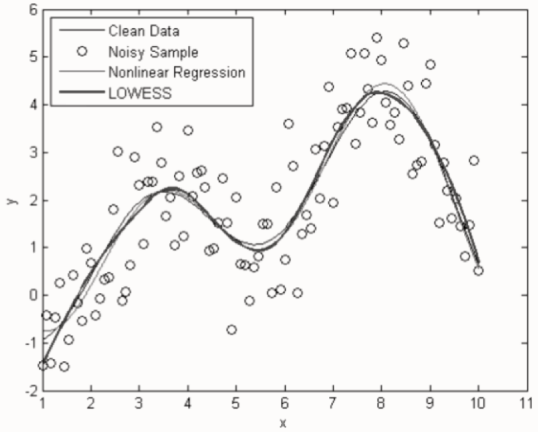
Price Prediction Based on Gi-Young Style Chart Analysis

Regression Problem

delivery distance	delivery time
100m	20min
150m	24min
300m	36min
400m	47min
130m	22min
240m	32min
350m	47min
200m	42min
100m	21min
110m	21min

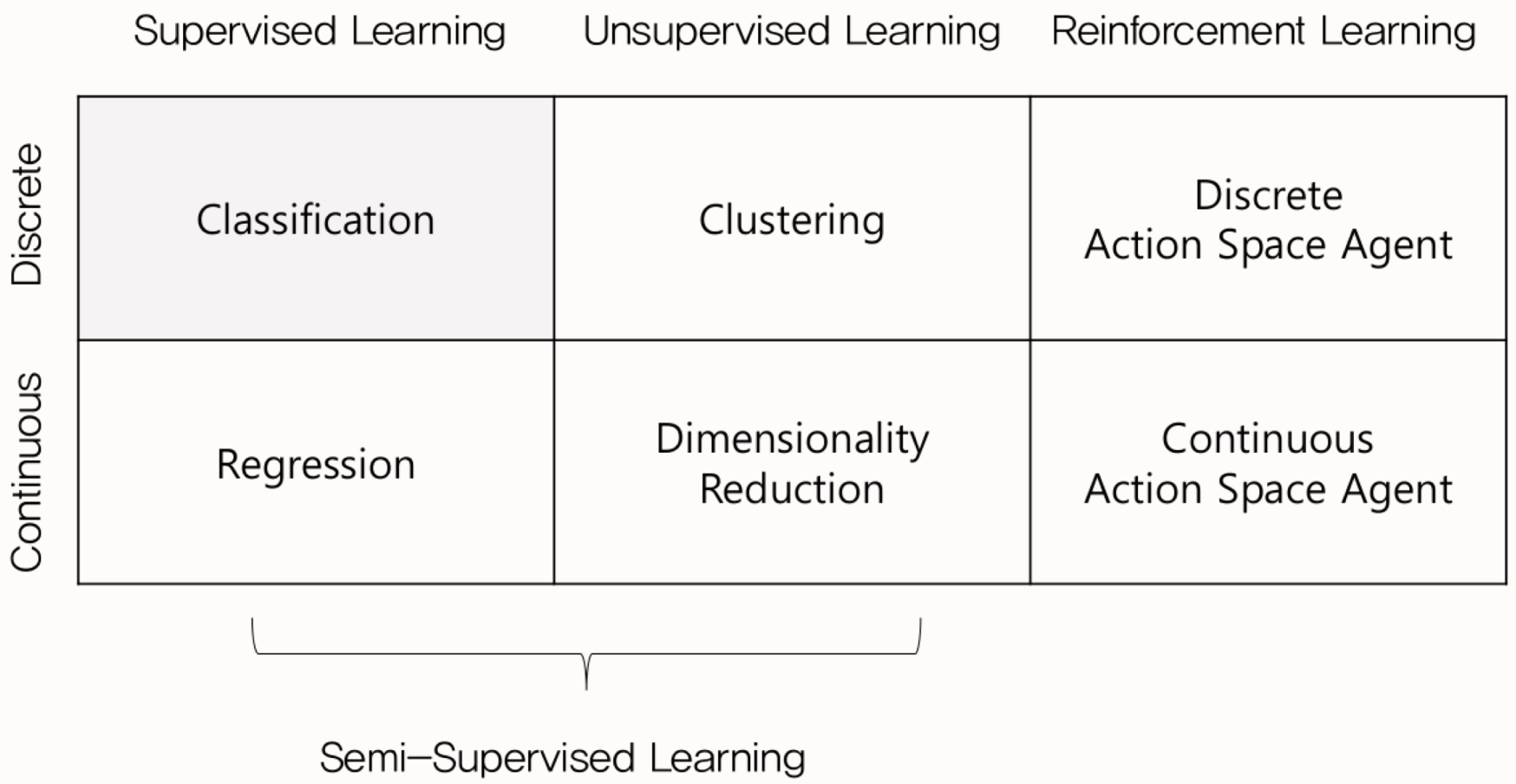


Linear regression



Nonlinear regression

Categories of ML Problems

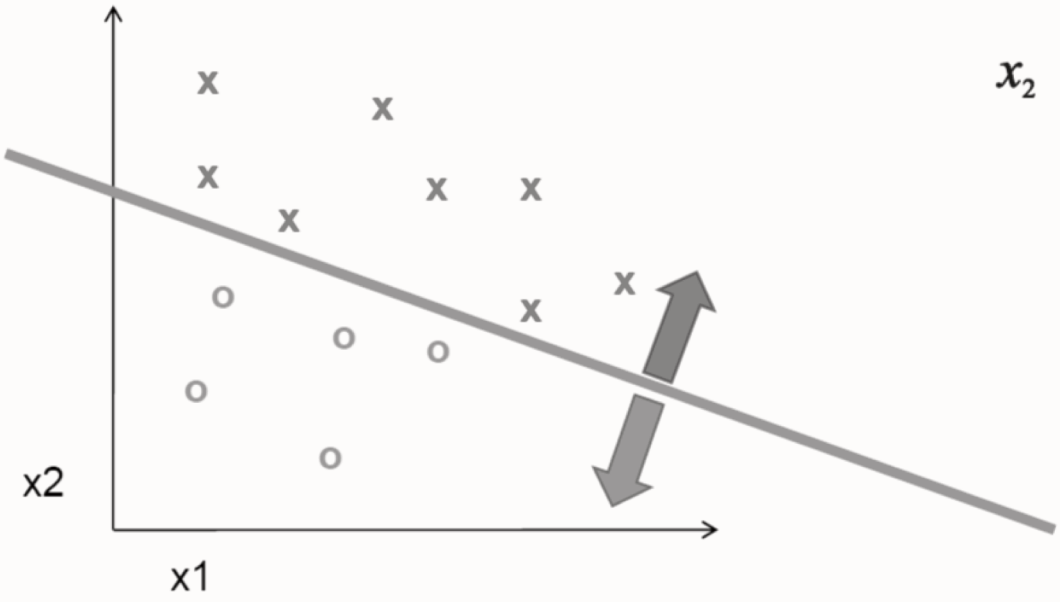


Classification Problem

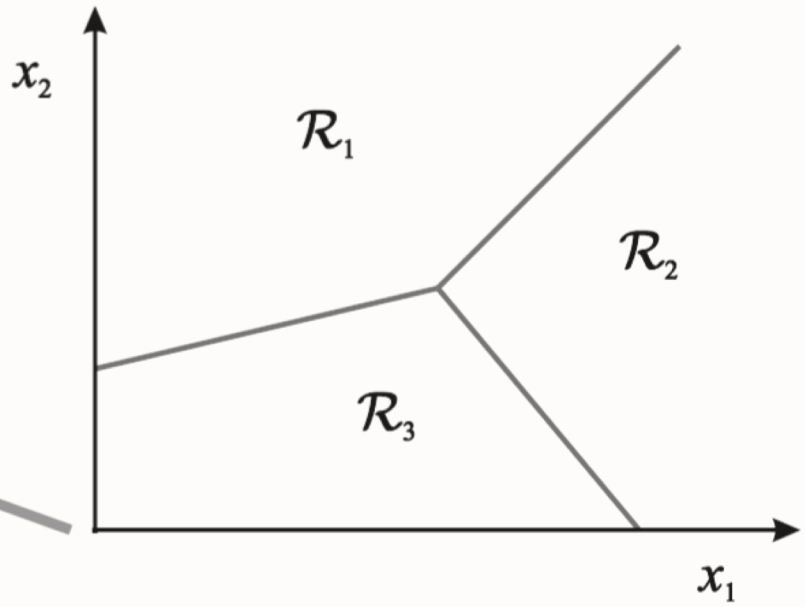
Chihuahua or Muffin?

Classification Problem

Identifying which of a set of categories a new instance belongs

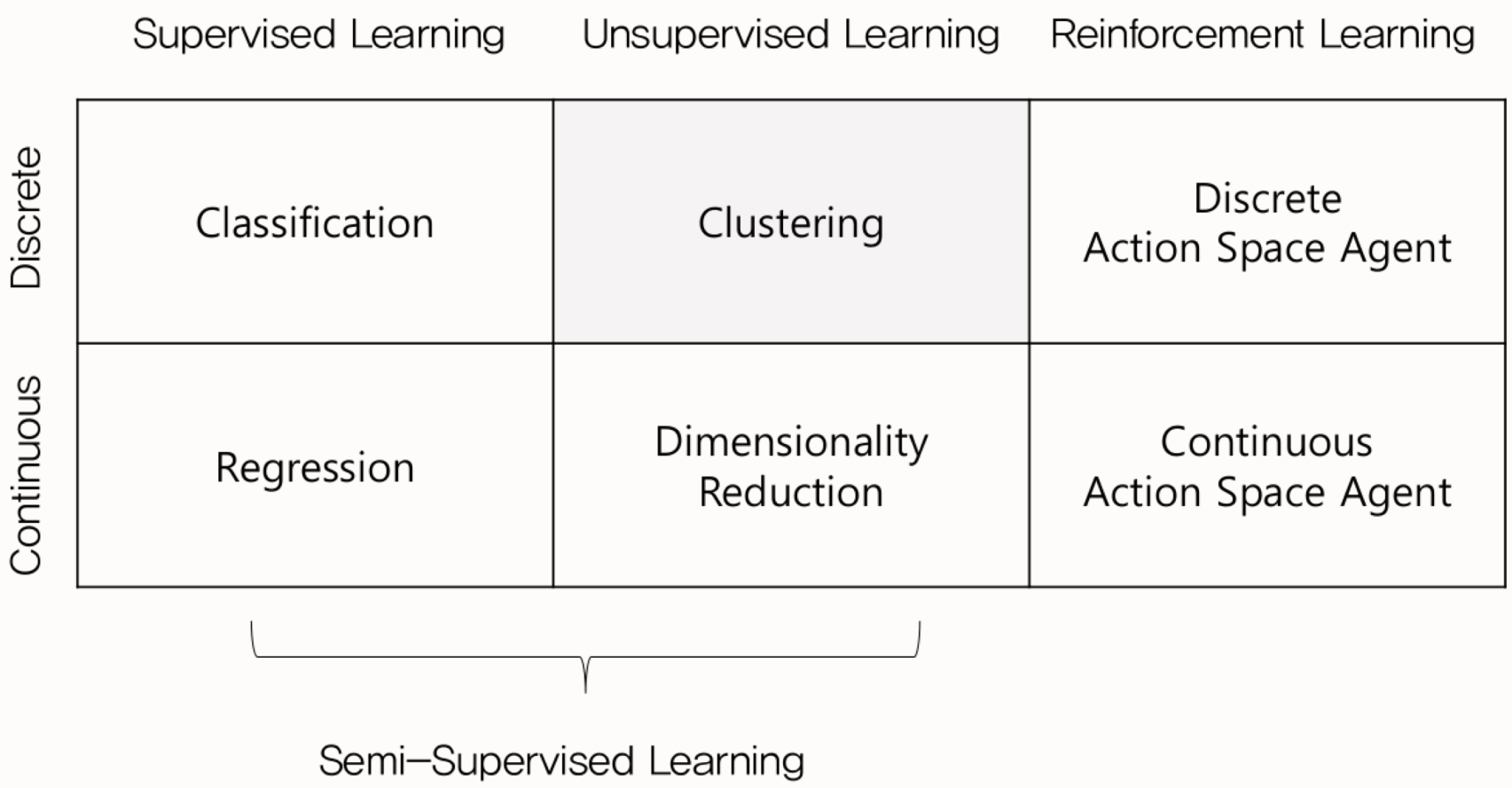


Binary Classification

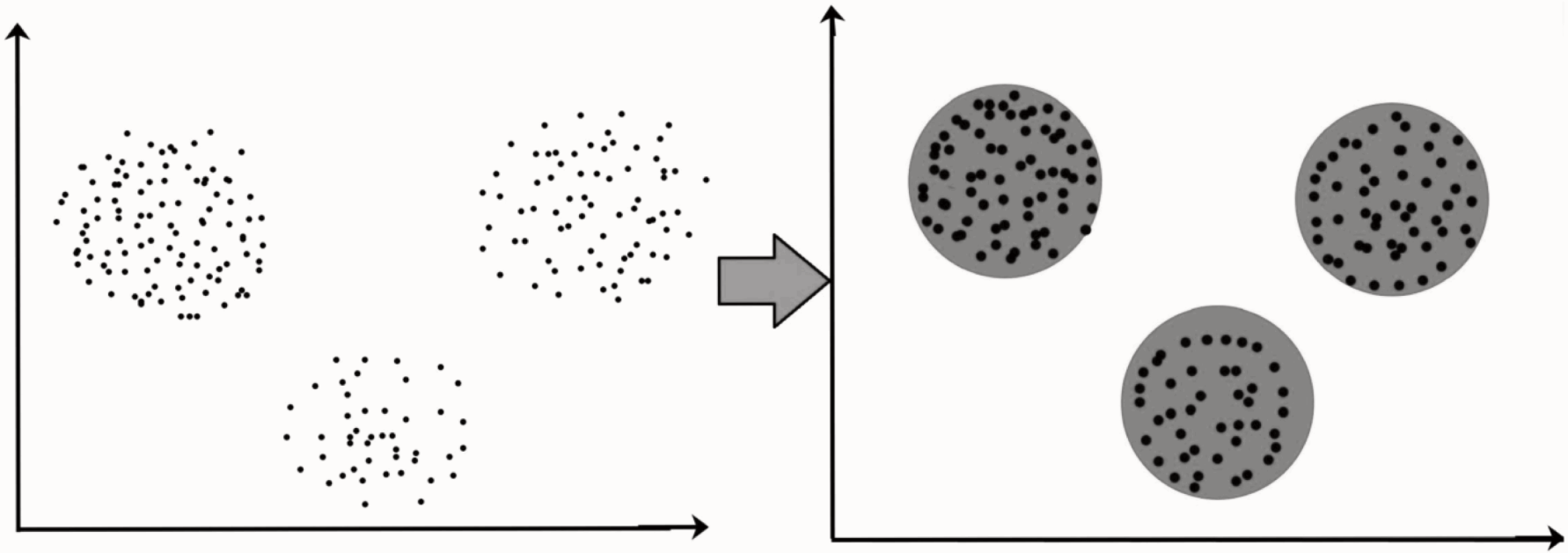


Multi-class Classification

Categories of ML Problems



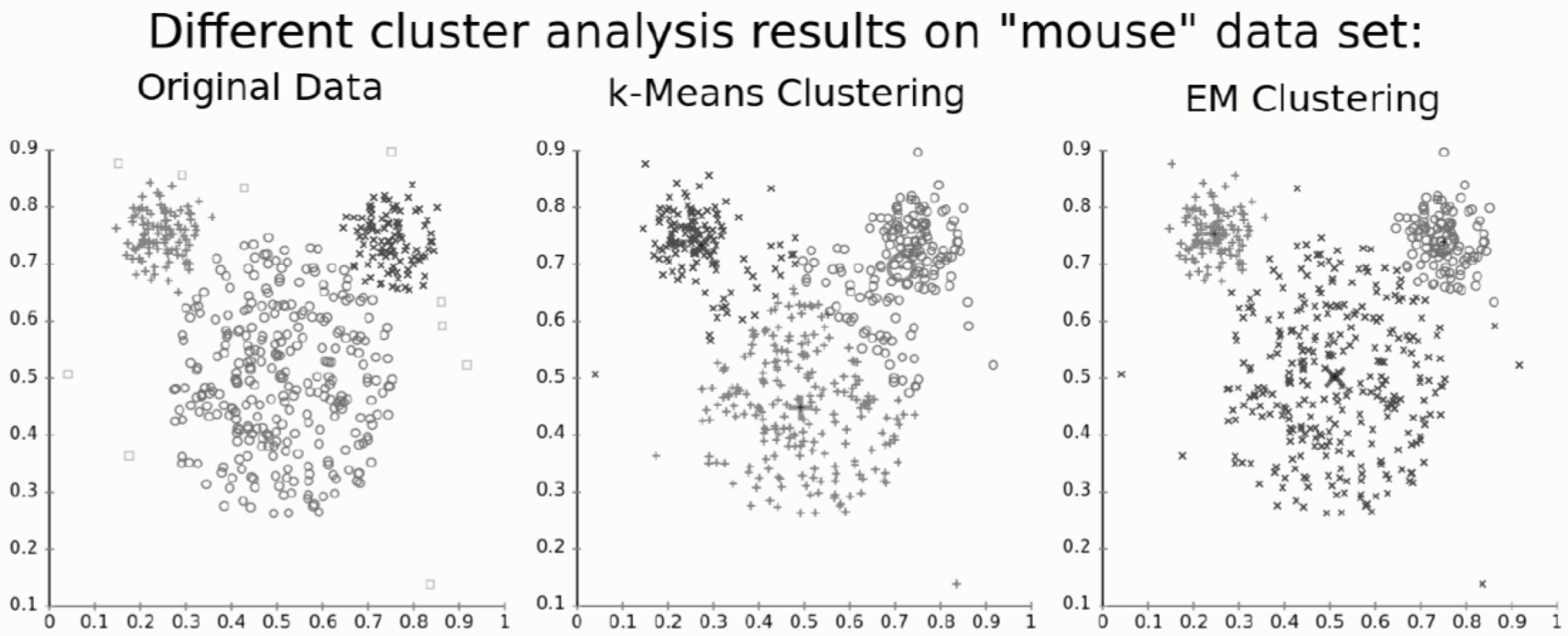
Clustering Problem



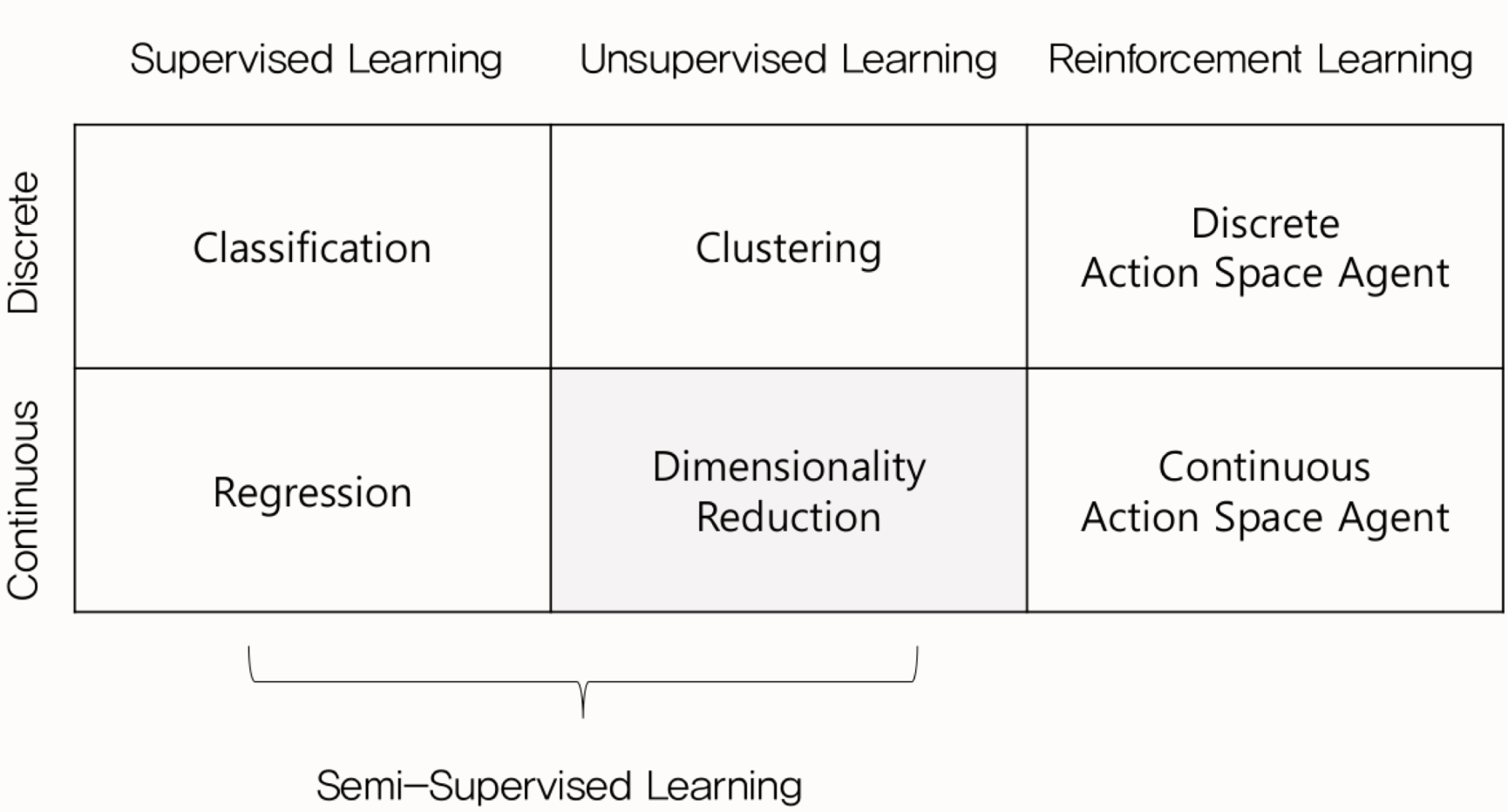
Grouping smilar samples into K groups

Clustering Problem

Automatic grouping of instances, such that the instances that belong to the same clusters are more similar to each other than to those in the other groups



Categories of ML Problems



Categories of ML Problems

Man or Woman?

- Short hair,
- Great physique.
- Two eyes.
- Two arms.
- One mouth.
- Two ears.
- ...

- Long hair,
- Small physique.
- Two eyes.
- Two arms.
- One mouth.
- Two ears.
- ...

Categories of ML Problems

Man or Woman?

- Short hair,
- Great physique.
- Two eyes.
- Two arms.
- One mouth.
- Two ears.
- ...

- Long hair,
- Small physique.
- Two eyes.
- Two arms.
- One mouth.
- Two ears.
- ...

Categories of ML Problems

Man or Woman?

- Short hair,
- Great physique.
- Two eyes.
- Two arms.
- One mou
- Two ears.
- ...

Too many dimensions

- Long hair,
- Small physique.
- Two eyes.
- Two arms.
- One mouth.
- Two ears.
- ...

Categories of ML Problems

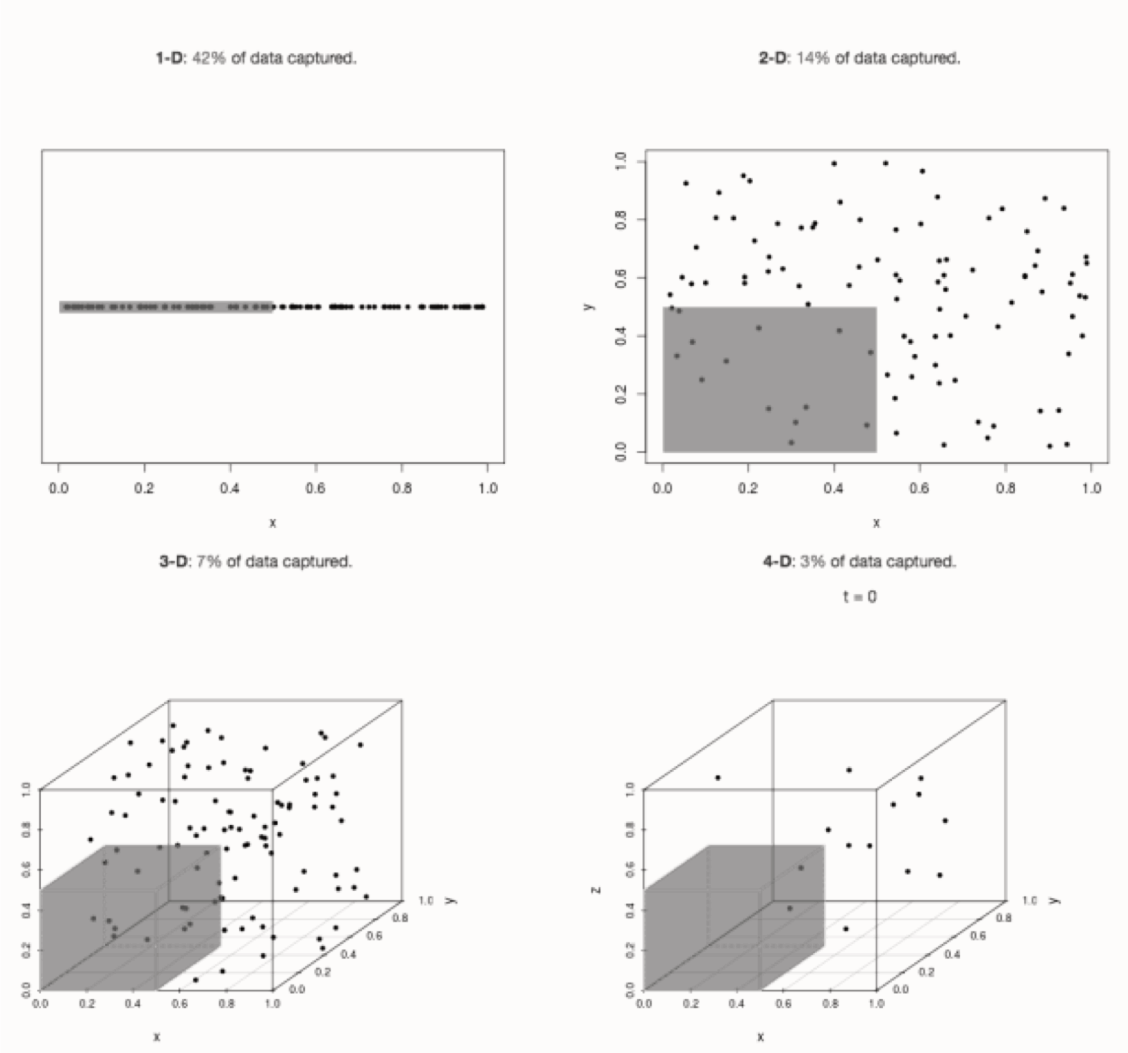
Man or Woman?

- Short hair.
- Great physique.

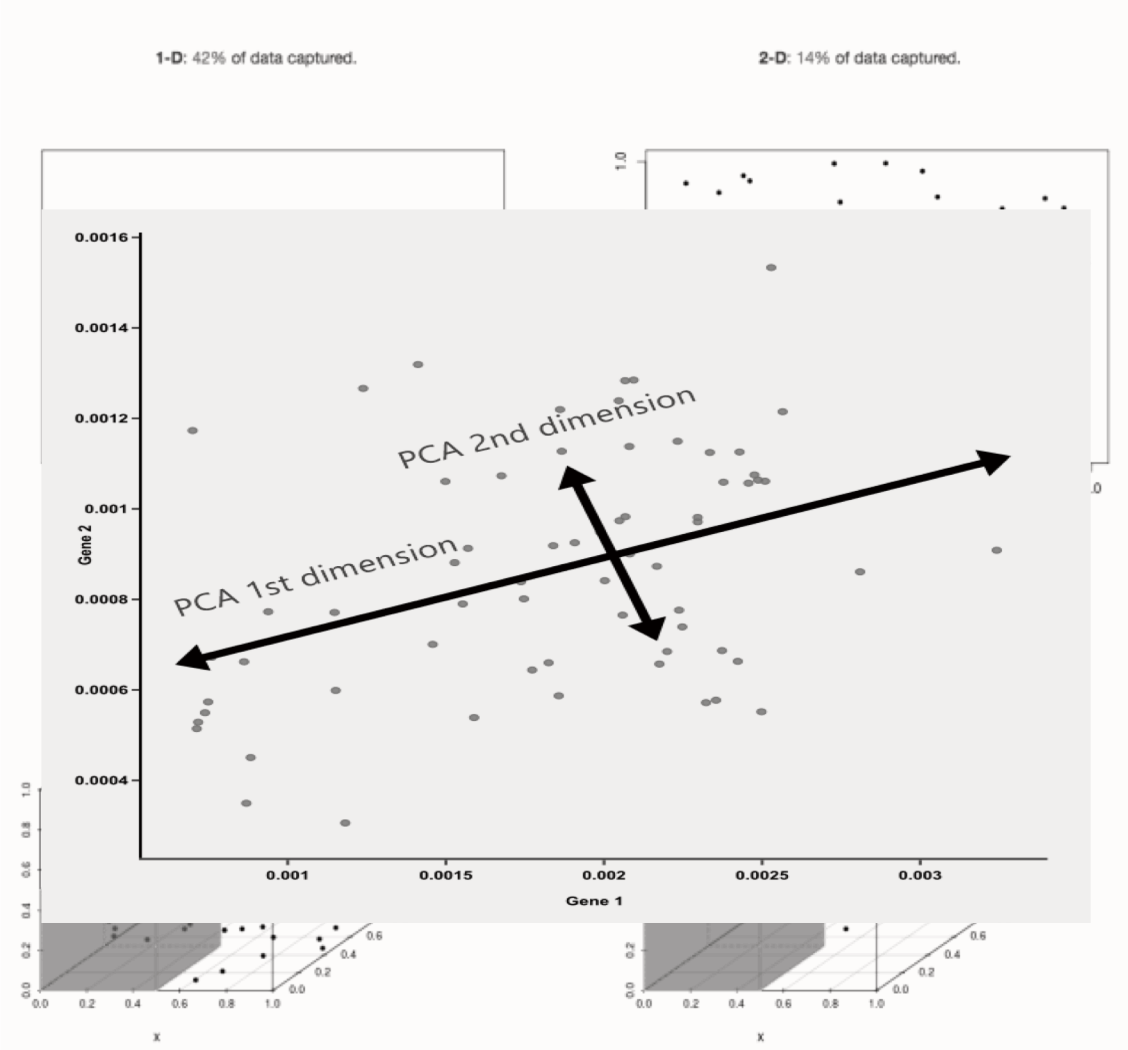
- Long hair.
- Small physique.

Good!

Dimensionality Reduction Problem

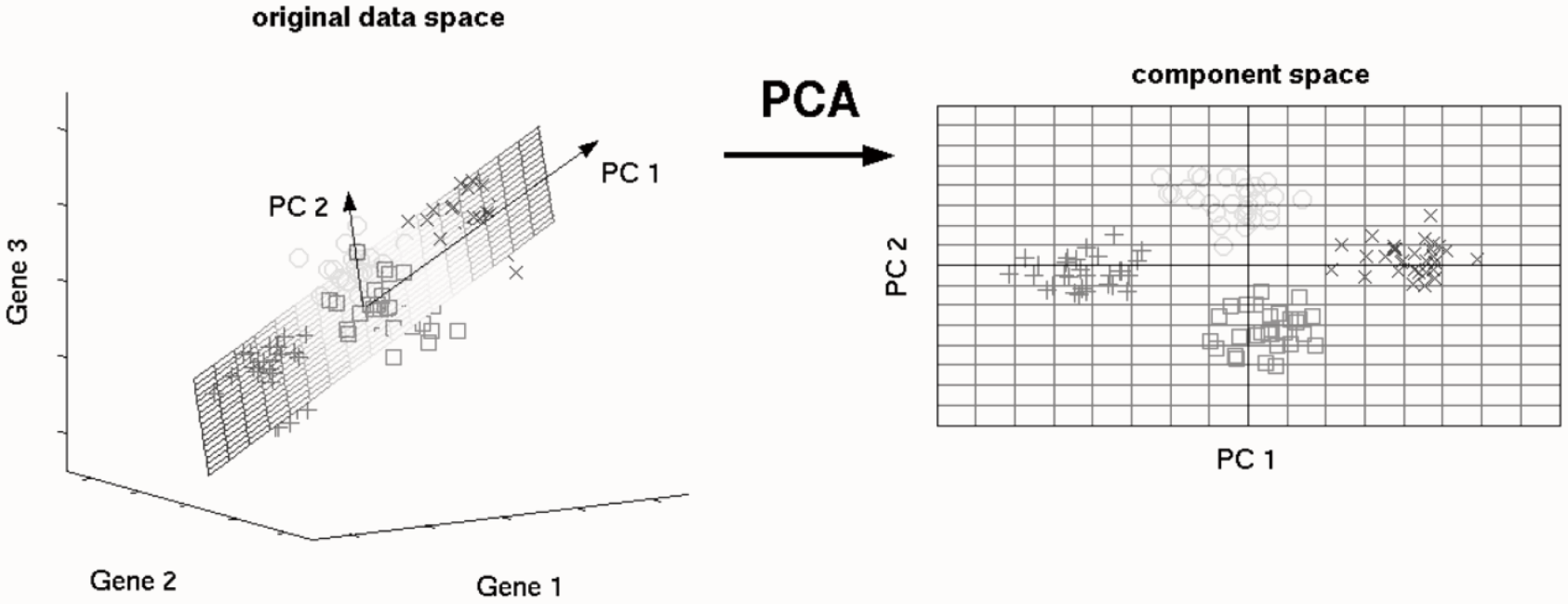


Dimensionality Reduction Problem



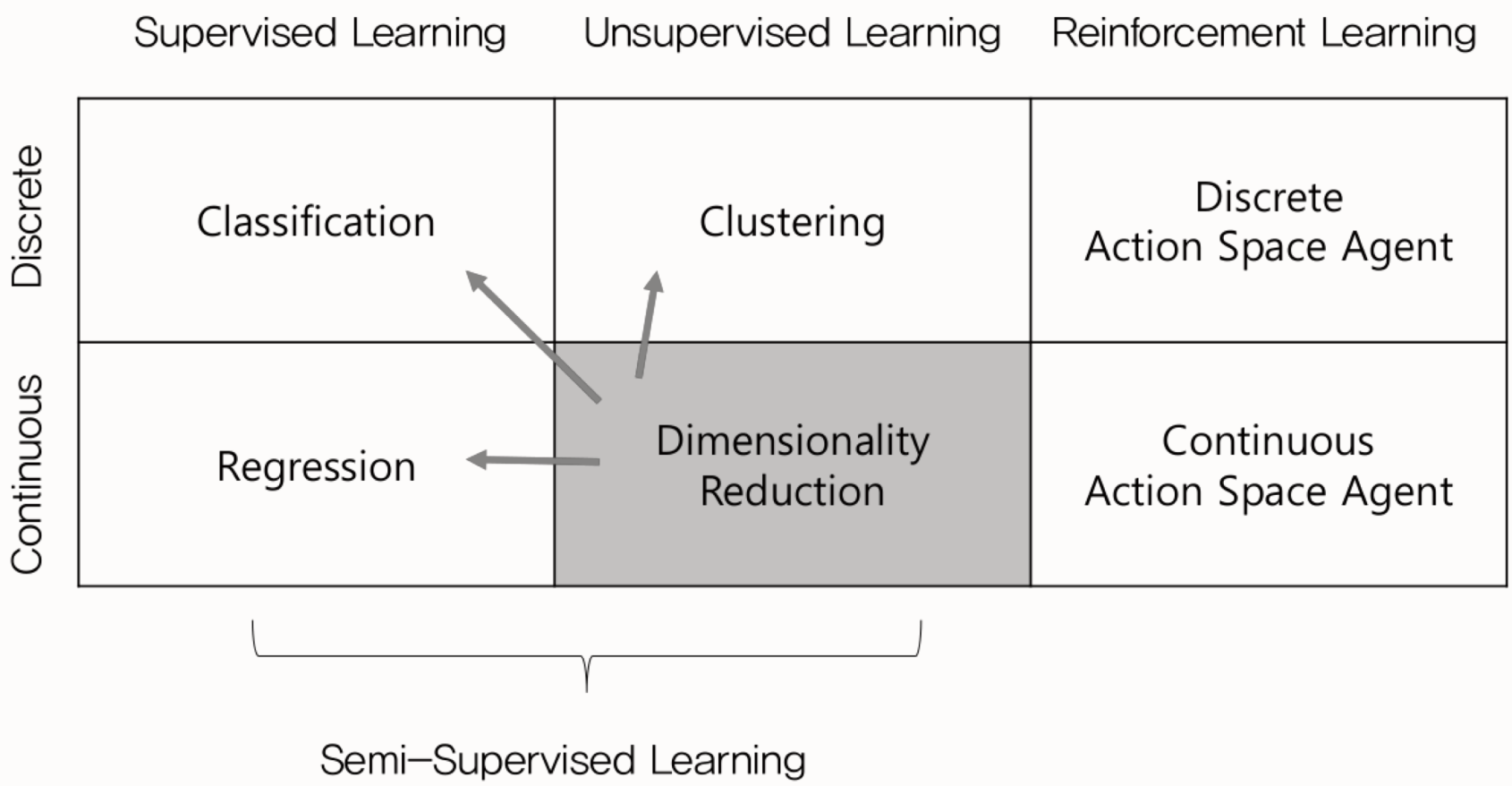
Dimensionality Reduction Problem

Reduce the dimension of input data, to avoid the effect of the curse of dimensionality

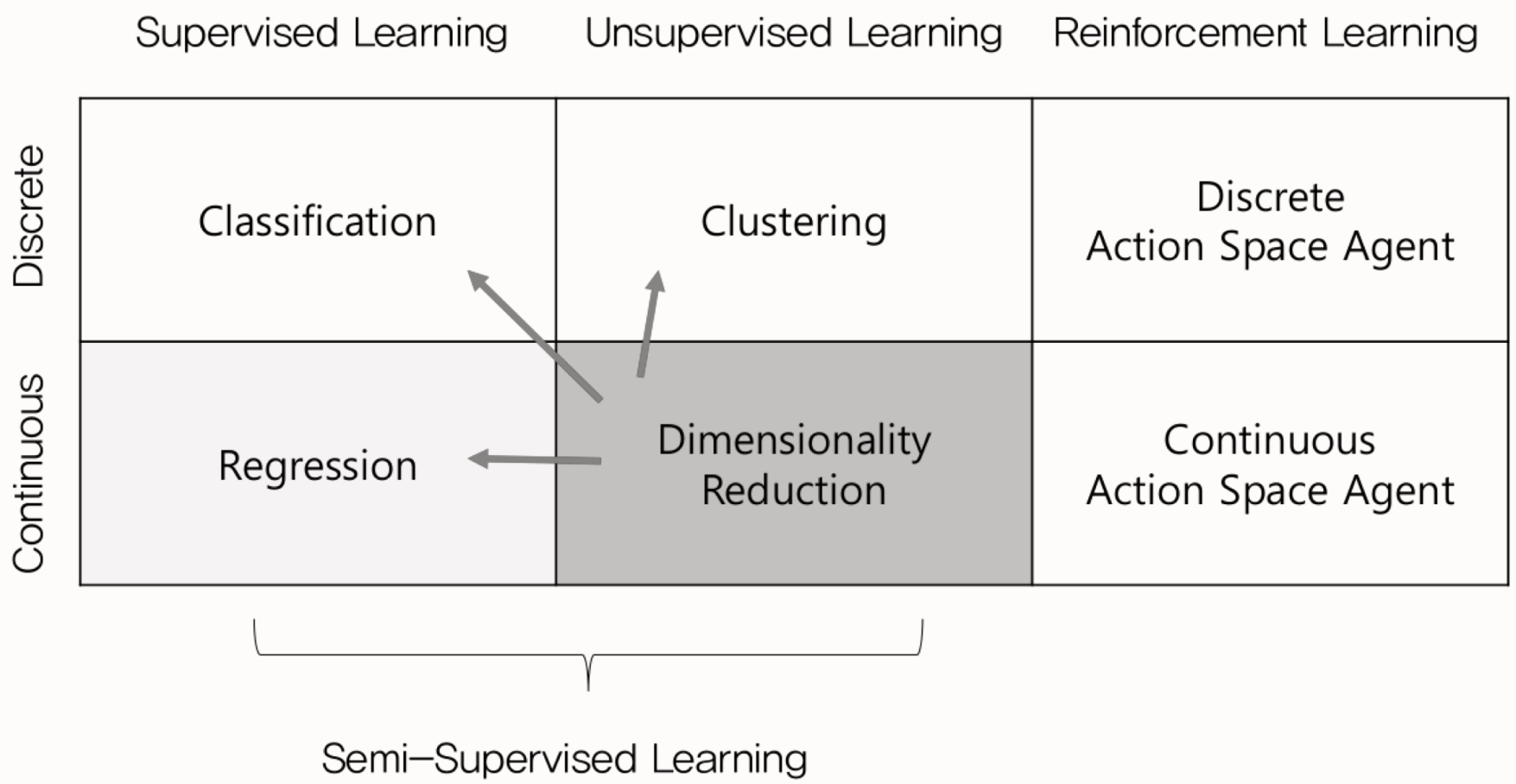


Reduce unnecessary representation axis

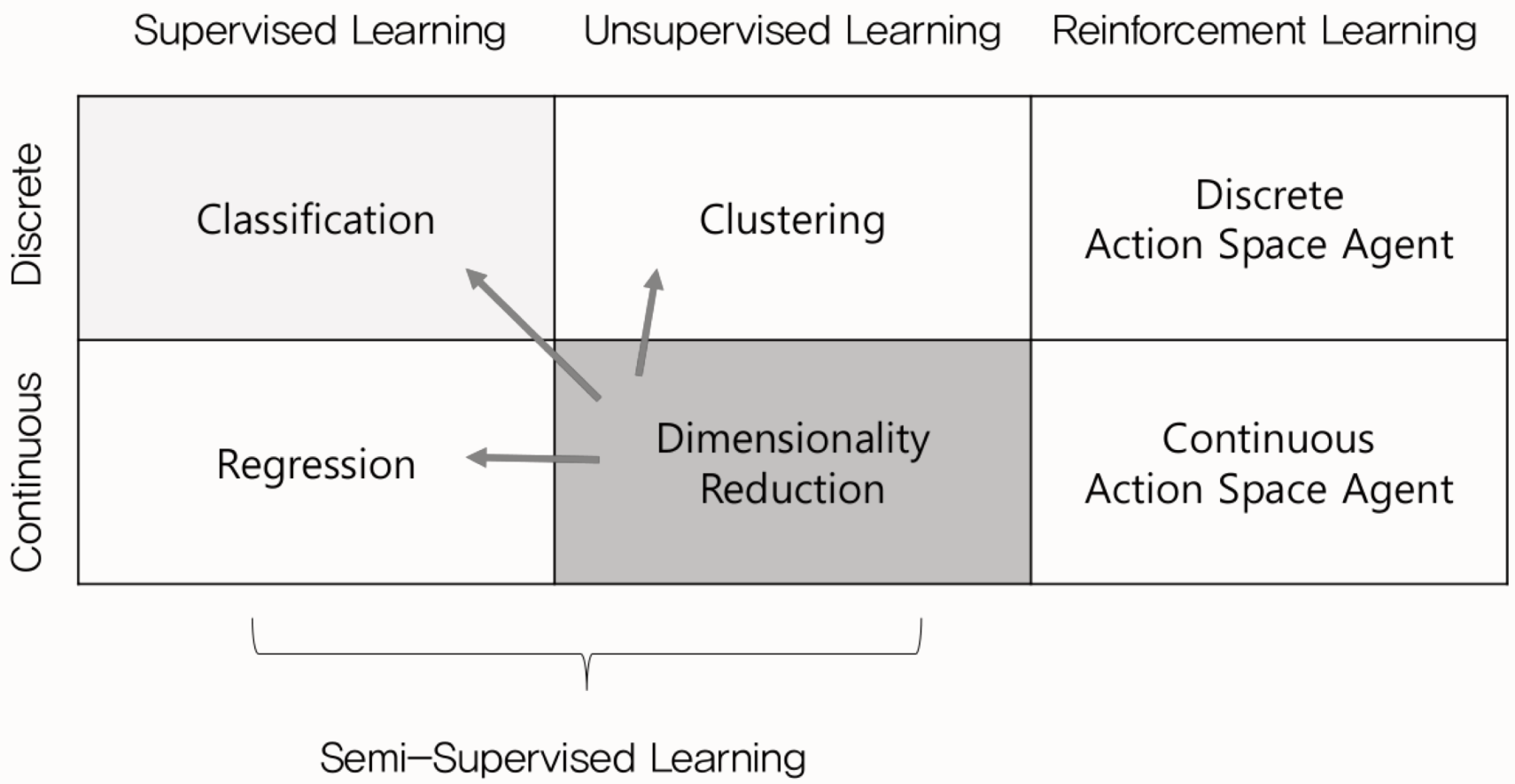
Categories of ML Problems



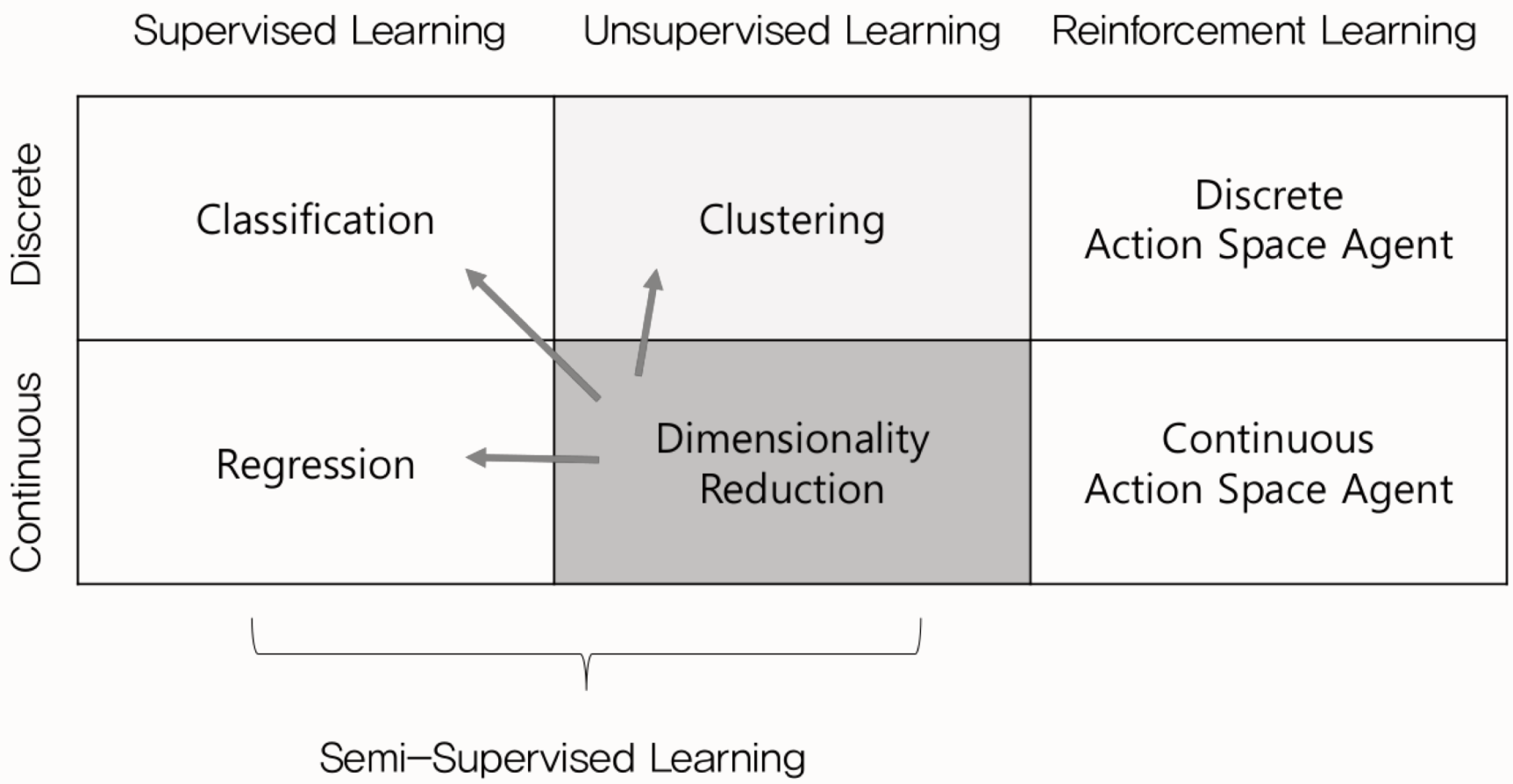
Categories of ML Problems



Categories of ML Problems



Categories of ML Problems

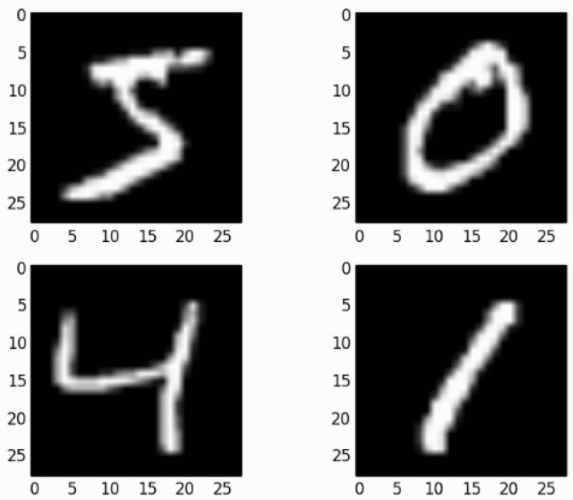


Feature & Data Representation

Case 1

X		Y
x1: quiz 1 score	ML →	y1: final exam score
x2: quiz 2 score		y2: grade
x3: study hour		

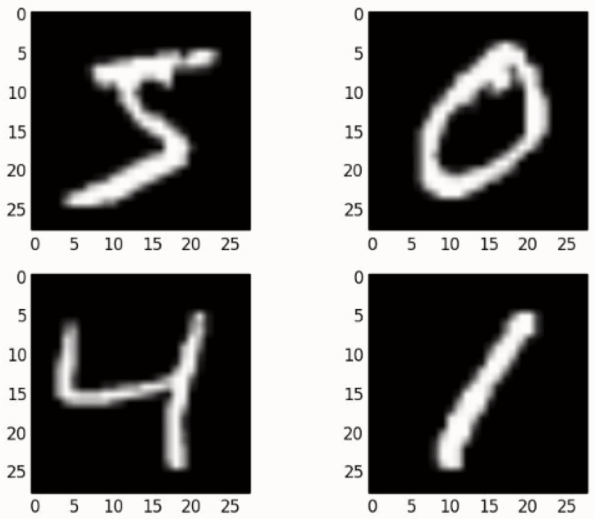
Feature & Data Representation



Case 2

X		Y
x1: first pixel value	→	y1: digit
x2: second pixel value		
x3: third pixel value		
...		
x784: 784^{th} pixel value		

Feature & Data Representation



Case 2

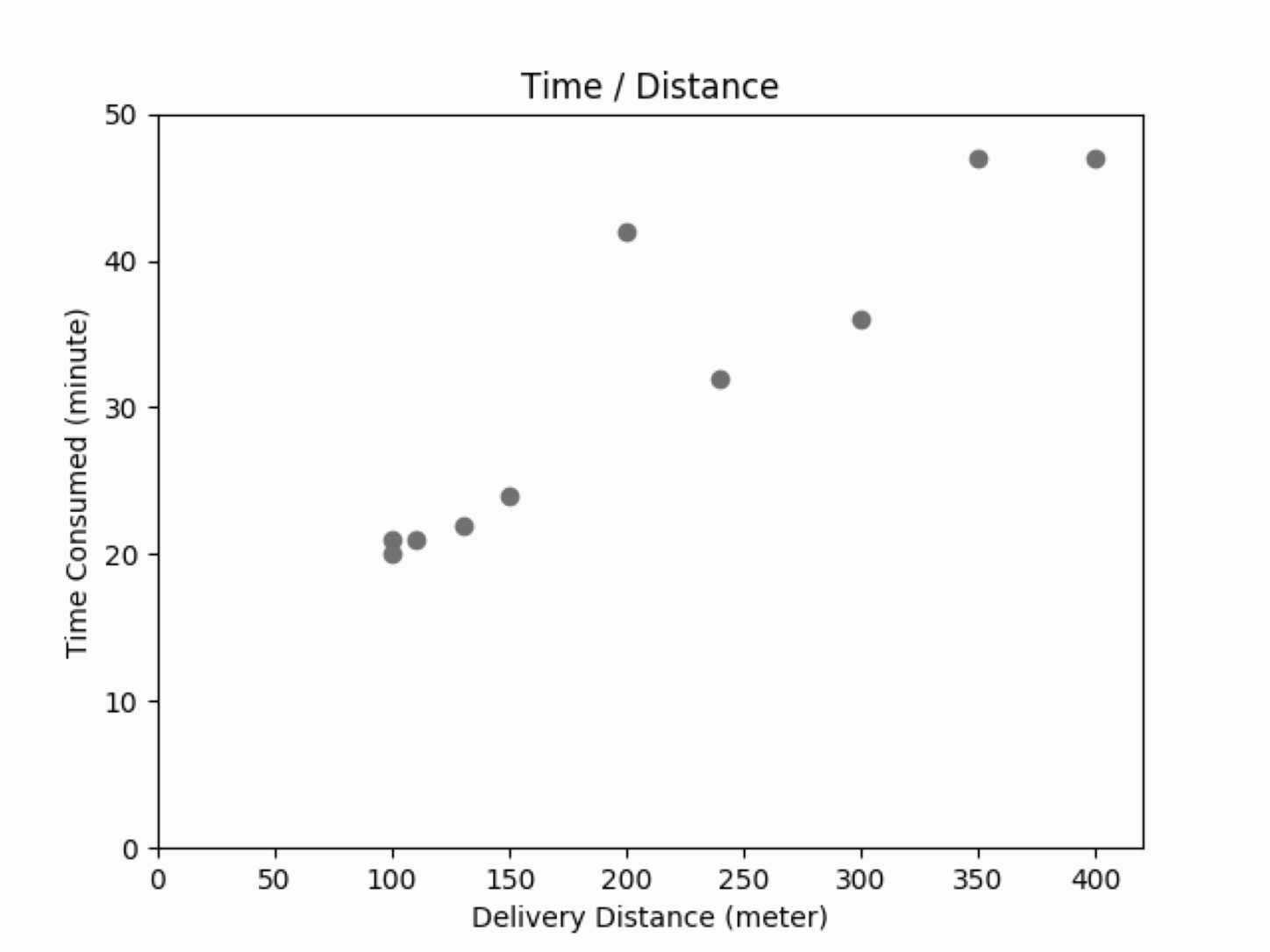
X		h		Y
	DL		ML	
x1: first pixel value	→	h1: feature1	→	y1: digit
x2: second pixel value		h2: feature2		
x3: third pixel value		h3: feature3		
...				
x784: 784^{th} pixel value				

Linear Regression

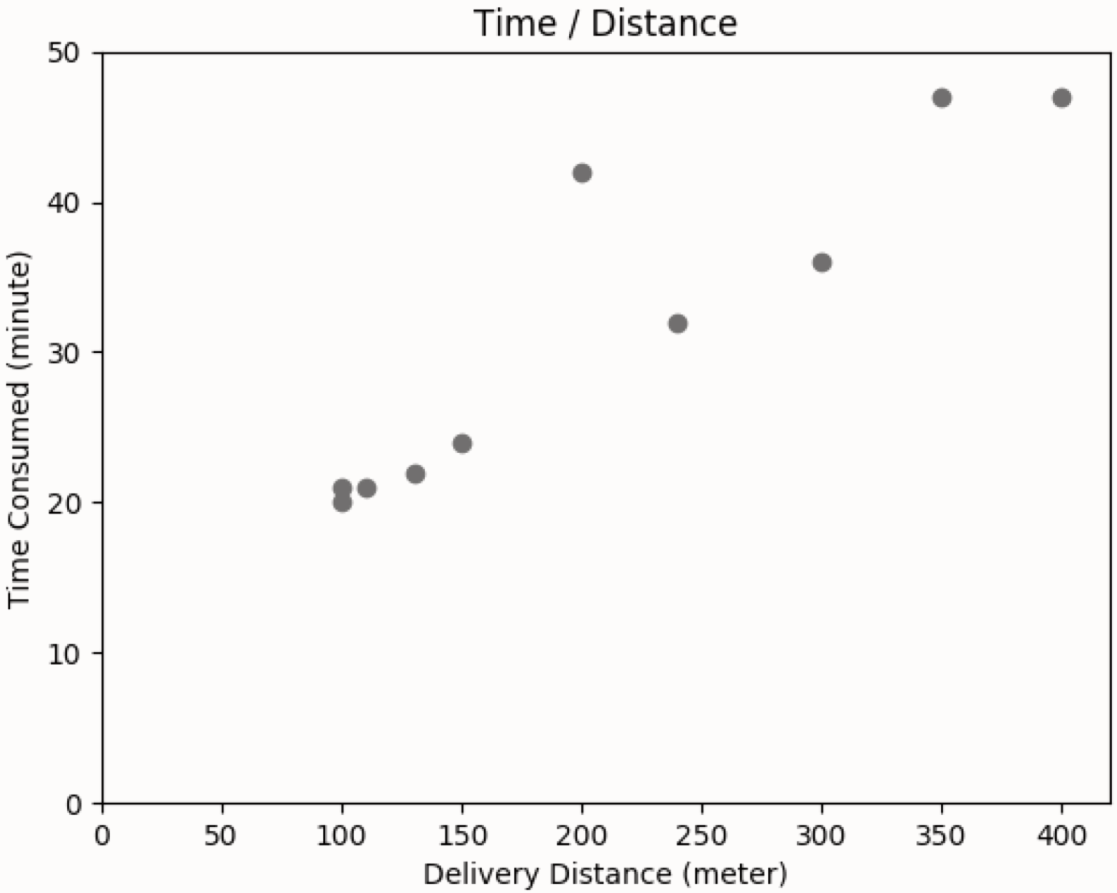
Pizza delivery time by distance

delivery distance	delivery time
100m	20min
150m	24min
300m	36min
400m	47min
130m	22min
240m	32min
350m	47min
200m	42min
100m	21min
110m	21min

Pizza delivery time by distance



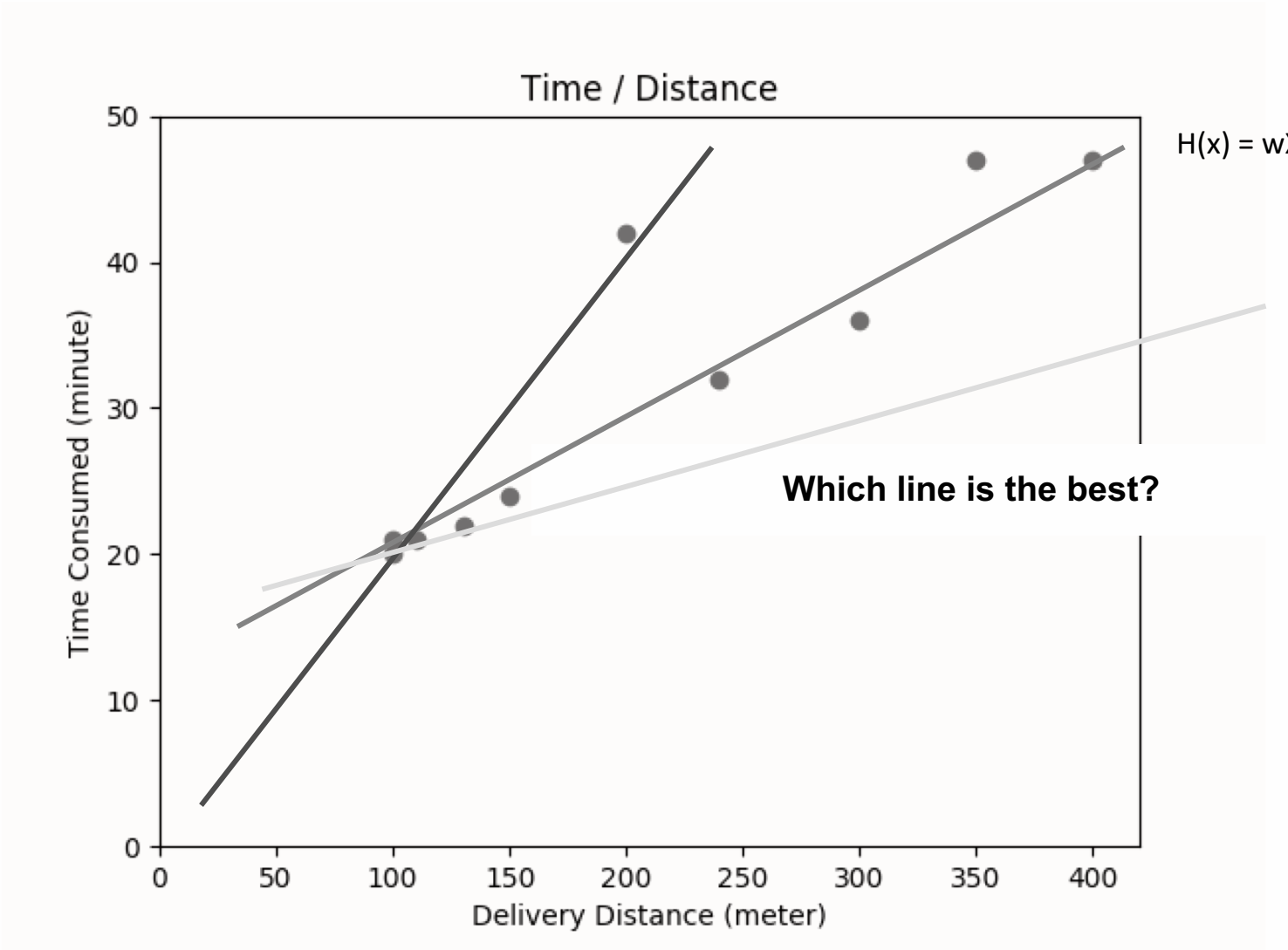
Pizza delivery time by distance



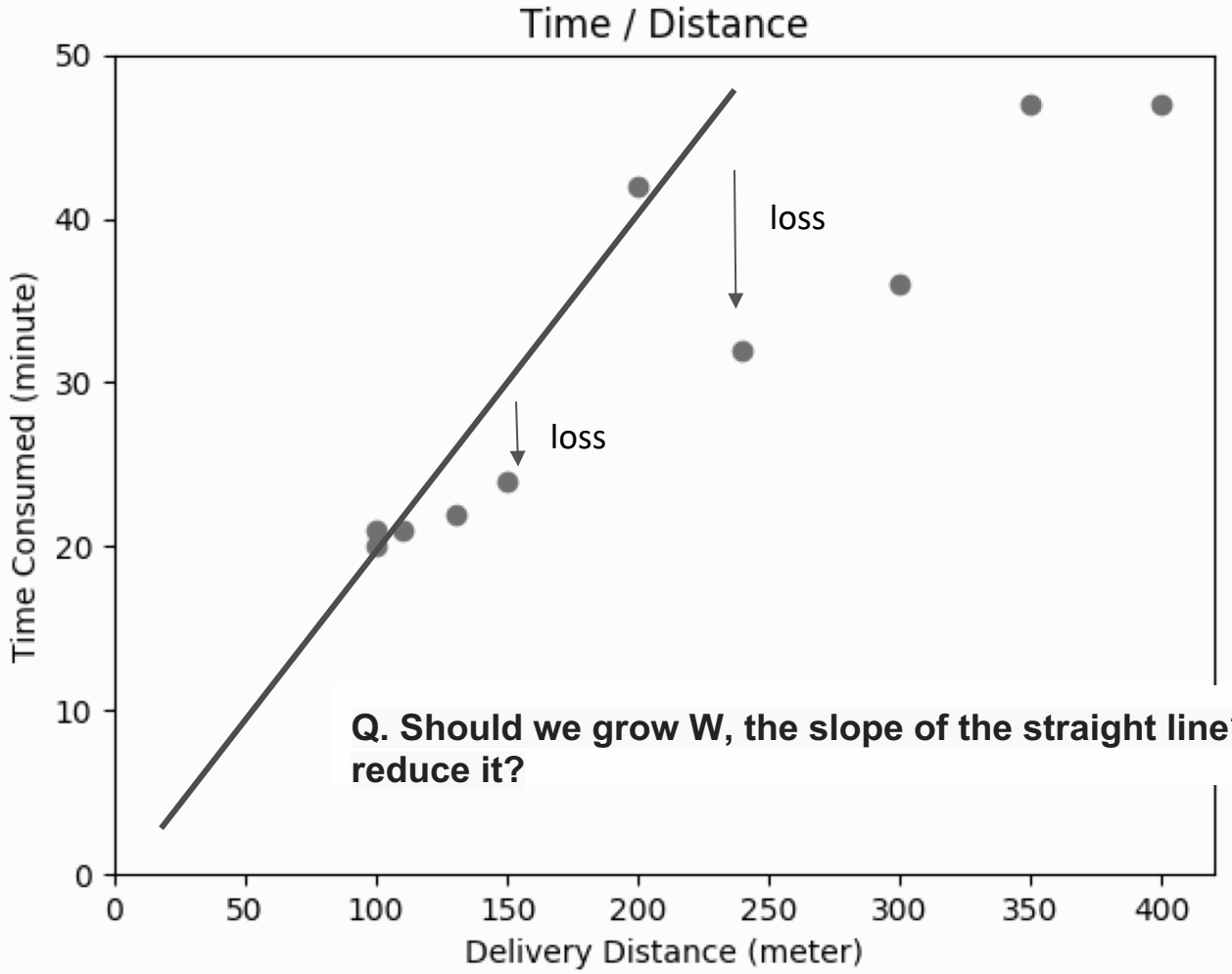
Q. How can we predict the delivery time based on distance?

Q. How can we find the best straight line?

Pizza delivery time by distance



Pizza delivery time by distance

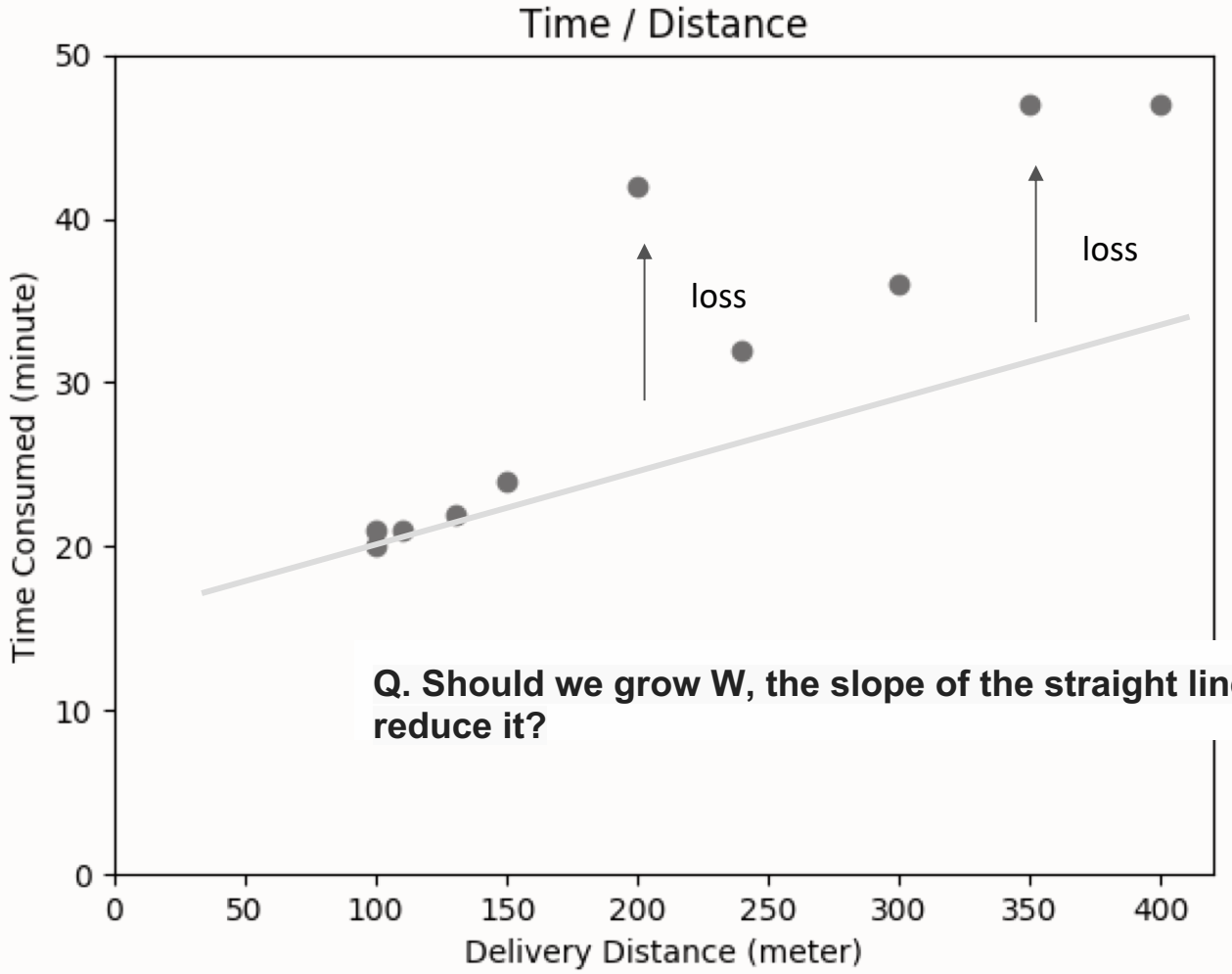


total loss =
(20-20) +
(23-21) +
(30-22) +
(40-41) +
(43-30) +
……..

result =
Positive number

Q. Should we grow W, the slope of the straight line? Should we reduce it?

Pizza delivery time by distance

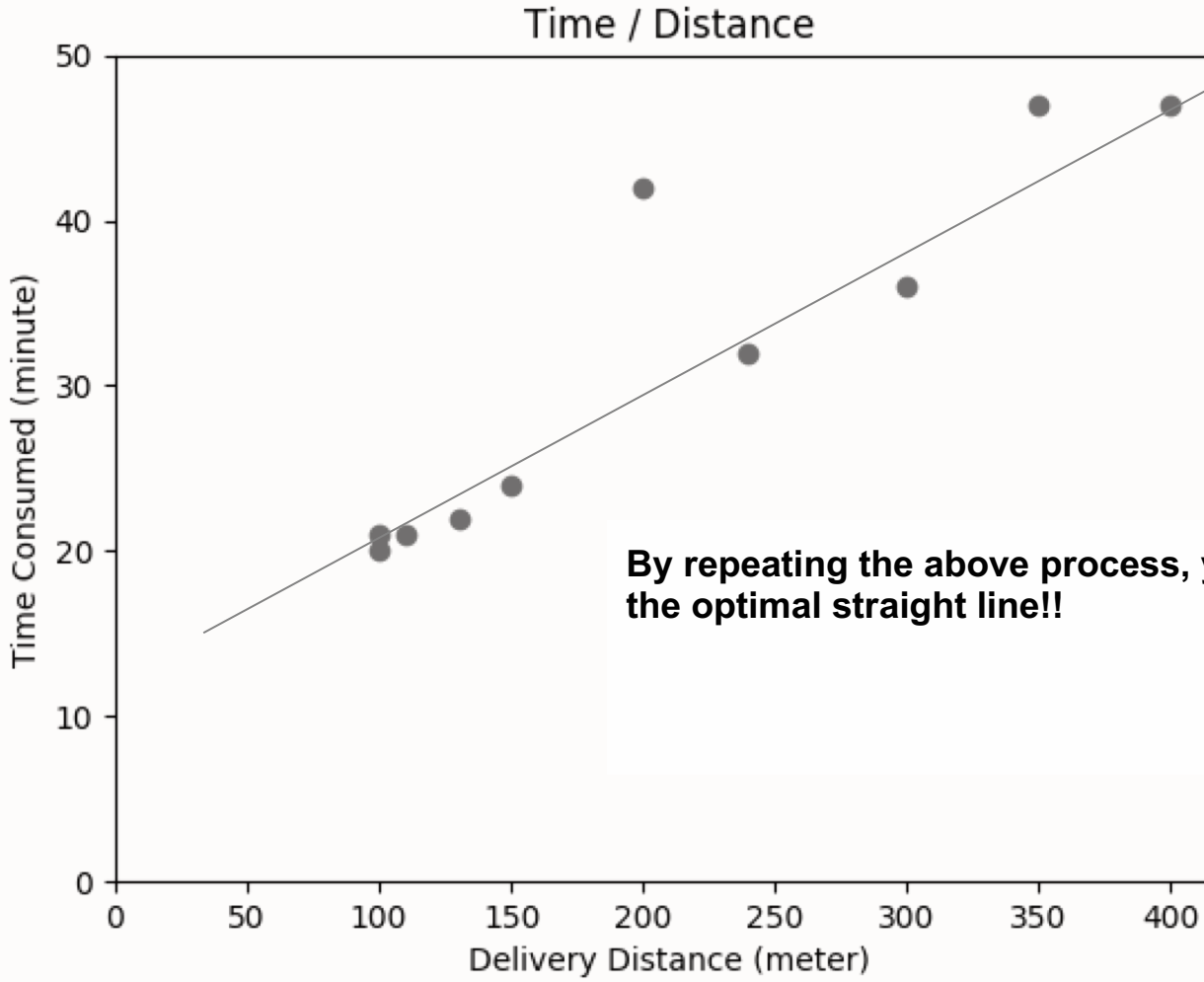


total loss =
(20-21) +
(23-20) +
(24-41) +
(25-31) +
(26-48) +
……..

result =
Negative number

Q. Should we grow W, the slope of the straight line? Should we reduce it?

Pizza delivery time by distance



$Y = wX + b$

By repeating the above process, you can find the optimal straight line!!

Mean Square Error(MSE)

$$cost = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$

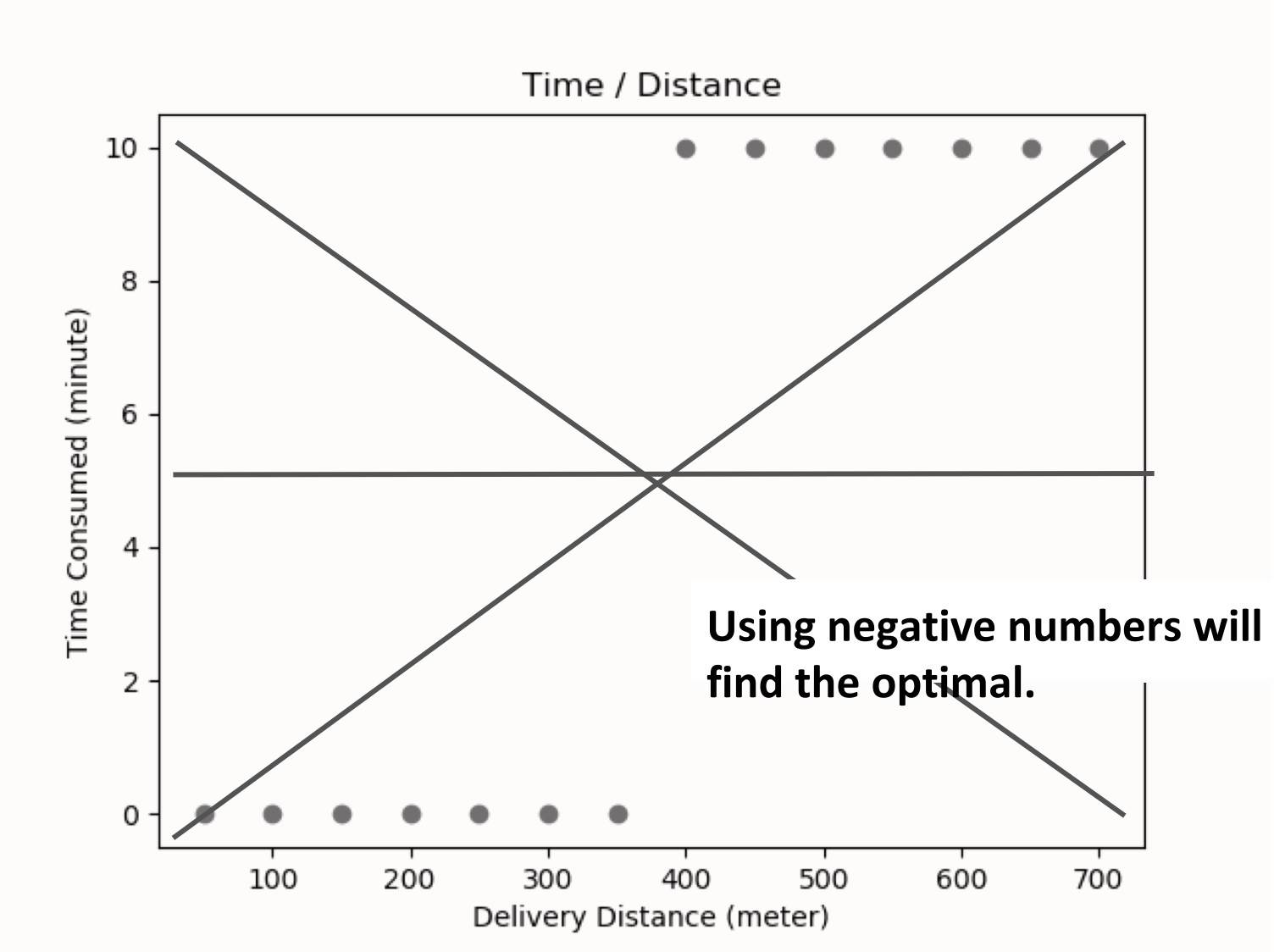
$$H(x) = Wx + b$$

Q. Why shouldn't we use negative numbers?

$$cost(W, b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$

Q. Why not use absolute values?

Why shouldn't we use negative numbers?



Using negative numbers will not find the optimal.

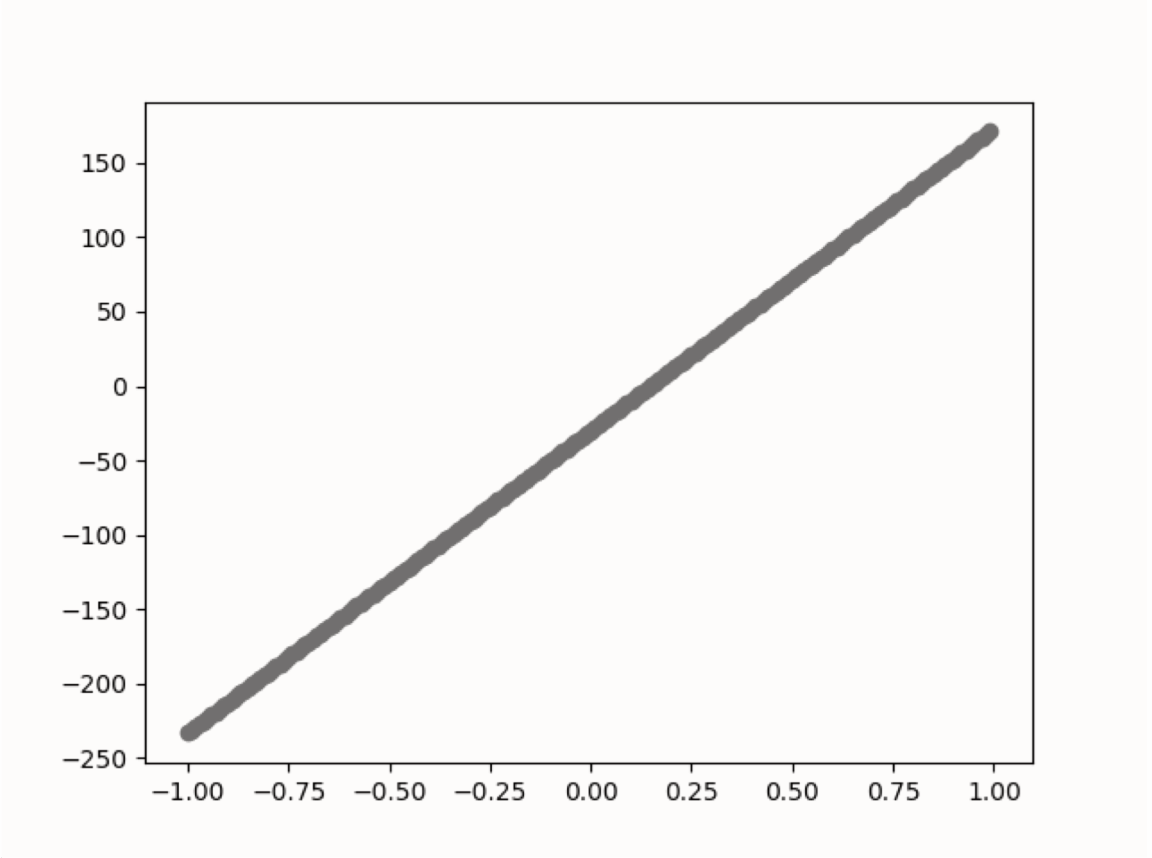
Why not use absolute values?

1. Faster learning with greater weight for large errors

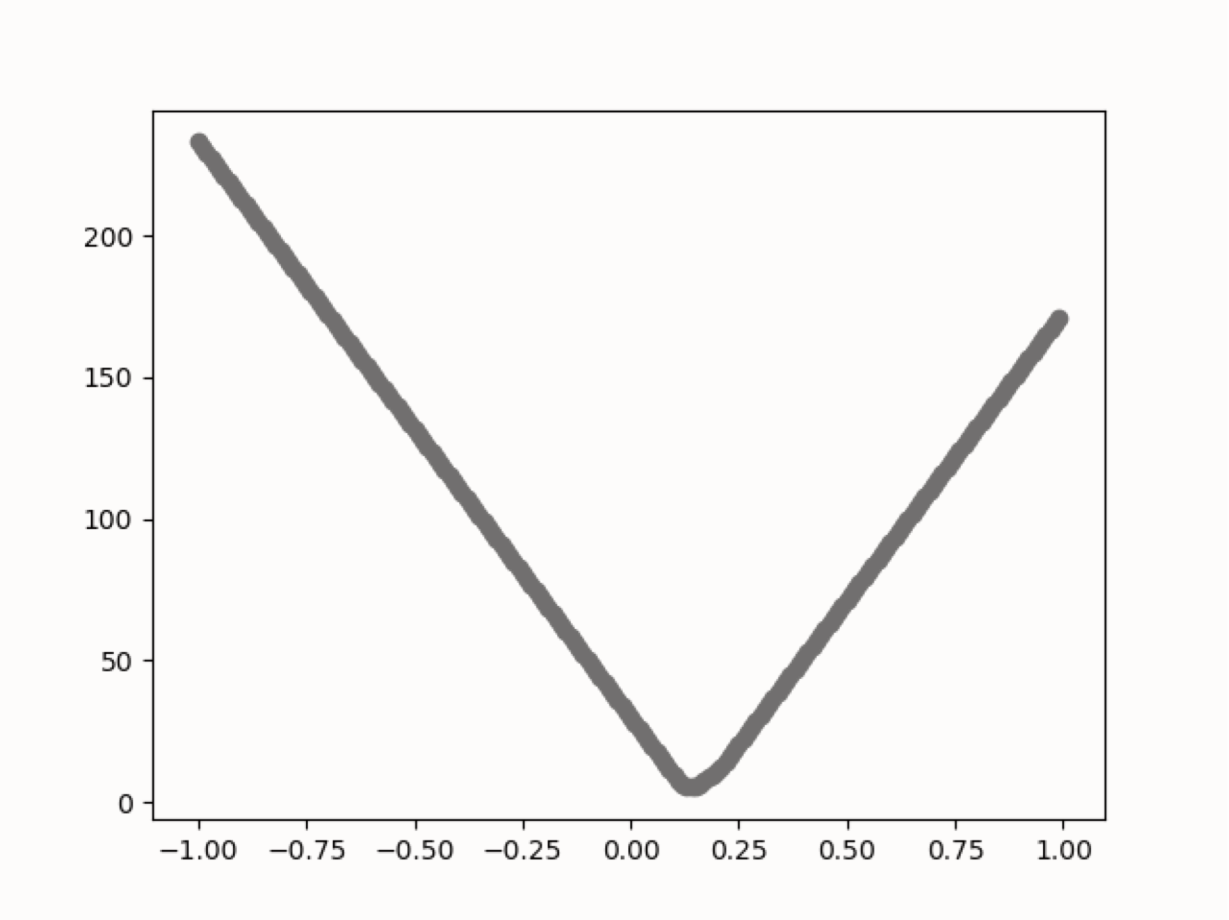
2. To find the optimal weight effectively by making MSE into a convex function.

3. Absolute values are not differentiable.

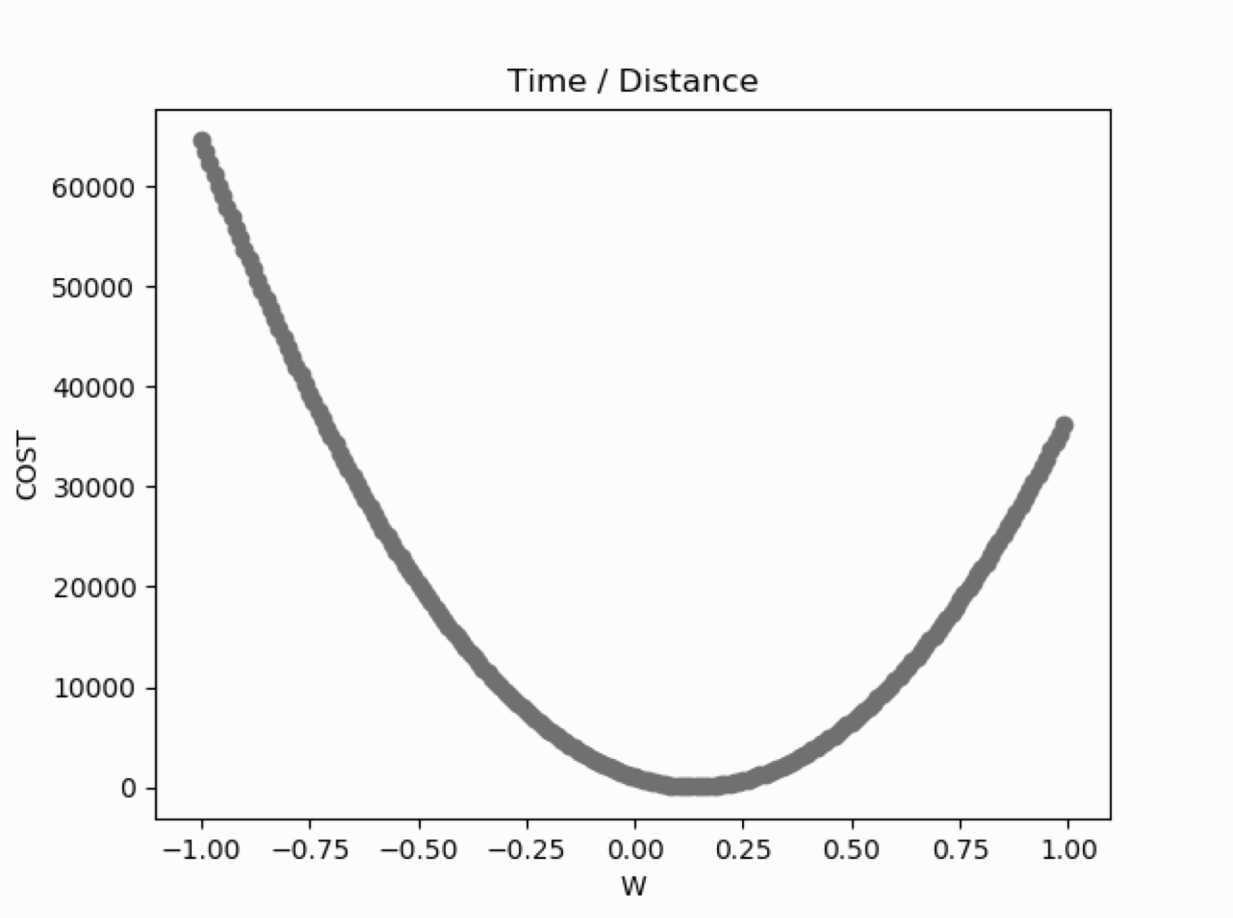
Compare


150
100
50
0
−50
−100
−150
−200
−250
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00

Compare



Compare


Time / Distance
COST
60000
50000
40000
30000
20000
10000
0
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
W

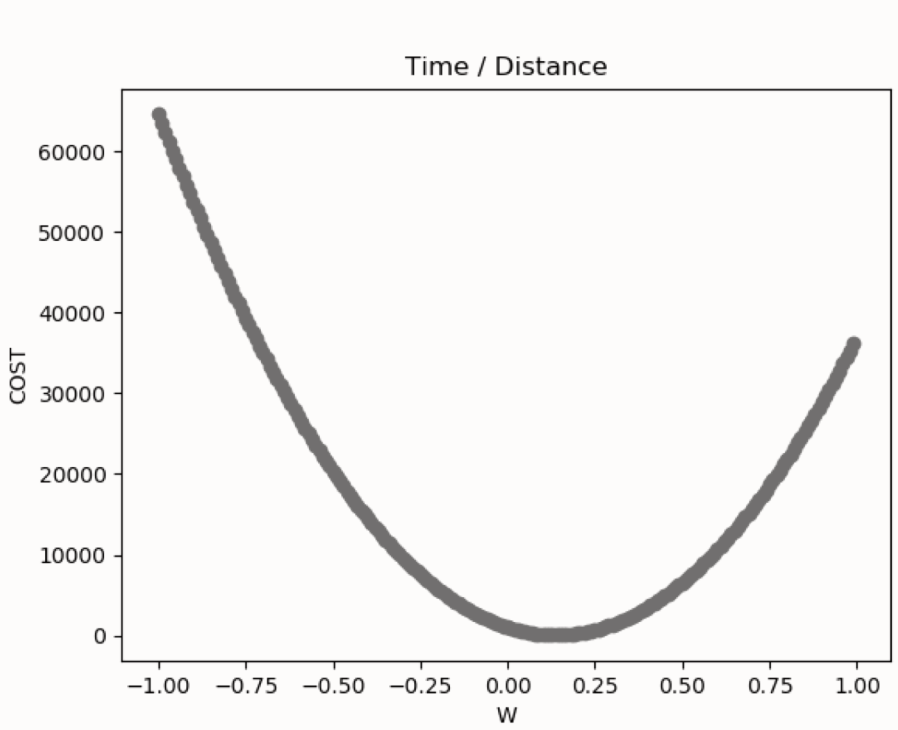
Gradient Descent: Intuition

- Let's go down the curve !

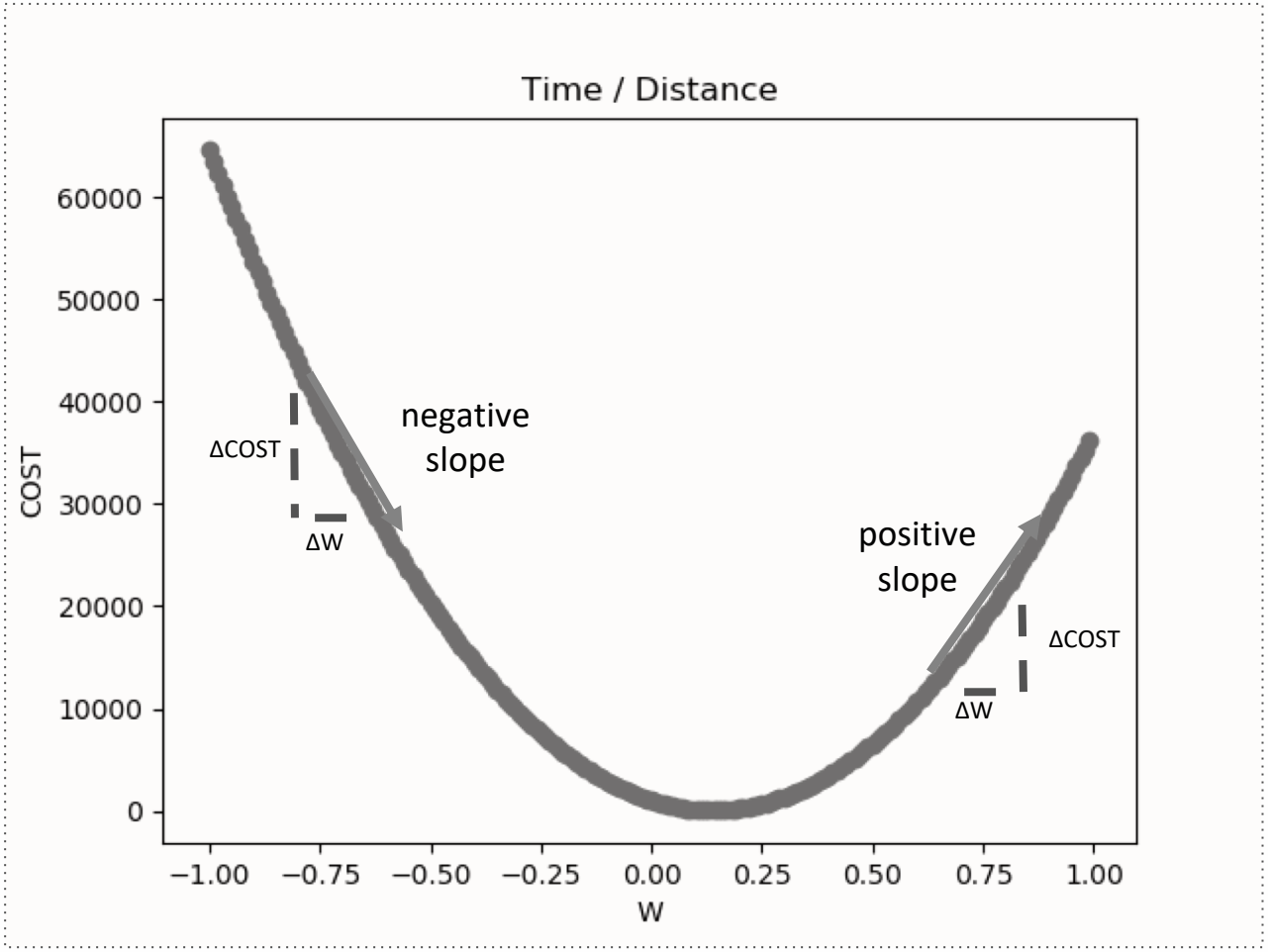
- The greater the slope, the farther!

- Calculate “Gradient” !

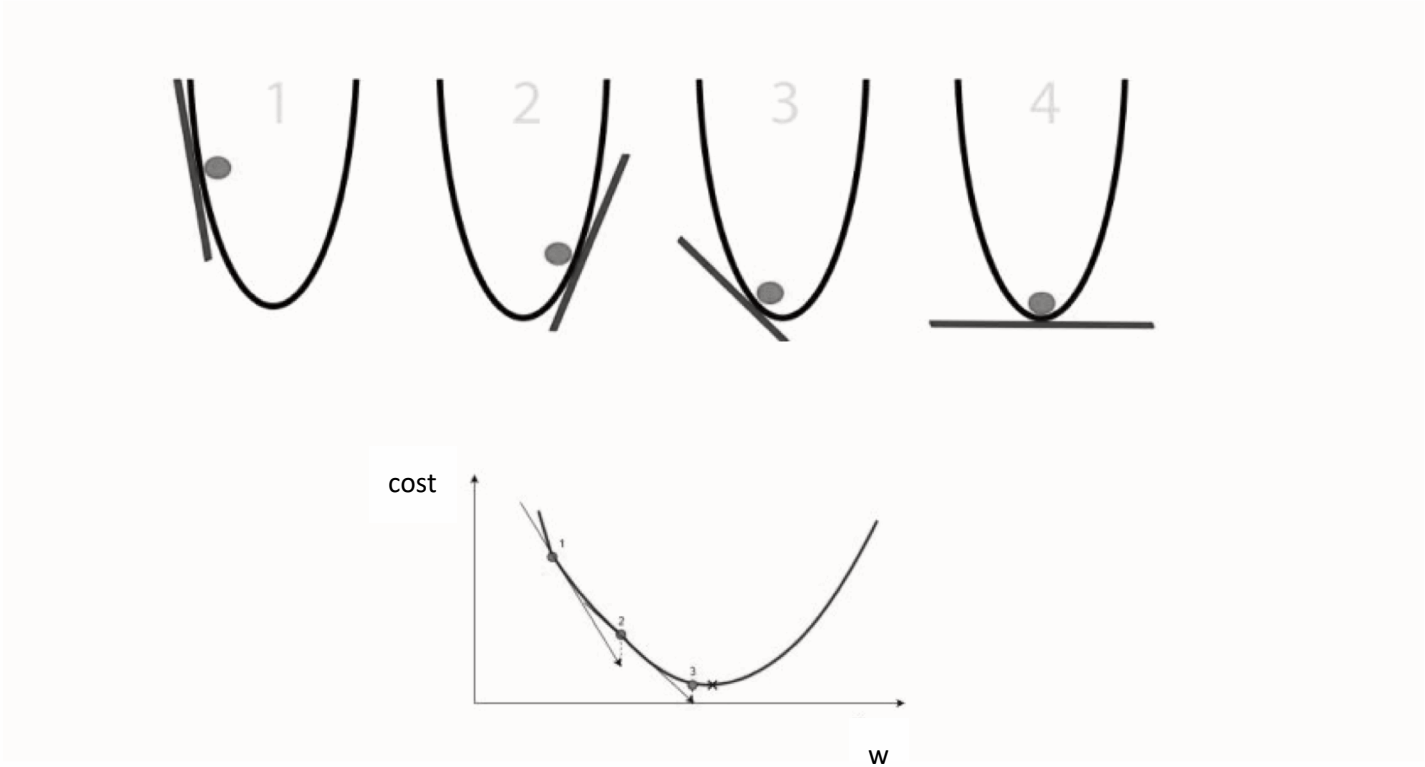
$$\frac{\partial cost}{\partial W} = \nabla W$$



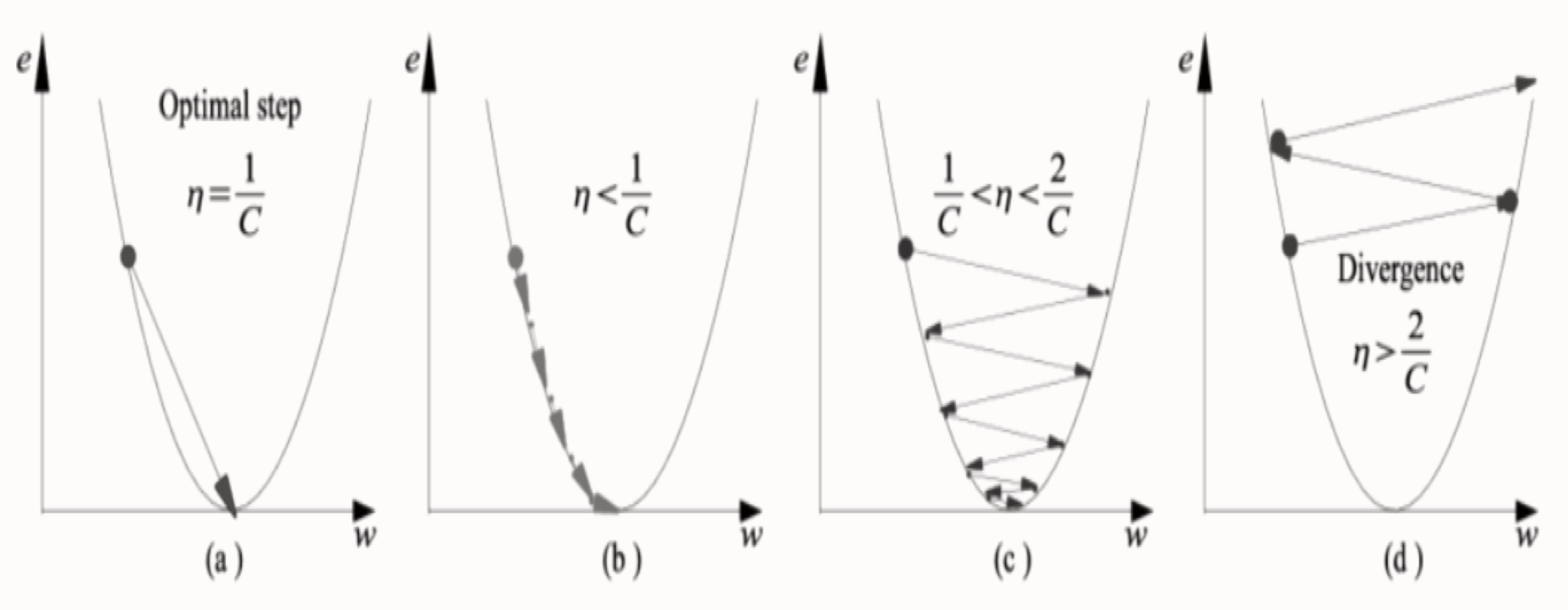
Gradient Descent



Gradient Descent: The Math



Gradient Descent



Gradient Descent: The Math

$$cost(W) = \frac{1}{m}\sum_{i=1}^{m}\left(W x^{(i)} - y^{(i)}\right)^2$$

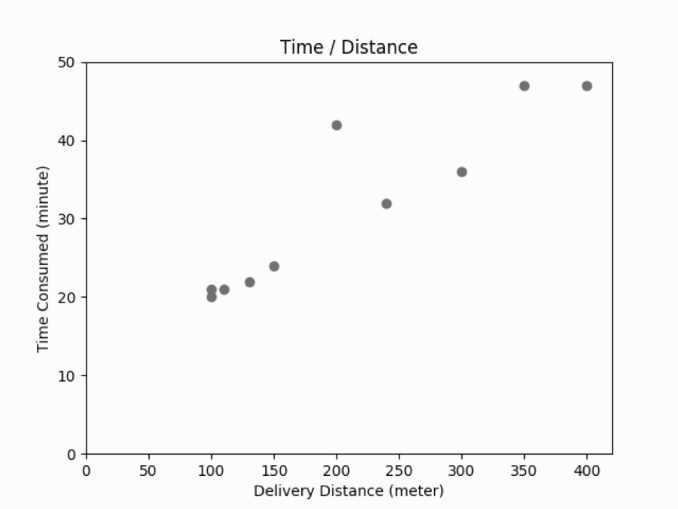
$$\nabla W = \frac{\partial cost}{\partial W} = \frac{2}{m}\sum_{i=1}^{m}\left(W x^{(i)} - y^{(i)}\right)x^{(i)}$$

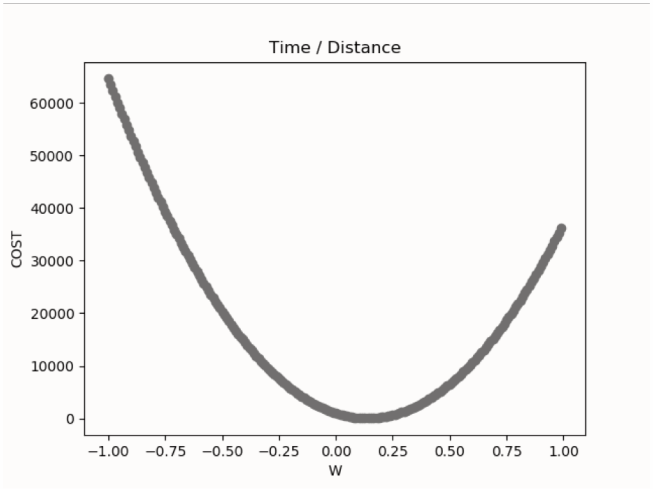
$$W := W - \alpha \nabla W$$

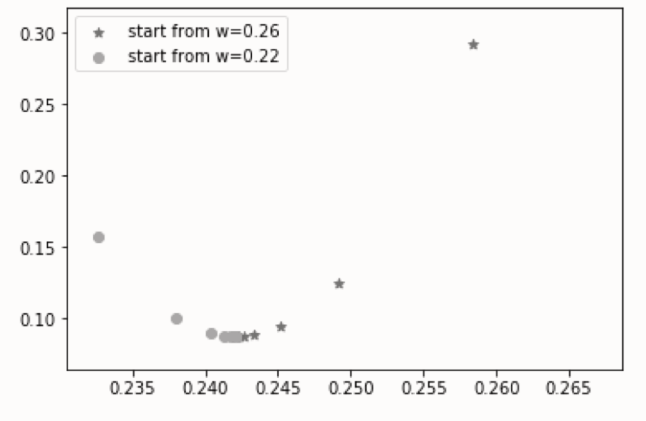
Learning rate

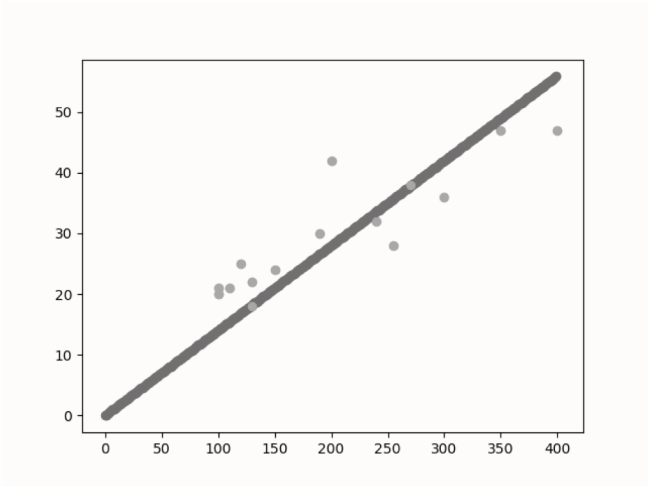
Gradient

Gradient descent Result





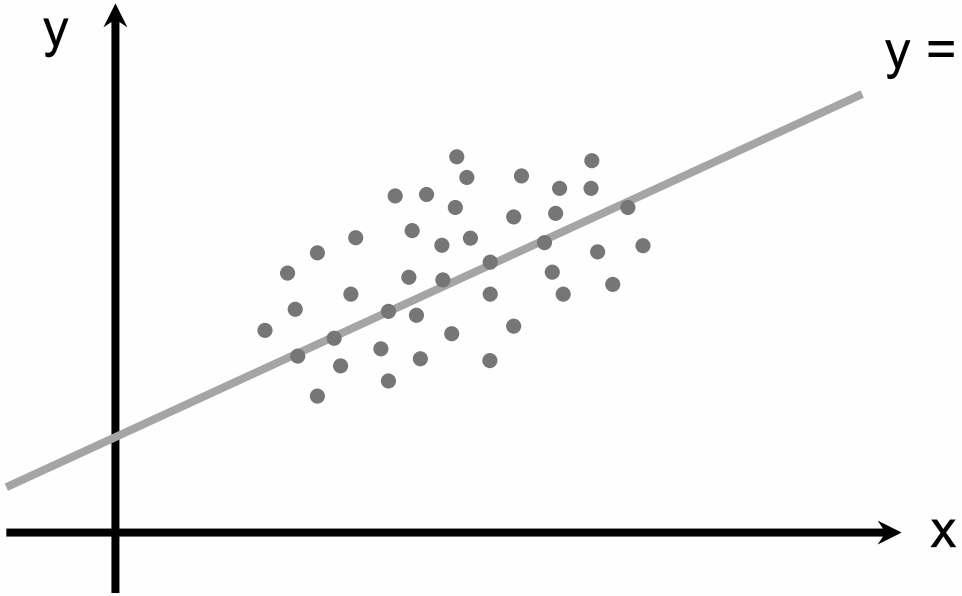




- Python practice (Room 410)

What is Linear Regression?

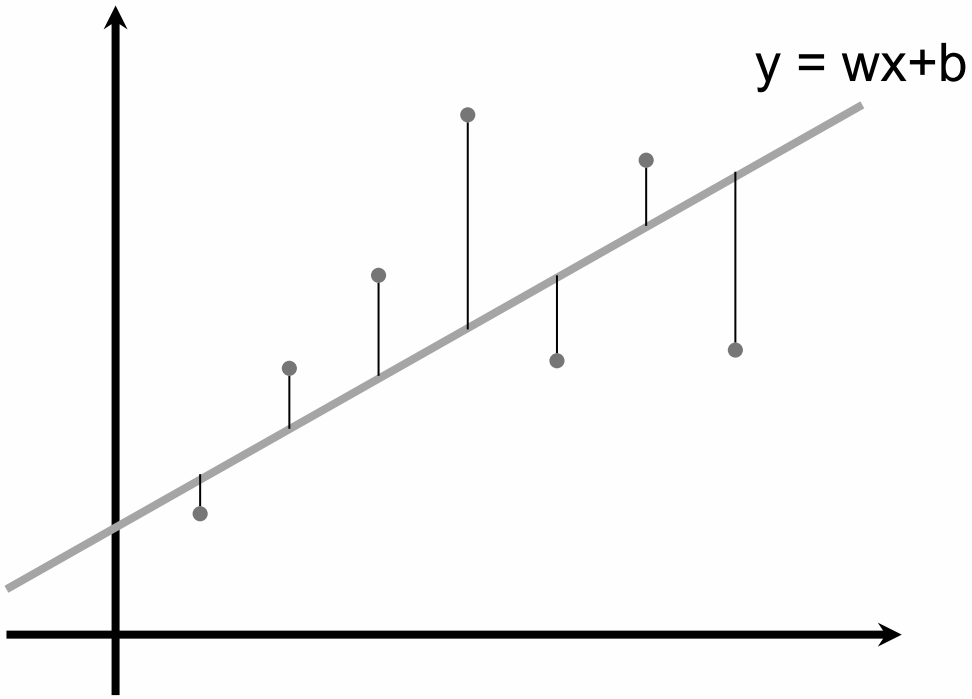
Finding the variables w and b that best describe the y value for x.



A regression technique that models the linear correlation of the dependent variable y with one or more independent variables x.

What is Linear Regression?

You need a **metric** to measure whether you predicted well or not.



Mean Squared Error (MSE)

$$MSE = \frac{(y_1 - y_2)^2}{n}$$

The mean of the square of the distance of two values.

$$Loss = \frac{(y^* - y)^2}{n} = \frac{(w^*x + b^* - y)^2}{n}$$

If any w * and b * are initial values, the Loss value is obtained by w *, b * in fixed x, y (data).

ex) When y = 0.5 * x + 4 and w * = 3 and b * = 2, what is the loss at x = 5?

$$Loss = \frac{(17 - 5.5)^2}{n} = \frac{(11.5)^2}{n}$$

What is Linear Regression?

Goal: Find (w, b) to minimize Loss.
Random Search? **NO!**
∴ Many times + impossible to be sure of correct answers

Can't we get it through the Loss value? **Gradient!**

Descent!