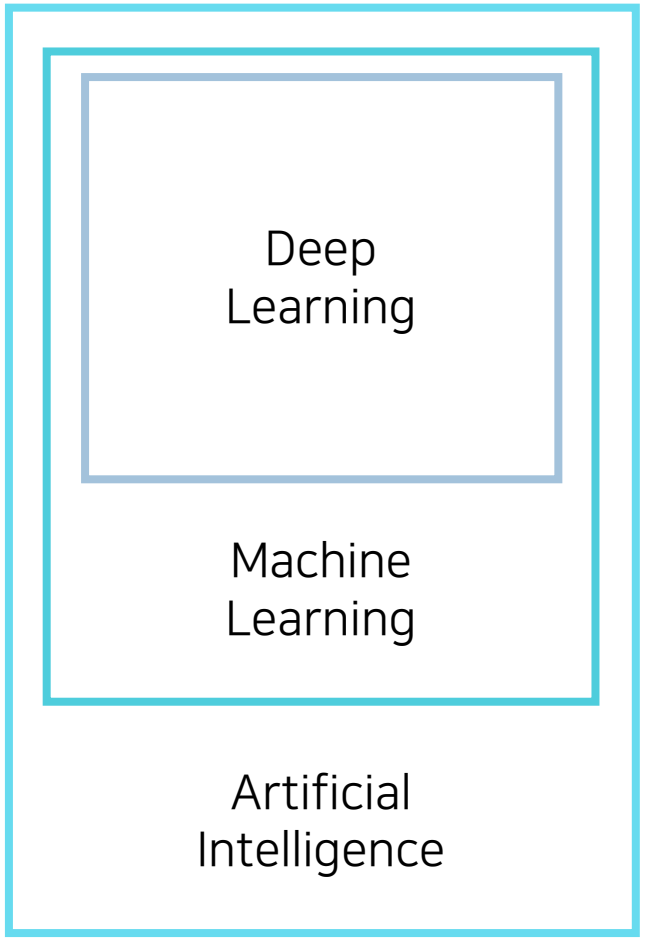
Overview Of Machine Learning #2

2019 / Dec / 31

By SuWhan Baek
zhsjzhsj@gmail.com

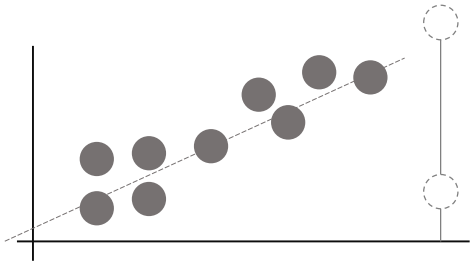
☑ Review Previous Lecture

Method	Machine Learning	Deep Learning	
Preprocessing	Denoise	Cleaning	Transformation
Model	Supervised	Unsupervised	Reinforcement
Problem Type	Classification	Regression	Segmentation
Function	Cost	Loss	Activate



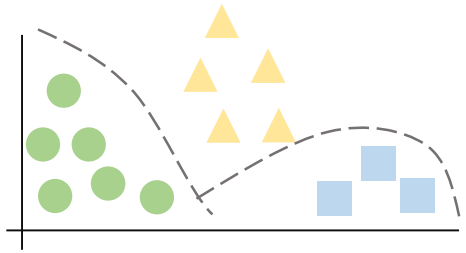
☑ Review Previous Lecture

Supervised



Get Output Result
Based on Given X to Y

Unsupervised



Get Output Result
Based on Data Itself

Only Data

Data Visualization is recommended!!

✅ Machine Learning Words

모델 : Model

how inputs are analyzed and manipulated according to various mathematical concepts and theory in order to achieve a certain task.

정확도 : Accuracy; Acc

Accuracy = Correct Predict / Total Predict

매개변수 : Hyperparameter

configuration variable that is internal to the model and whose value can be estimated from data.
[Hyperparameter is User input Model Parameters]

서브젝트 : Subject

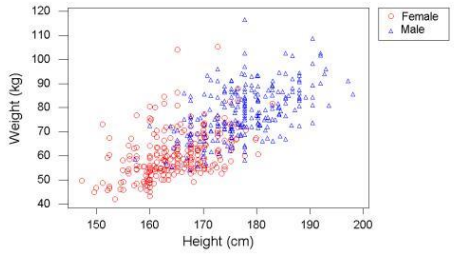
피험자; Can be interpreted as a bunch of data from one person
Intrasubject / Intersubject

배치 : Batch

Group of training samples

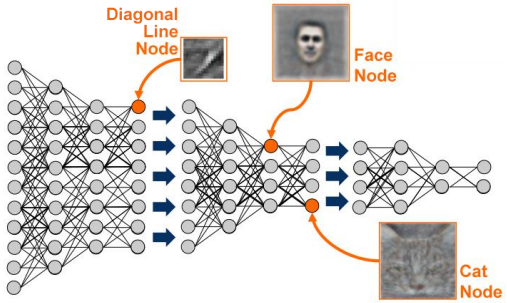
What Is The **Feature?**

Machine Learning



User Feature Based

Deep Learning



Auto Feature Extraction

What Is The **Feature?**

What Is The **Feature?**

Feature :

An individual measurable property or characteristic of a phenomenon being observed.

관찰되는 현상의 개별 측정 가능한 속성 또는 특성

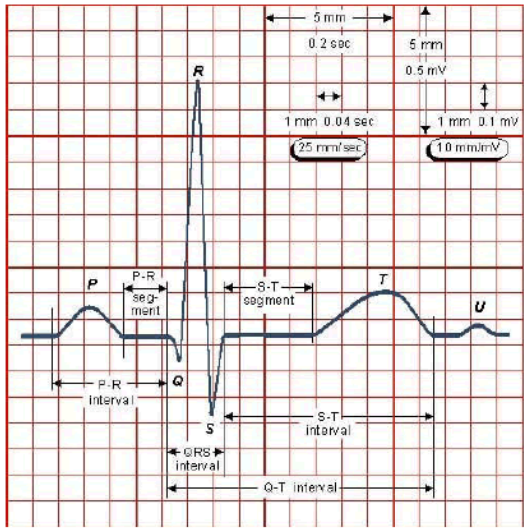
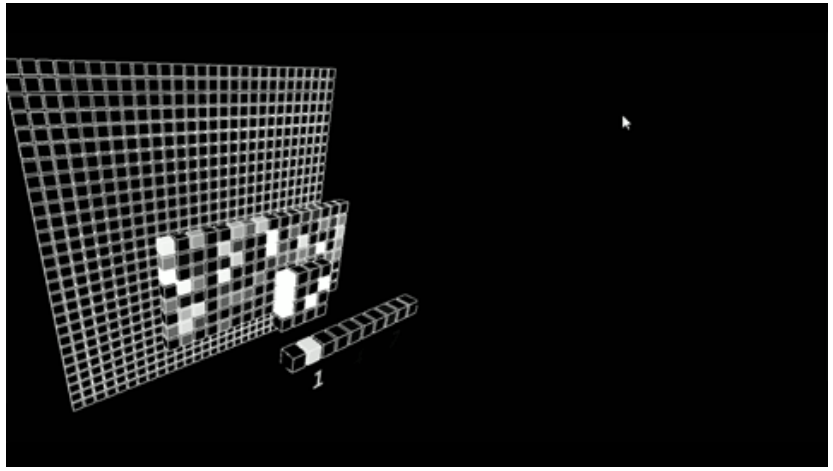


Figure 1. Normal ECG waveform [4].



https://okdalto.github.io/VisualizeMNIST_web/

What Is The **Feature?**

Feature :

An individual measurable property or characteristic of a phenomenon being observed.

관찰되는 현상의 개별 측정 가능한 속성 또는 특성

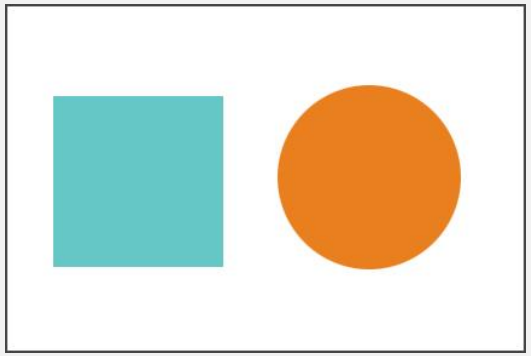
Feature Extraction

Feature Selec

Featuring Met

Feature Const

What About Image??



- Length
- Corner
- Edge
- Color

Computer Vision Area

What Is The **Feature?**

Feature :

An individual measurable property or characteristic of a phenomenon being observed.

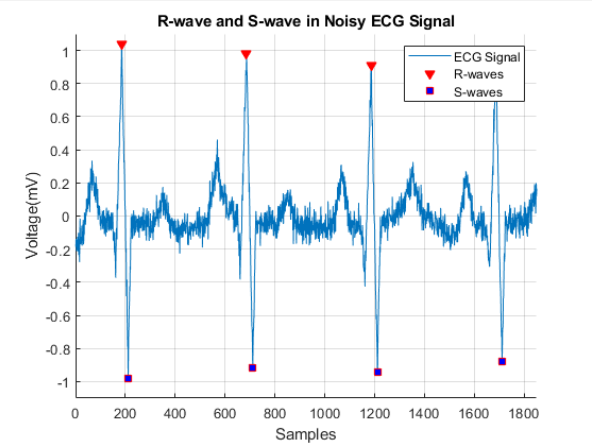
관찰되는 현상의 개별 측정 가능한 속성 또는 특성

Feature Extraction

Feature Selec

Featuring Met

Feature Const



- Length between each peak point
- Min Value between each peak
- Max Peak value
- Differential value of each rising point

What Is The **Feature?**

Feature :

An individual measurable property or characteristic of a phenomenon being observed.

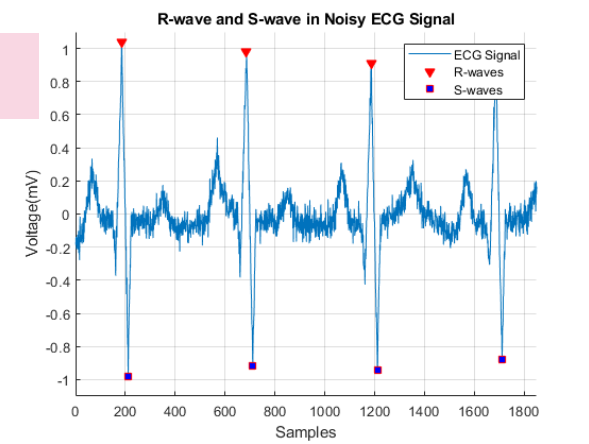
관찰되는 현상의 개별 측정 가능한 속성 또는 특성

Feature Extra

Feature Selection

Featuring Met

Feature Const



Select Good Features!

- Length between each peak point
- Min Value between each peak
- Max Peak value
- Differential value of each rising point

What Is The **Feature?**

Feature :

An individual measurable property or characteristic of a phenomenon being observed.

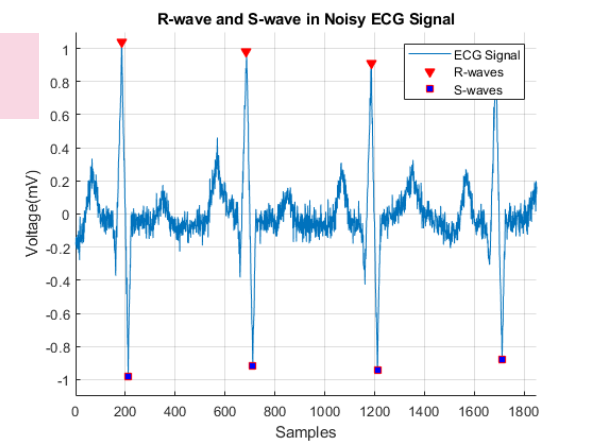
관찰되는 현상의 개별 측정 가능한 속성 또는 특성

Feature Extra

Feature Selection

Featuring Met

Feature Const



- Length between each peak point
- Min Value between each peak

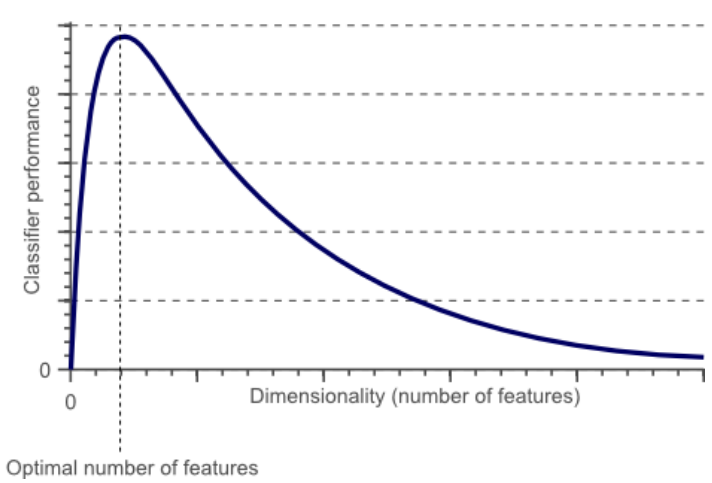
Time	Preak	Peak Len	Min
1	1	0.125	-0.91
2	2	0.100	-0.8
3	3	0.126	-0.9
4	4	0.148	-0.87
5	5	0.150	-0.99
6	6	0.152	-0.88
7	7	0.178	-0.78

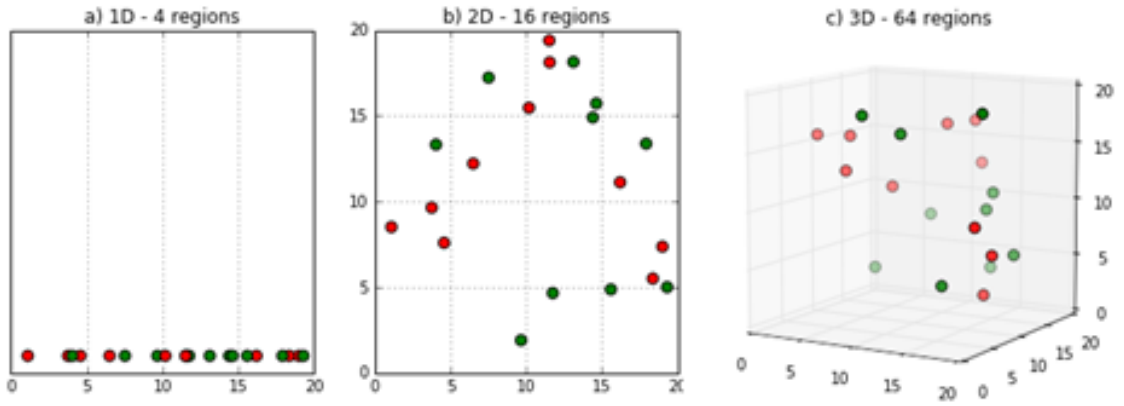
About Number of Feature

Too Many!

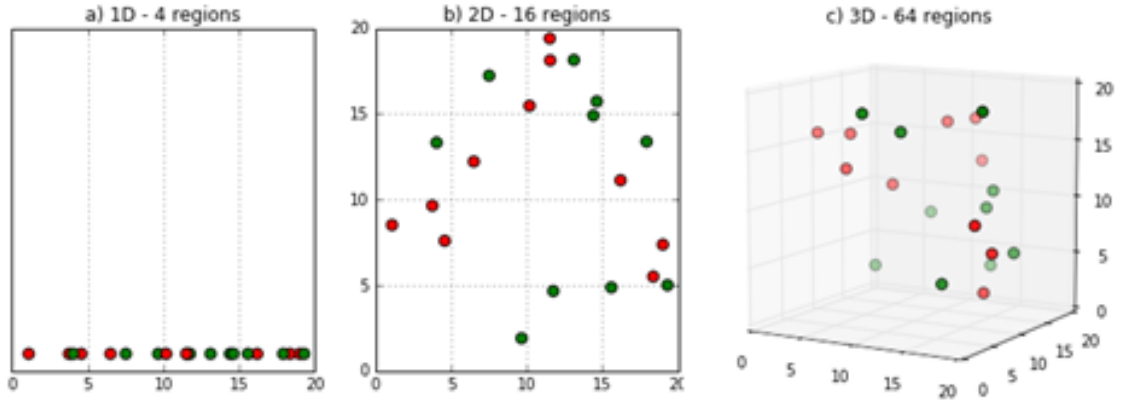
Time	Preak	Peak Len	Min	Light	Weight	OA	Temp	EE m	Peak Term	Second DR	ARE	Gender	Age
1	1	0.125	-0.91	15	88	0	26	0.1548	0.15	15	0.8410	Male	24
2	2	0.100	-0.8	84	87	1	26	0.1048	0.49	66	0.8461	Male	24
3	3	0.126	-0.9	39	64	1	27	0.1048	0.4	68	0.8484	Male	28
4	4	0.148	-0.87	48	54	1	28	0.0448	0.98	15	0.8418	Male	23
5	5	0.150	-0.99	26	67	1	26	0.0548	0.43	23	0.8465	Male	30
6	6	0.152	-0.88	74	91	0	24	0.0148	0.26	84	0.8415	Male	41
7	7	0.178	-0.78	18	76	1	22	0.0548	0.48	23	0.8415	Male	18

More Feature → More Dimension Required → More Compute Resource Required

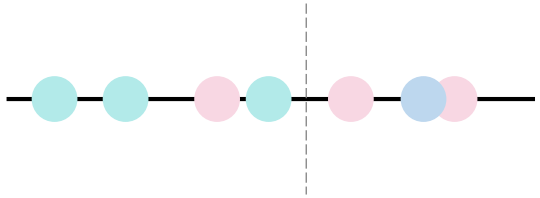




About Dimension

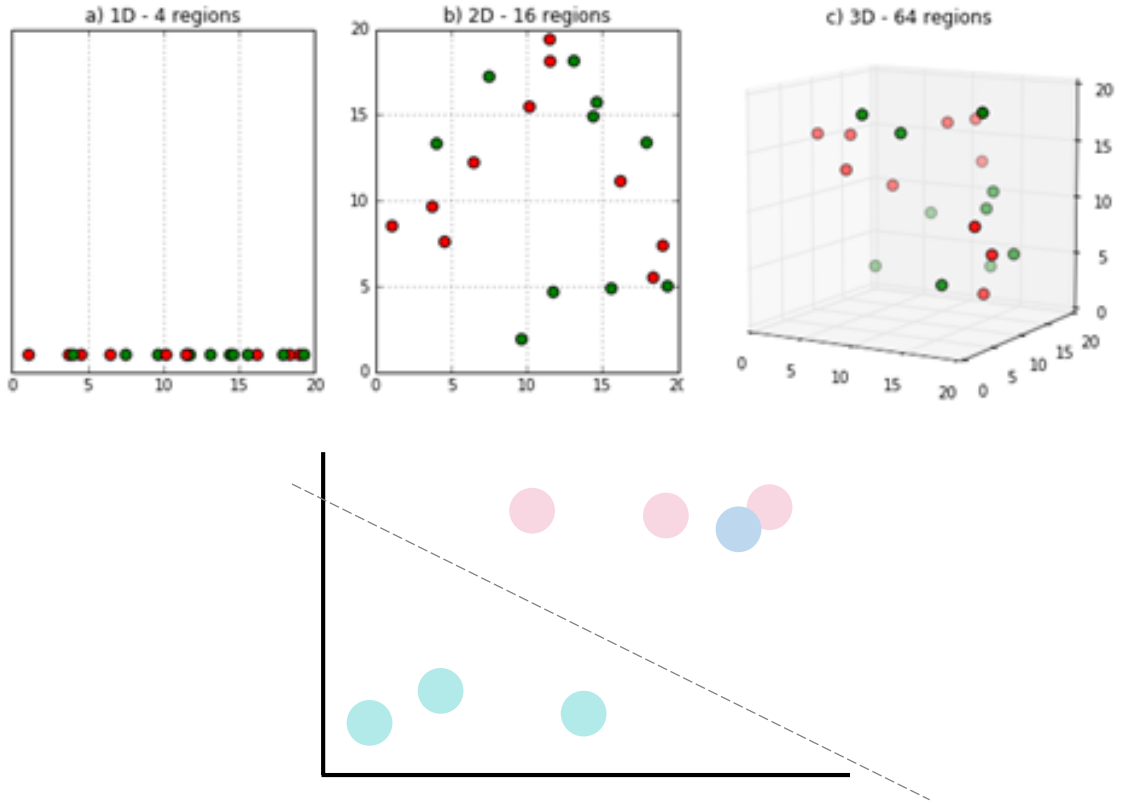


Multiple Classifier : 다중 분류기



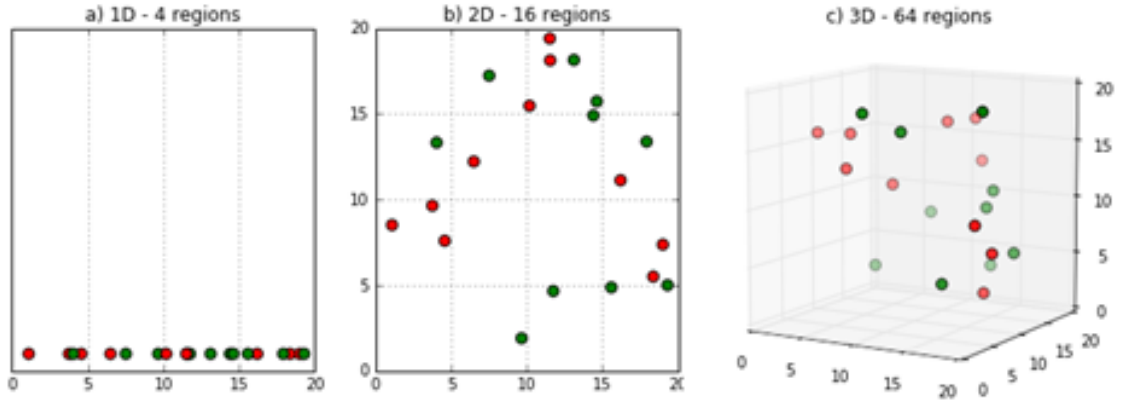
On 1D Space : Nothing Classified Clearly

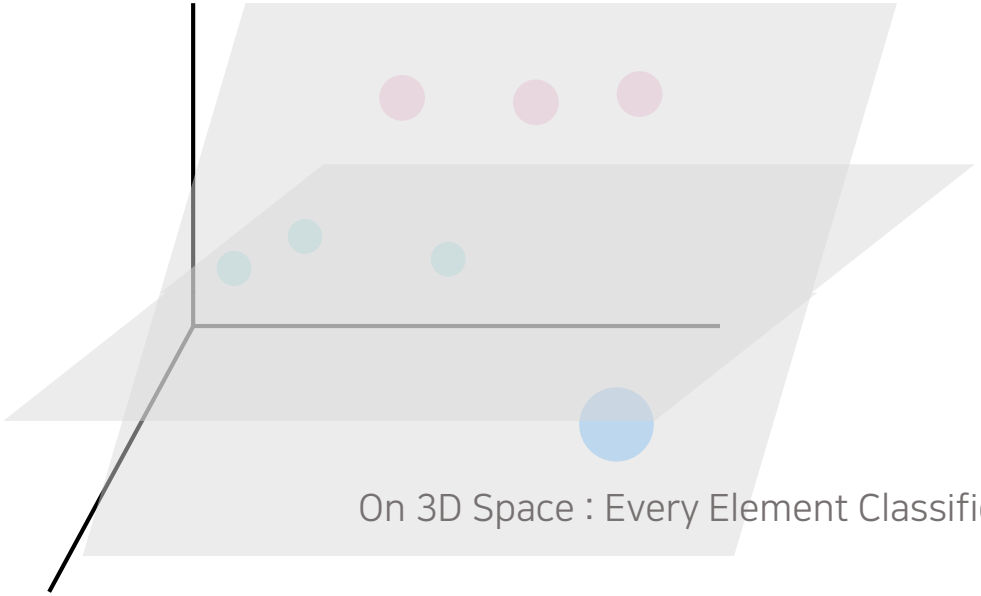
About Dimension



On 2D Space : Green Classified Clearly Others is Not

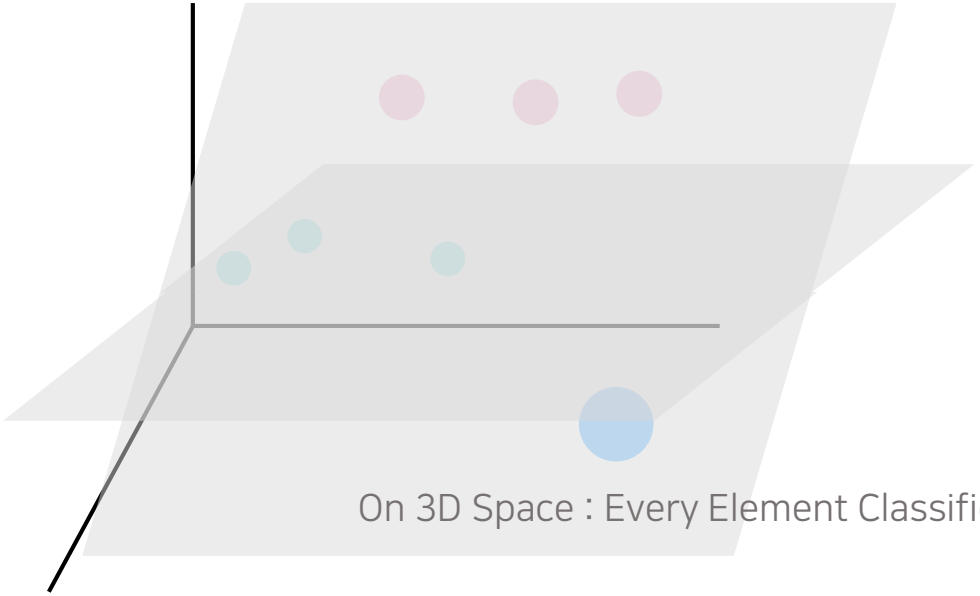
About Dimension



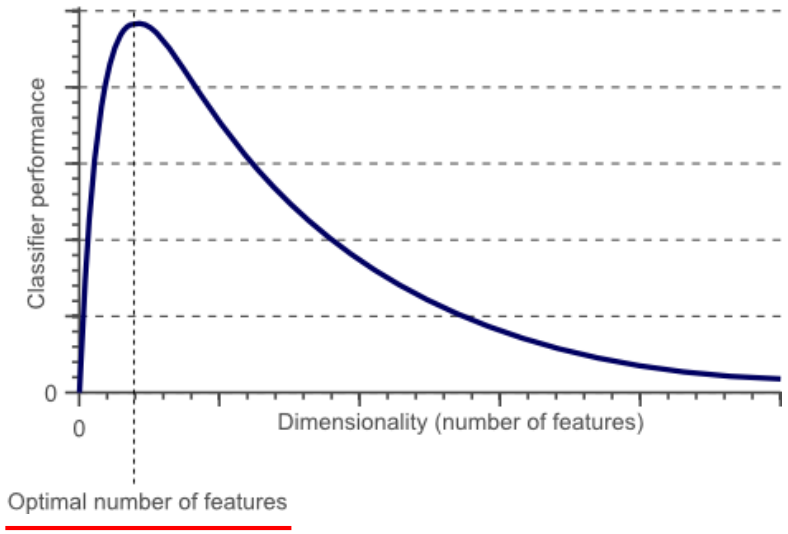


On 3D Space : Every Element Classified Clearly

About Dimension



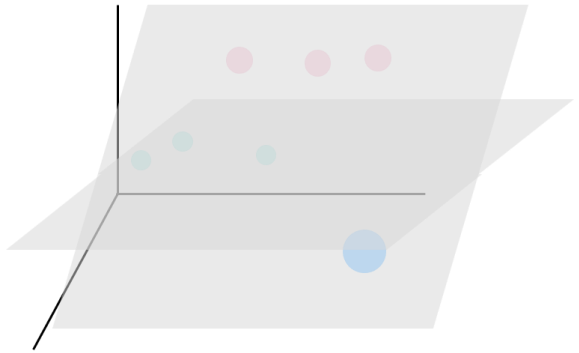
On 3D Space : Every Element Classified Clearly



Finding the optimal dimension is the most important

About Dimensionality with Visualization

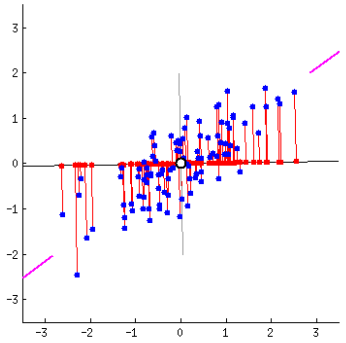
You can see, imagine, draw 3D Fields

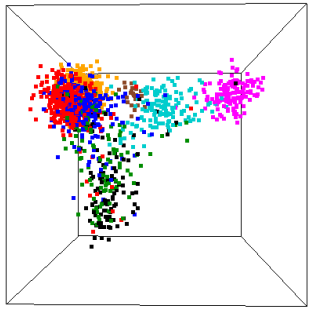


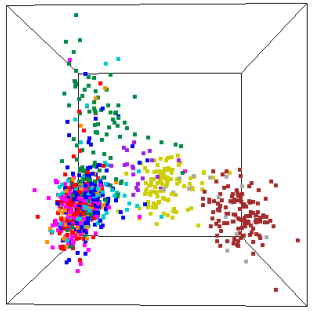
Than, What About [n] D?

About Dimensionality with Visualization

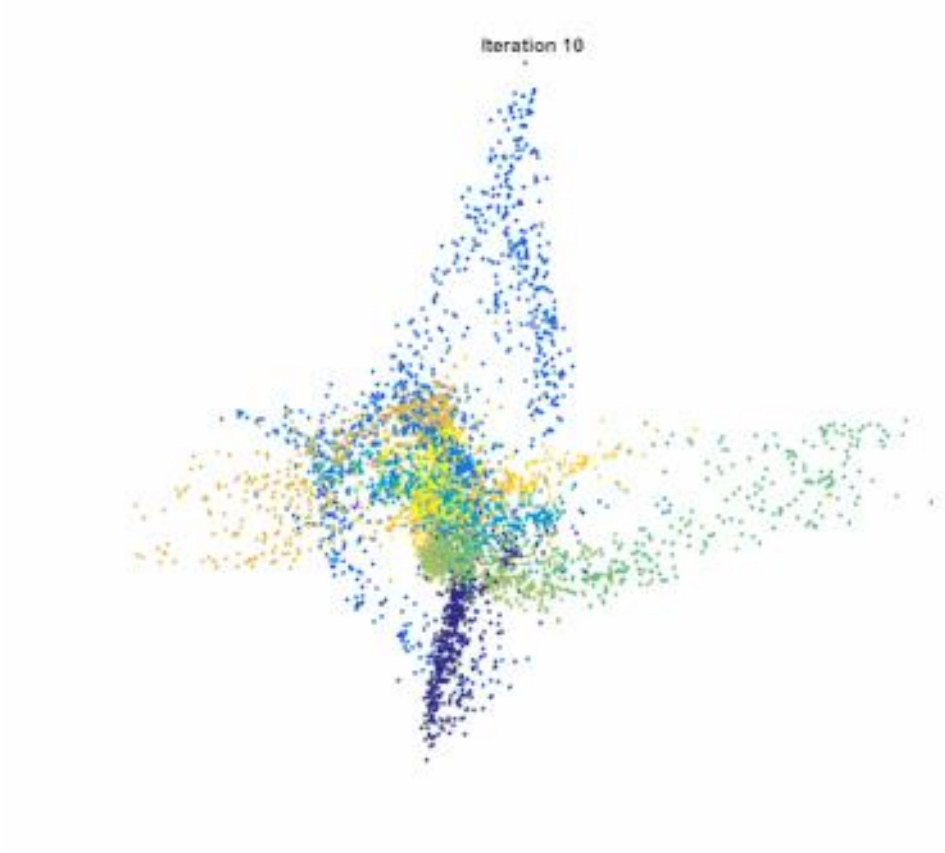
Data Visualization
PCA







Data Visualization
T-SNE



Feature Preprocessing

Time	Preak	Peak Len	Min	EE m	Peak Term	Second DR	ARE	Gender	Age
1	1	0.125	-0.91	0.1548	0.15	15	0.8410	Male	24
2	2	0.100	-0.8	0.1048	0.49	66	0.8461	Male	24
3	3	0.126	-0.9	0.1048	0.4	68	0.8484	Male	28
4	4	0.148	-0.87	0.0448	0.98	15	0.8418	Male	23
5	5	0.150	-0.99	0.0548	0.43	23	0.8465	Male	30
6	6	0.152	-0.88	0.0148	0.26	84	0.8415	Male	41
7	7	0.178	-0.78	0.0548	0.48	23	0.8415	Male	18

Feature Scaling :

Normalization

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Standardization

$$x_{new} = \frac{x - \mu}{\sigma}$$

Prevents features from being ignored by the relative size

Feature Preprocessing

Normalization

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Standardization

$$x_{new} = \frac{x - \mu}{\sigma}$$

Time	Preak	Peak Len	Min	EE m	Peak Term	Second DR	ARE	Gender	Age
1	1	0.125	-0.91	0.1548	0.15	15	0.8410	Male	24
2	2	0.100	-0.8	0.1048	0.49	66	0.8461	Male	24
3	3	0.126	-0.9	0.1048	0.4	68	0.8484	Male	28
4	4	0.148	-0.87	0.0448	0.98	15	0.8418	Male	23
5	5	0.150	-0.99	0.0548	0.43	23	0.8465	Male	30
6	6	0.152	-0.88	0.0148	0.26	84	0.8415	Male	41
7	7	0.178	-0.78	0.0548	0.48	23	0.8415	Male	18

Prevents features from being ignored by the relative size

Feature Preprocessing

Normalization

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Standardization

$$x_{new} = \frac{x - \mu}{\sigma}$$

Time	Preak	N_Peak Len	Min	EE m	Peak Term	Second DR	ARE	Gender	N_Age
1	1	0.44	-0.91	0.1548	0.15	15	0.8410	Male	0.21
2	2	0	-0.8	0.1048	0.49	66	0.8461	Male	0.21
3	3	0.54	-0.9	0.1048	0.4	68	0.8484	Male	0.23
4	4	0.74	-0.87	0.0448	0.98	15	0.8418	Male	0.45
5	5	0.97	-0.99	0.0548	0.43	23	0.8465	Male	0.8
6	6	0.98	-0.88	0.0148	0.26	84	0.8415	Male	1
7	7	1	-0.78	0.0548	0.48	23	0.8415	Male	0

Prevents features from being ignored by the relative size

Feature Preprocessing

Time	Preak	Peak Len	Min	EE m	Peak Term	Second DR	ARE	Gender	Age
1	1	0.125	-0.91	0.1548	0.15	15	0.8410	Male	24
2	2	0.100	-0.8	0.1048	0.49	66	0.8461	Male	24
3	3	0.126	-0.9	0.1048	0.4	68	0.8484	Male	28
4	4	0.148	-0.87	0.0448	0.98	15	0.8418	Male	23
5	5	0.150	-0.99	0.0548	0.43	23	0.8465	Male	30
6	6	0.152	-0.88	0.0148	0.26	84	0.8415	Male	41
7	7	0.178	-0.78	0.0548	0.48	23	0.8415	Male	18

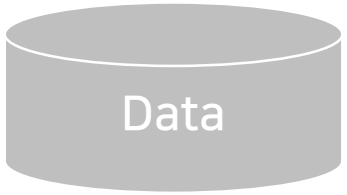
Feature Preprocessing

Time	Preak	Peak Len	Min	EE m	Peak Term	Second DR	ARE	Gender	Age
1	1	0.125	-0.91	0.1548	0.15	15	0.8410	Male	24
2	2	0.100	-0.8	0.1048	0.49	66	0.8461	Male	24
3	3	0.126	-0.9	0.1048	0.4	68	0.8484	Male	28
4	4	0.148	-0.87	0.0448	0.98	15	0.8418	Male	23
5	5	0.150	-0.99	0.0548	0.43	23	0.8465	Male	30
6	6	0.152	-0.88	0.0148	0.26	84	0.8415	Male	41
7	7	0.178	-0.78	0.0548	0.48	23	0.8415	Male	18

Considerable Things :

- Data Balancing (Especially Binary Classifier)
- Too Much Transformation
- Outlier range
- Filter Strength
- Academic evidence of data transformation

What is Cost?



Input

DNN

SVM

DQN

CNN

Random Forest

DDPG

What is Cost?

REDUCE COST

Cost function

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

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

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

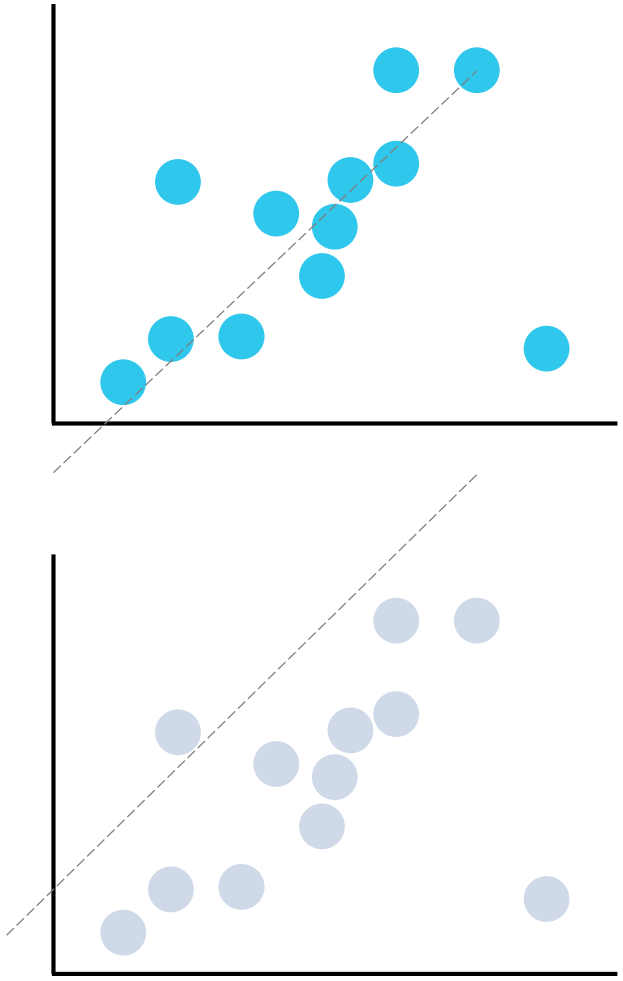
What is Cost?

Cost function

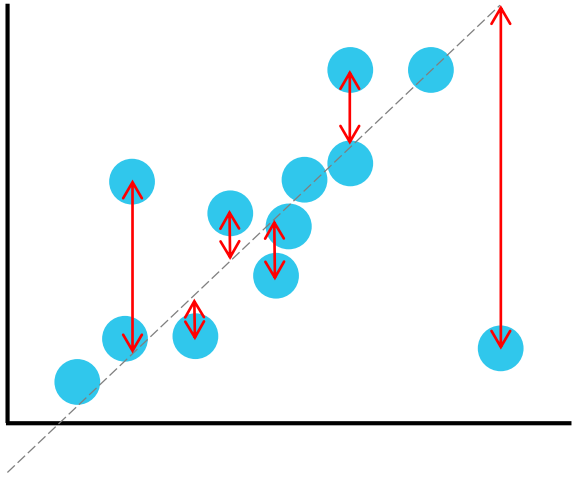
$$cost = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$

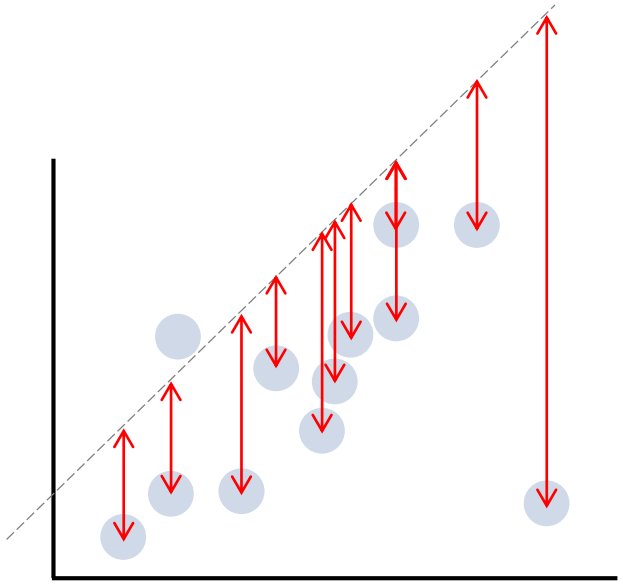
$$H(x) = Wx + b$$

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



What is Cost?





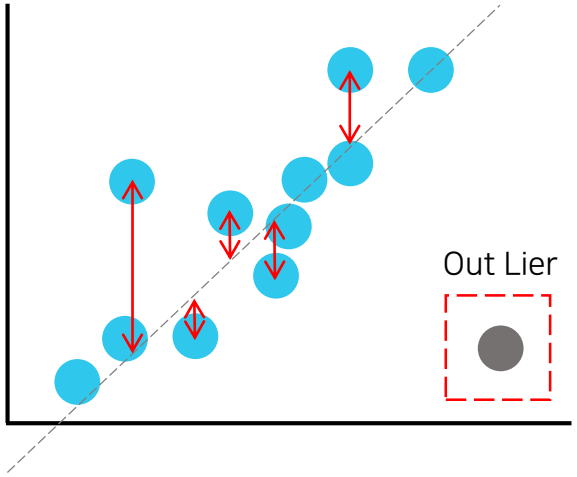
$$y = wx + b$$

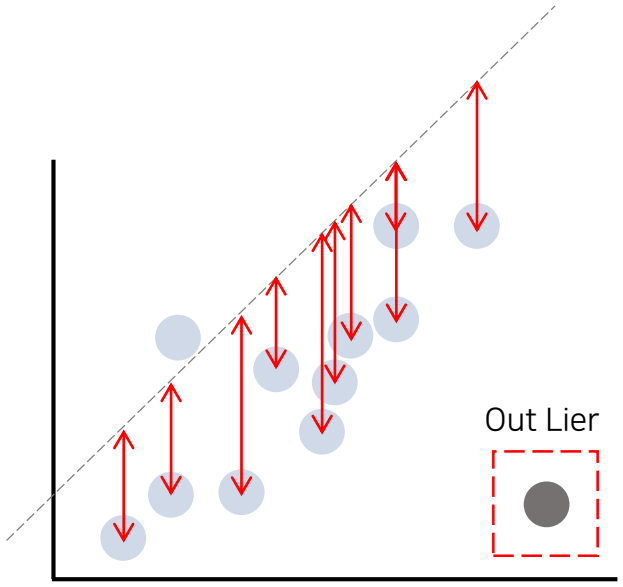
$$w : Weight$$

$$b : Bias$$

Find Minimum cost value via changing w and b

What is Cost?





$$y = wx + b$$

$$w : Weight$$

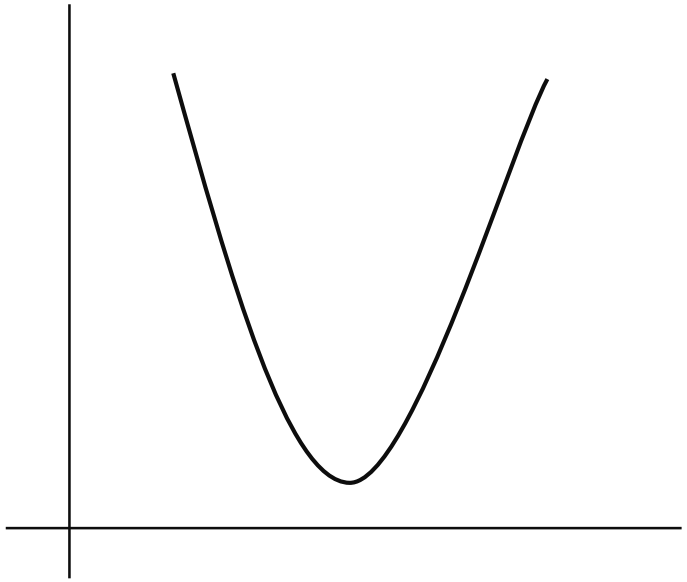
$$b : Bias$$

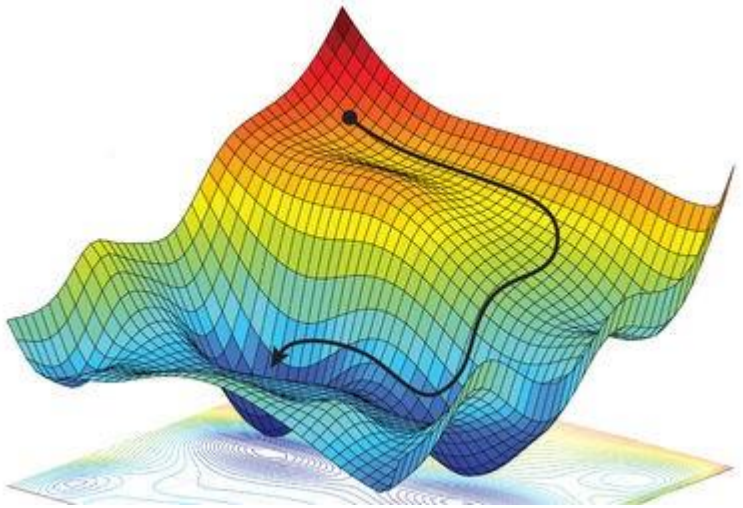
Find Minimum cost value via changing w and b

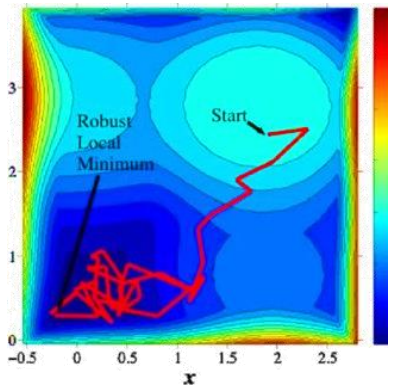
What is **Cost?**

$$y = wx + b$$

$w : Weight$
$b : Bias$



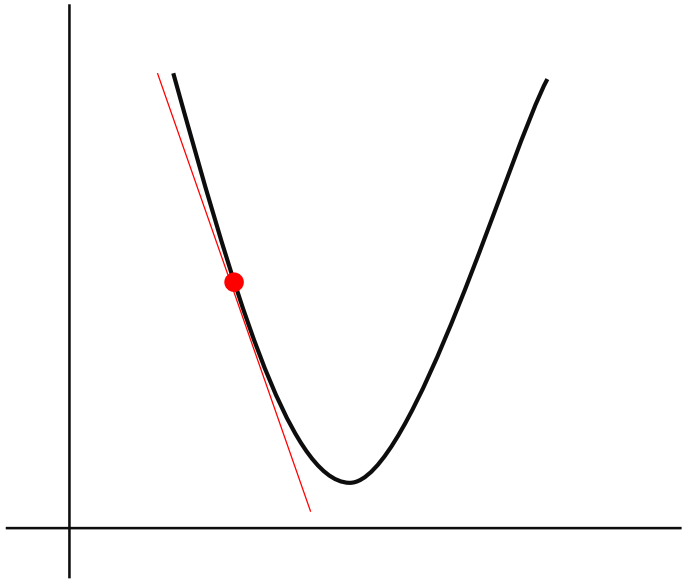


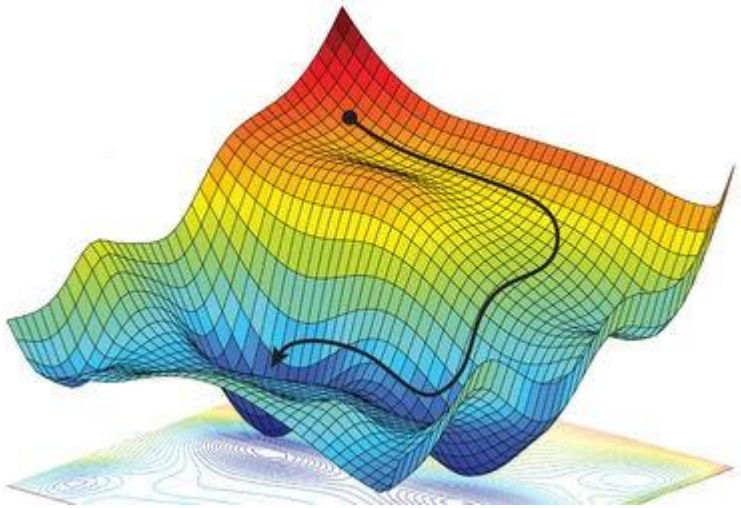


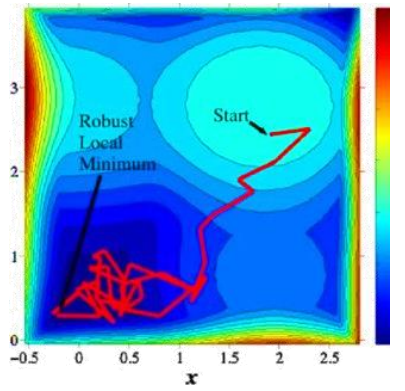
What is Cost?

$$y = wx + b$$

$w : Weight$
$b : Bias$



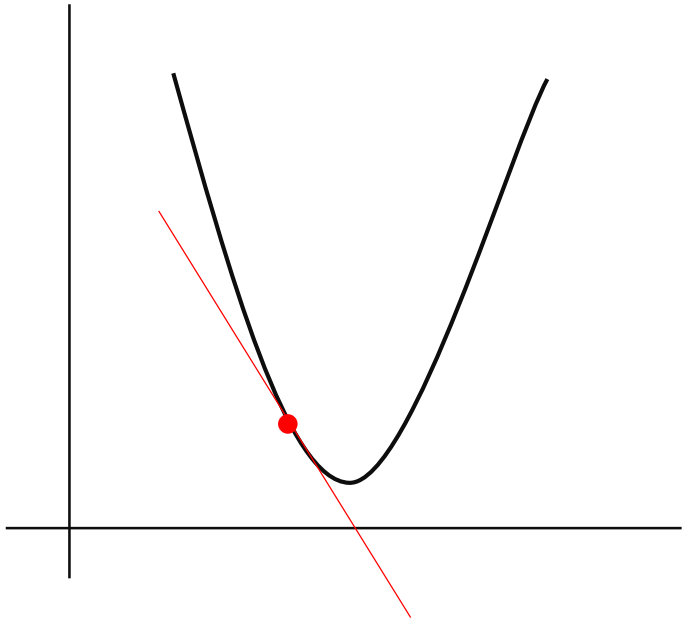


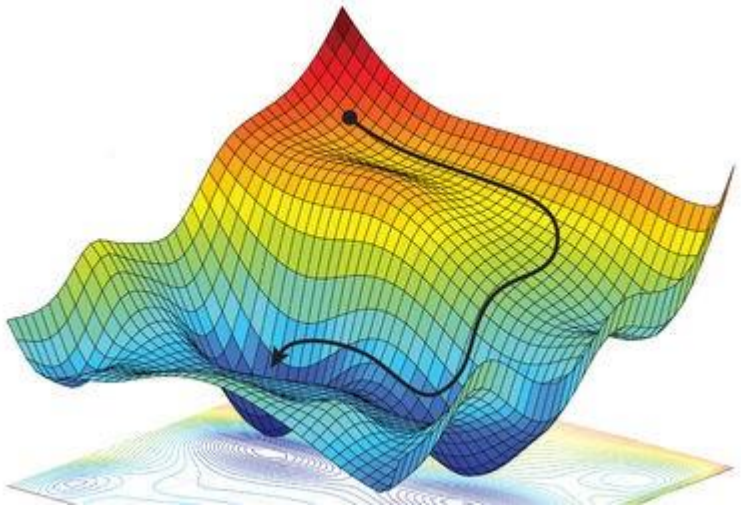


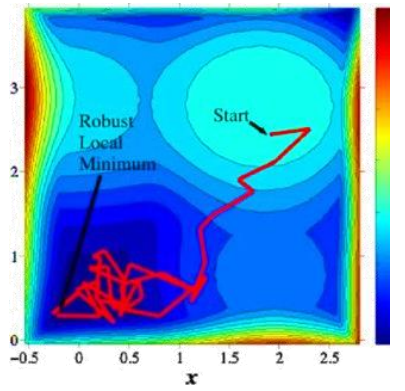
What is **Cost**?

$$y = wx + b$$

$w : Weight$
$b : Bias$



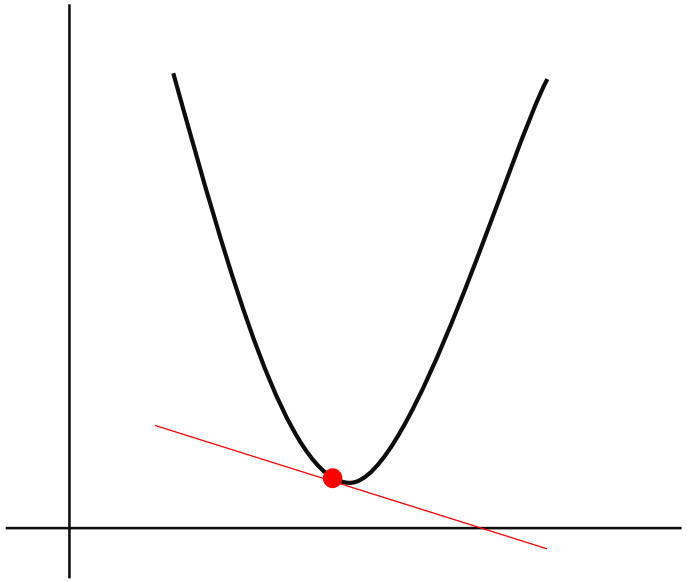


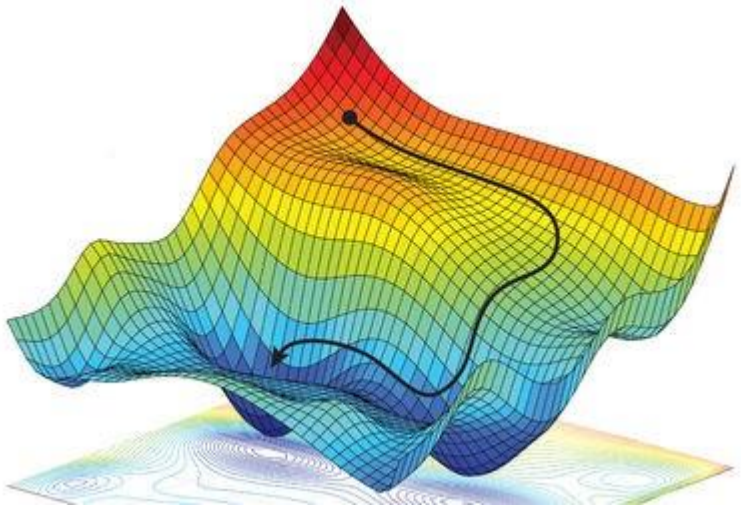


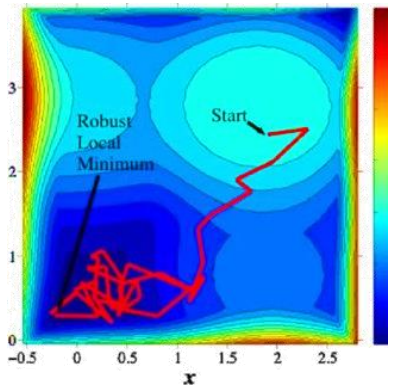
What is **Cost**?

$$y = wx + b$$

$w : Weight$
$b : Bias$



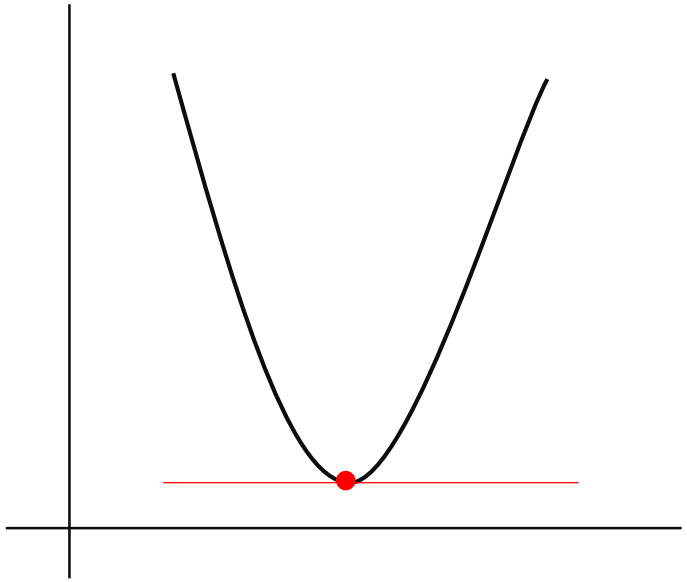


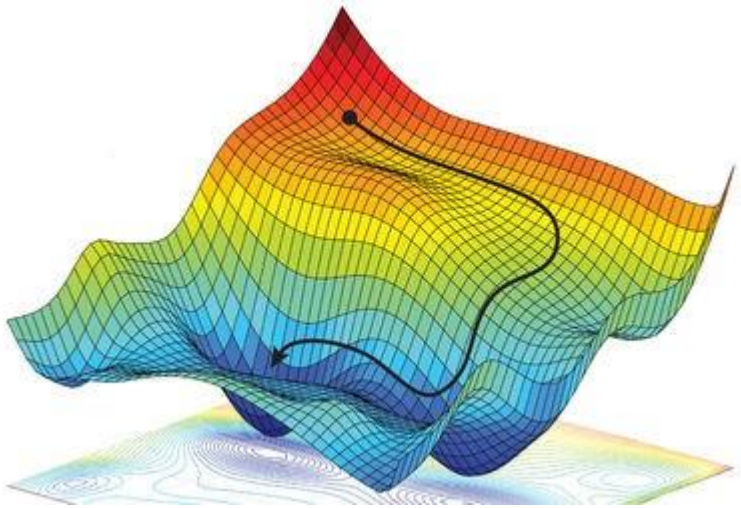


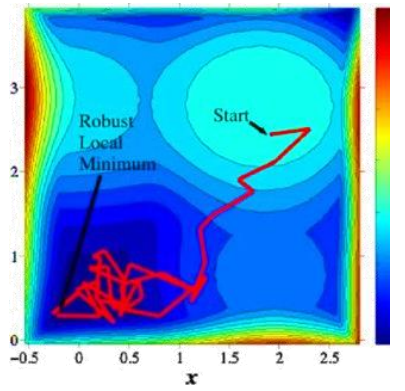
What is **Cost**?

$$y = wx + b$$

$w : Weight$
$b : Bias$



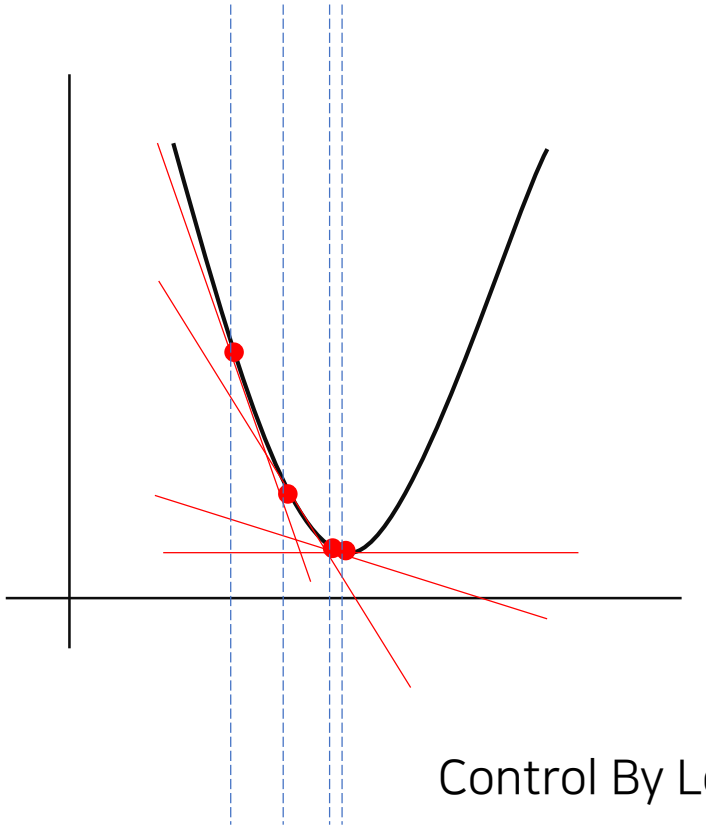




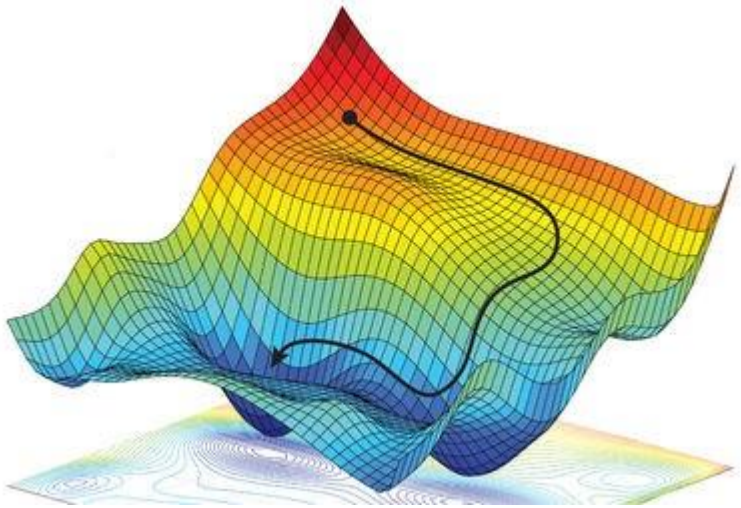
What is **Cost**?

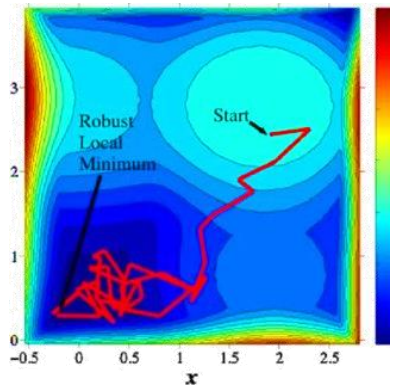
$$y = wx + b$$

$w : Weight$
$b : Bias$



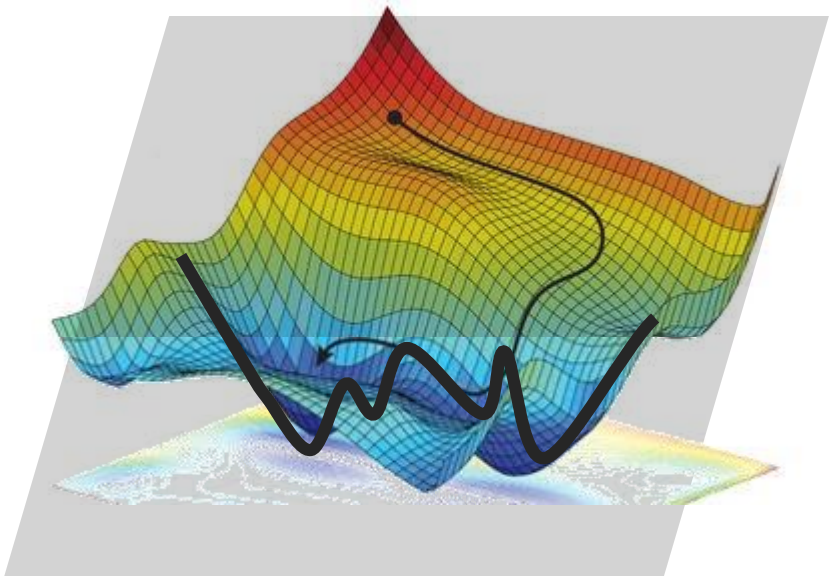
Control By Learning Rate



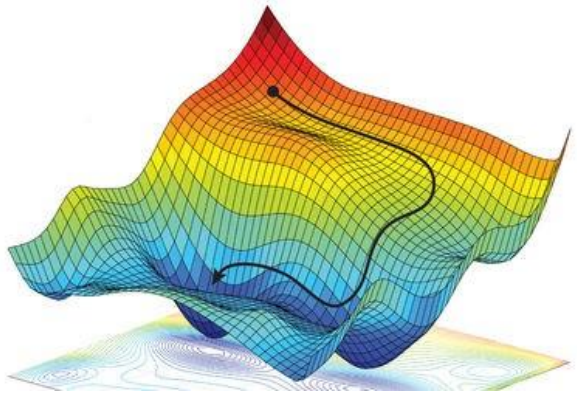


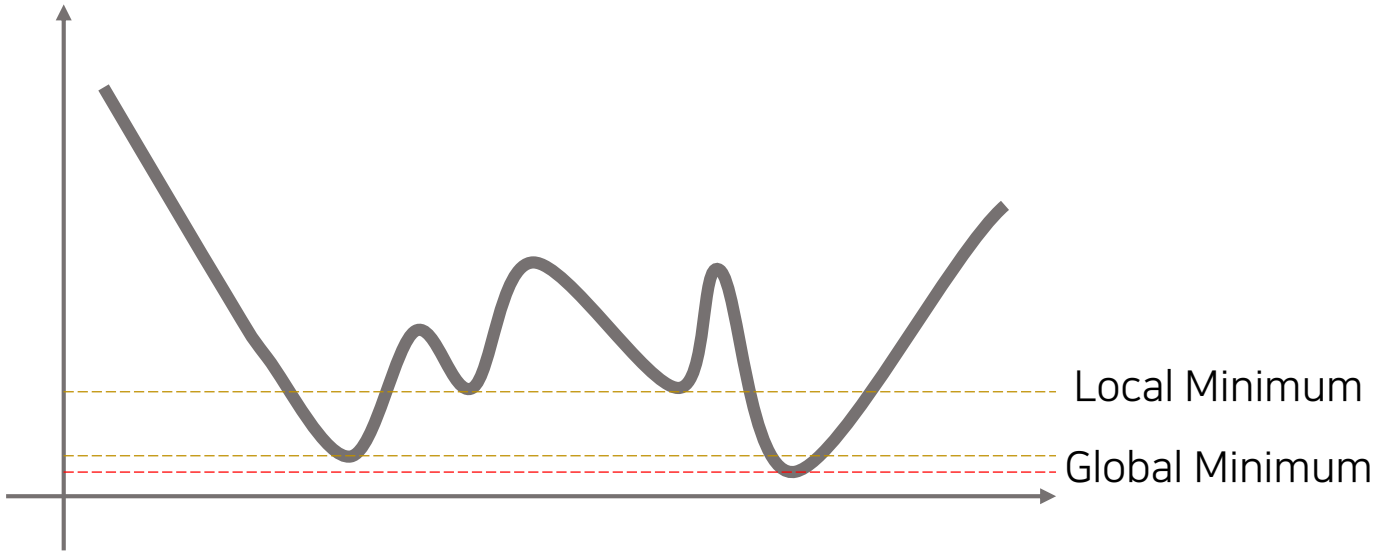
More dimension is possible
by Hyperparameters

💰 Local / Global Minima

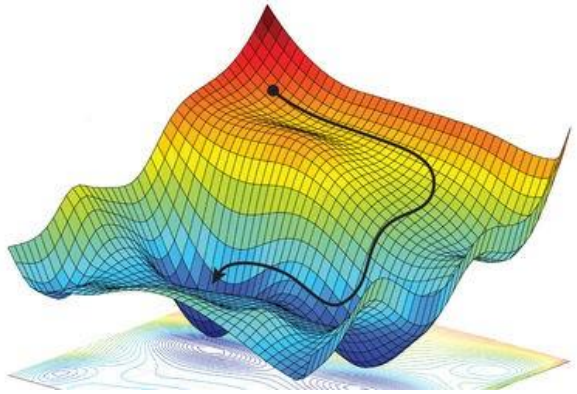


💰 Local / Global Minimum

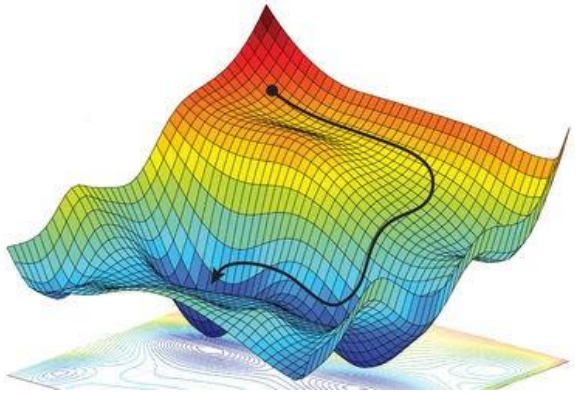




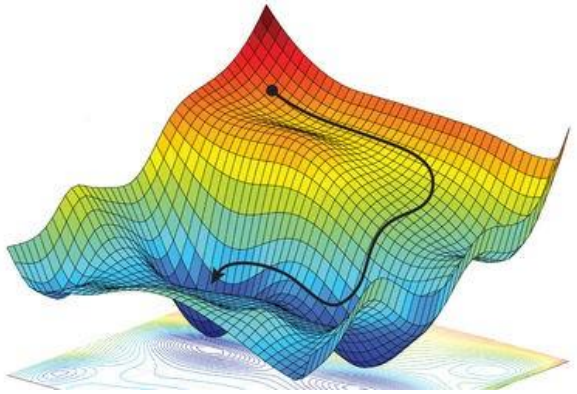
💰 Local / Global Minimum



Hyper_1 = 0
Hyper_2 = 10
Hyper_3 = 50
…



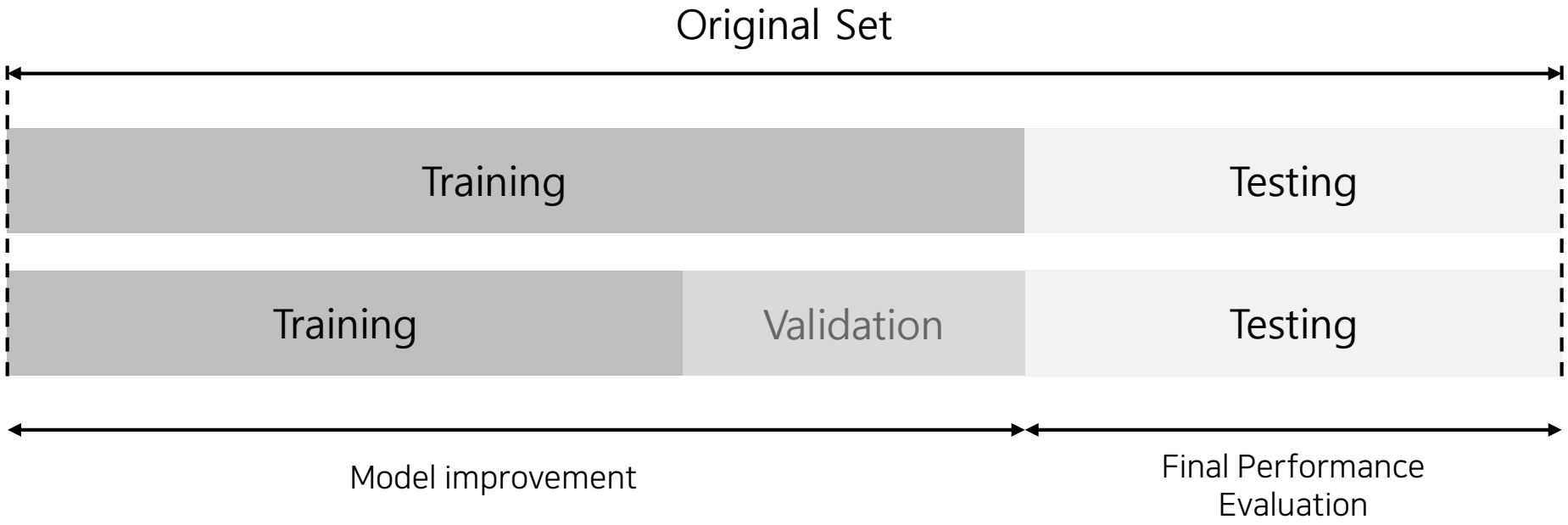
Hyper_1 = 1
Hyper_2 = 10
Hyper_3 = 50
…



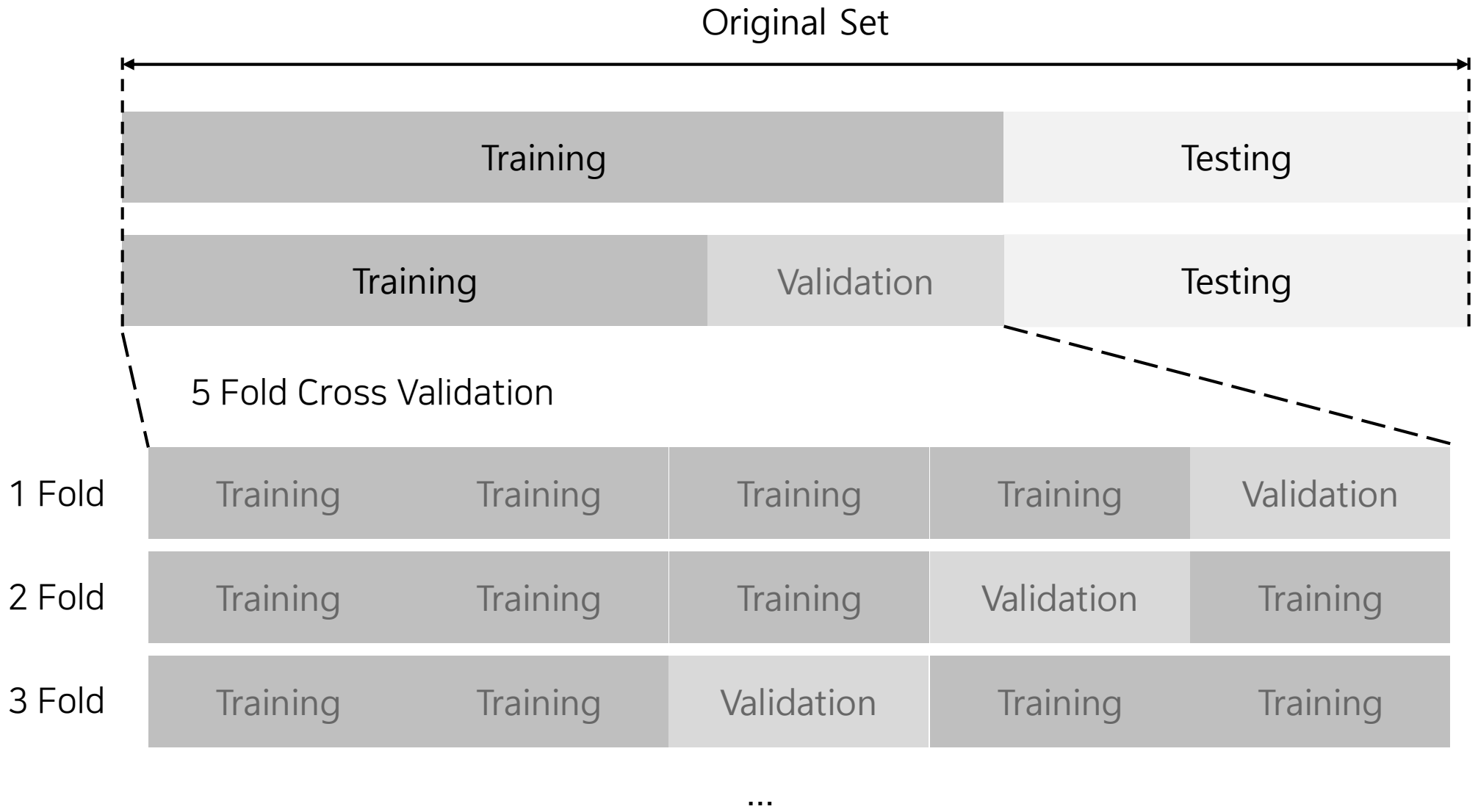
Hyper_1 = 2
Hyper_2 = 10
Hyper_3 = 50
…

Should Know!
Dimension is REALLY BIG!!

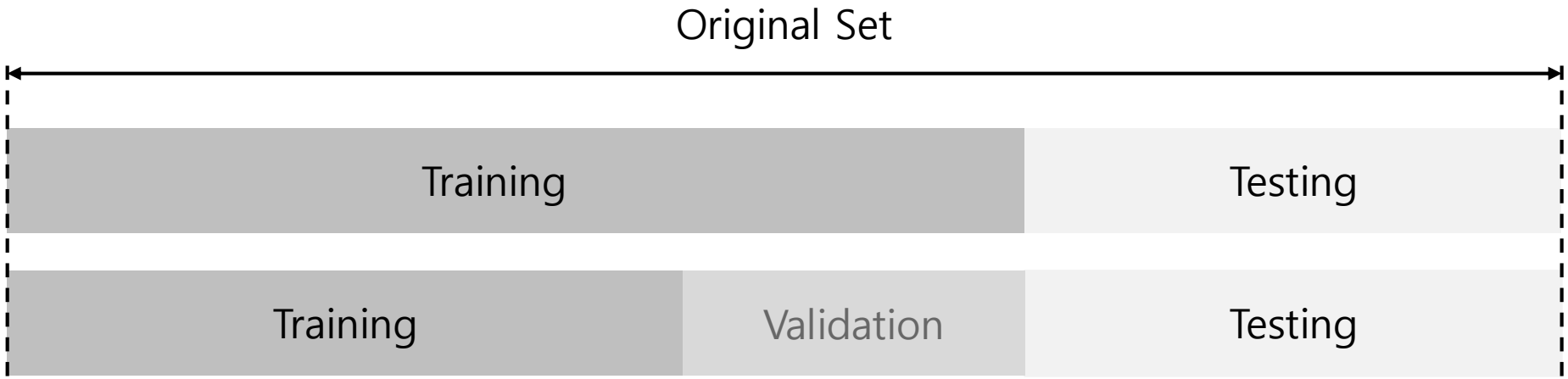
Test / Validation



Test / Validation



Test / Validation



Accuracy :

Model improvement

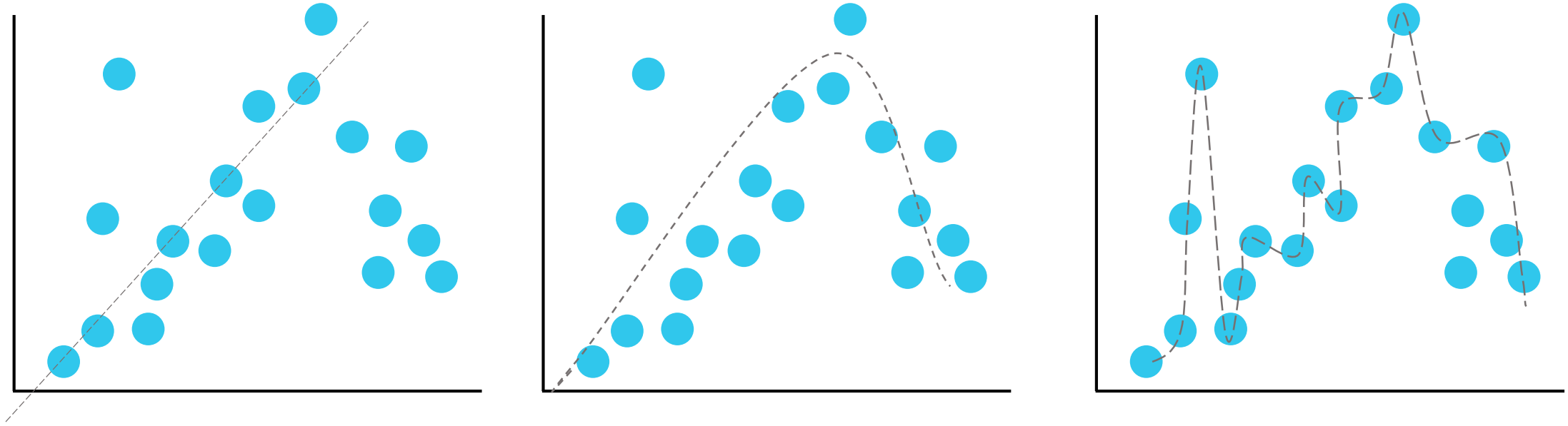
ROC Curve :

Model improvement

Confusion Matrix :

Model improvement

Over / Under Fitting

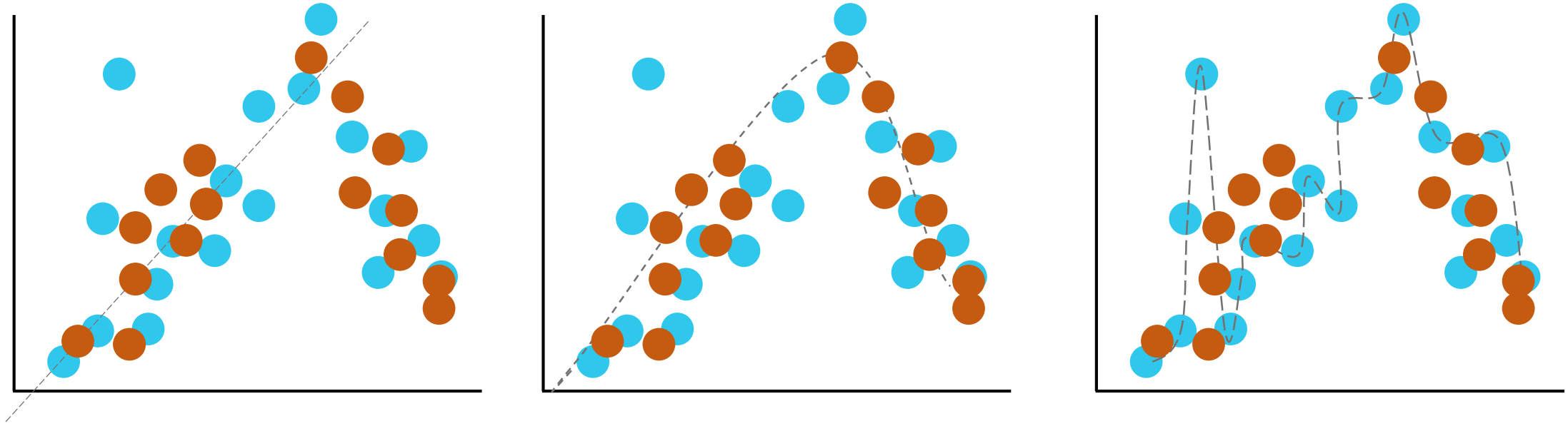


Less Cost Function

Our Main Purpose is Minimizing Cost Function.
So the Right side model is the Best Model.

Over / Under Fitting

New Inputs



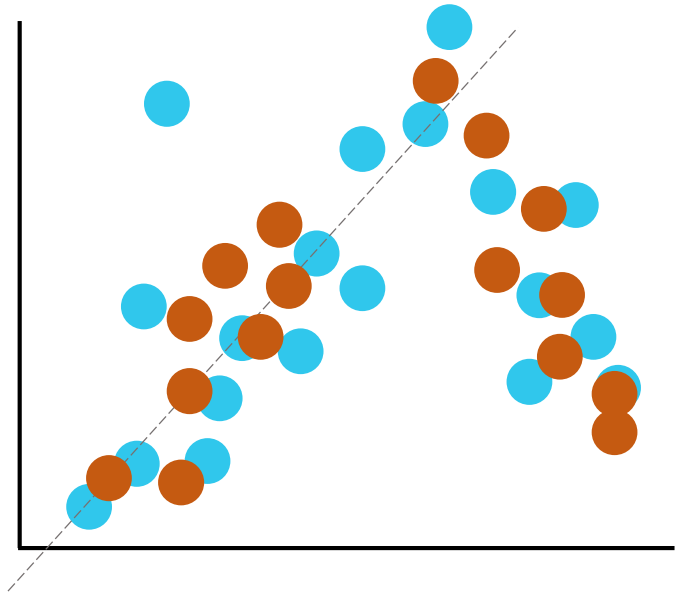
Less Cost Function

More Cost Function

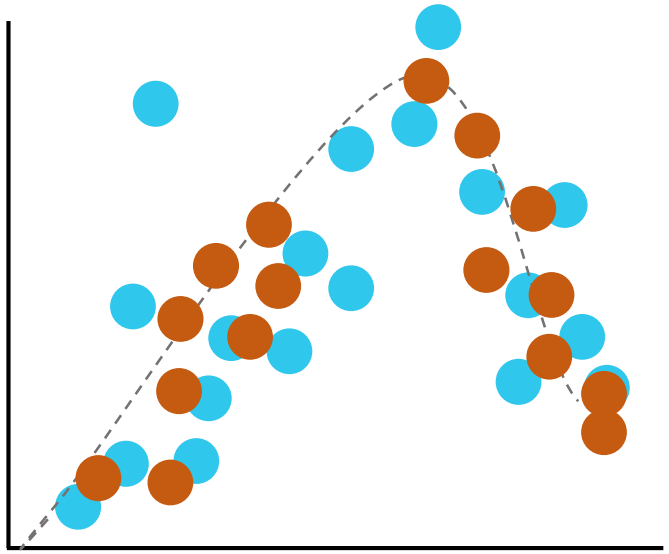
Our Main Purpose is Minimizing General Problems Cost Function.
So the Right side model is the Center Model.

Over / Under Fitting

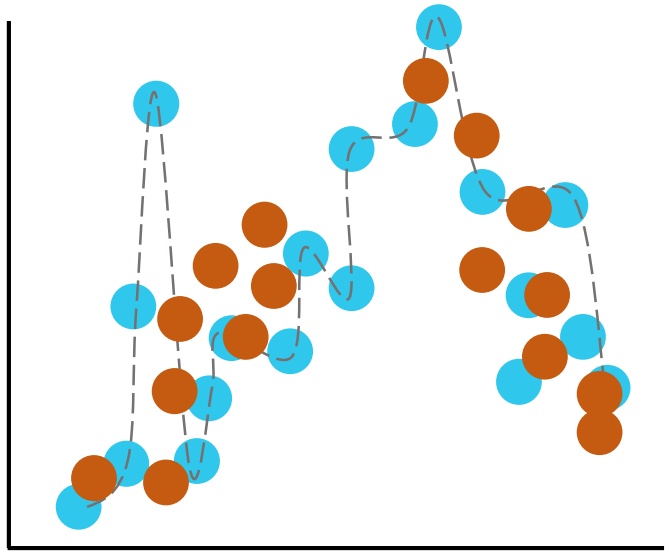
New Inputs



Under Fitting



Good Fitting



Over Fitting

(Curse of Dimension!!)

-Drop Out!

✅ Midterm Check

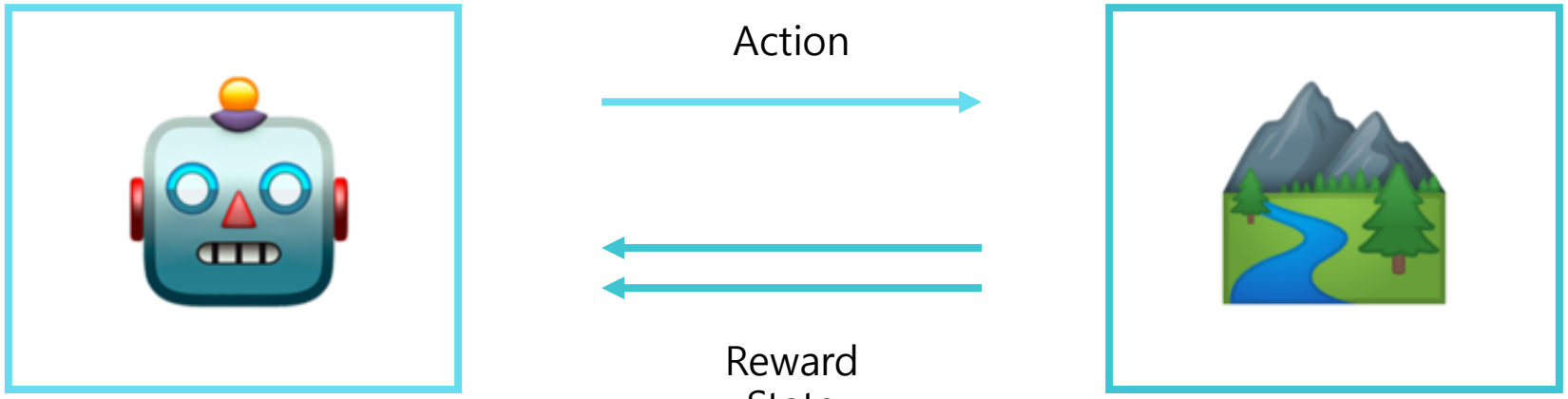
Preprocessing	Model Training	Model Evaluation
Feature extracting Data cleaning Selection	Cost function Gradient Descent Back propagation	Test / Validate set Cross Validation ROC Curve

✅ Midterm Check

Preprocessing	Model Training	Model Evaluation
Feature extracting Data cleaning Selection	Cost function Gradient Descent Back propagation	Test / Validate set Cross Validation ROC Curve

REINFORCEMENT LEARNING

Reinforcement Learning



Sequential Action Decision Problem

: 순차적 행동 결정문제

→

Dynamic Programming

: 다이나믹 프로그래밍

Reinforcement Learning

Action

Reward

Policy

State

Static / Dynamic State

State Define is important

Agent determine Action based on state

(Kind of Feature??)

Reinforcement Learning

State

Reward

Policy

Action

Agent Determine Action

Things that agent can do in specific state

Action → Reward (by Env)

Reinforcement Learning

State

Action

Policy

Reward

Key of Reinforcement Learning

Env give Reward to Agent

Action → Reward (by Env)

Reinforcement Learning

State

Action

Reward

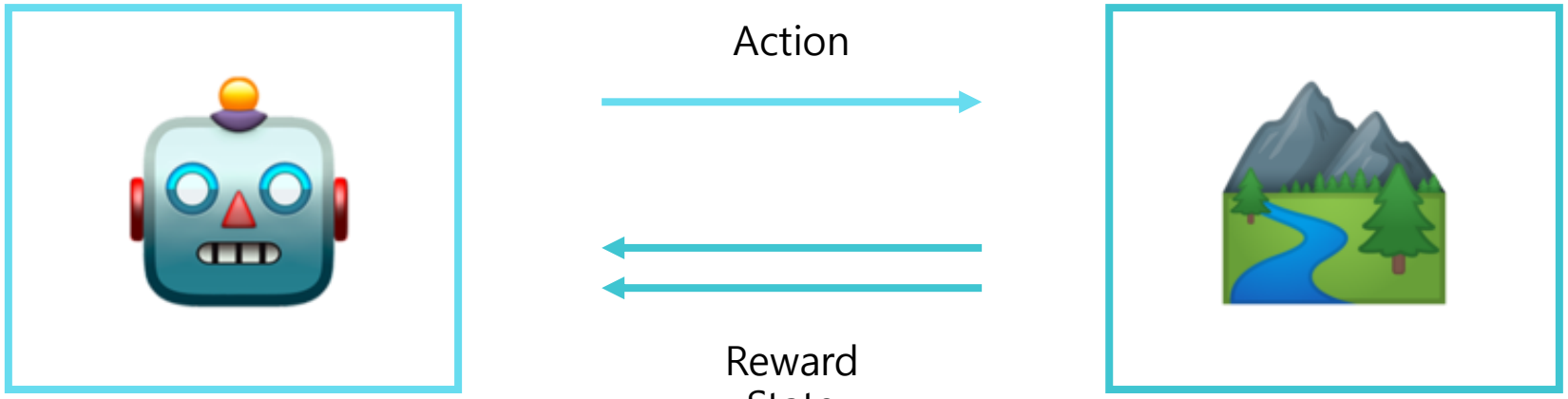
Policy

Result of Reinforcement Learning

Set of Action Rule for Agent

(Book of Action based on State)

Reinforcement Learning



Sequential Action Decision Problem

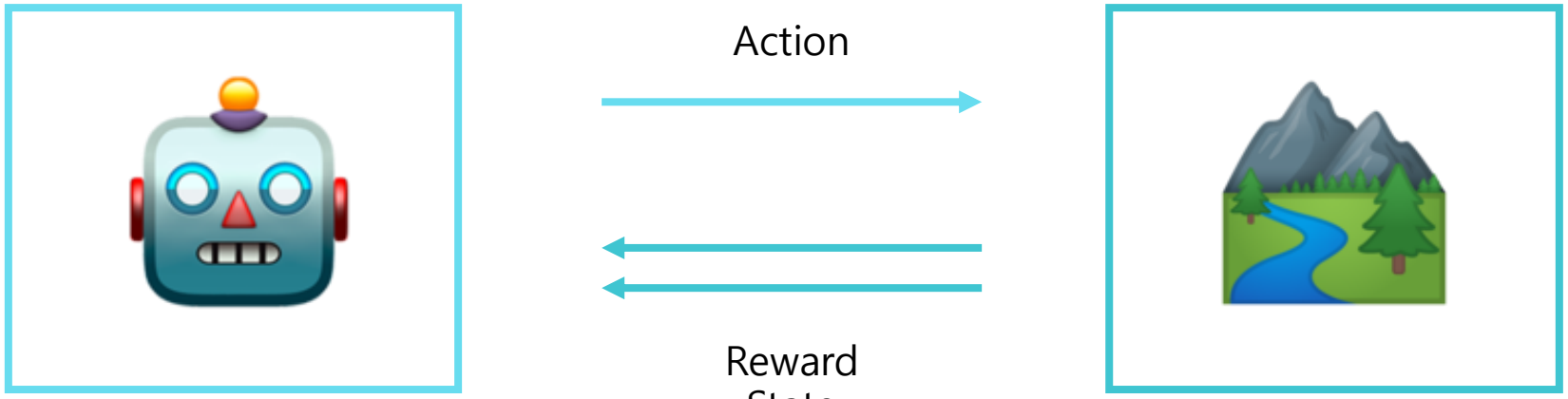
: 순차적 행동 결정문제

→

Dynamic Programming

: 다이나믹 프로그래밍

Reinforcement Learning



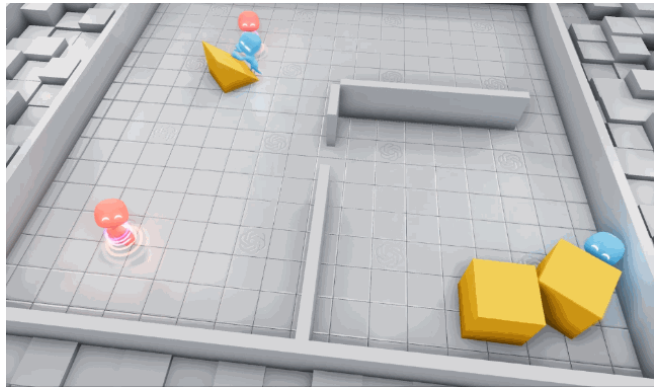
Sequential Action Decision Problem

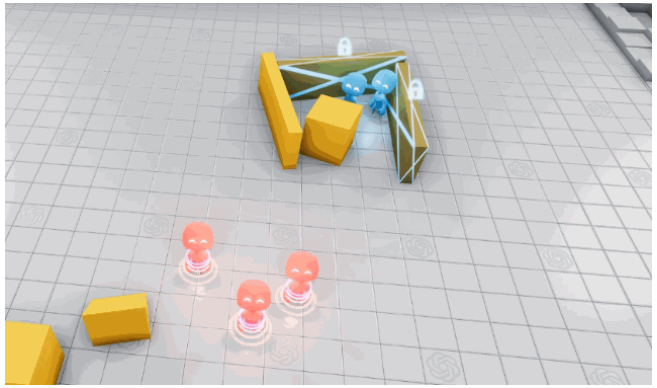
: 순차적 행동 결정문제

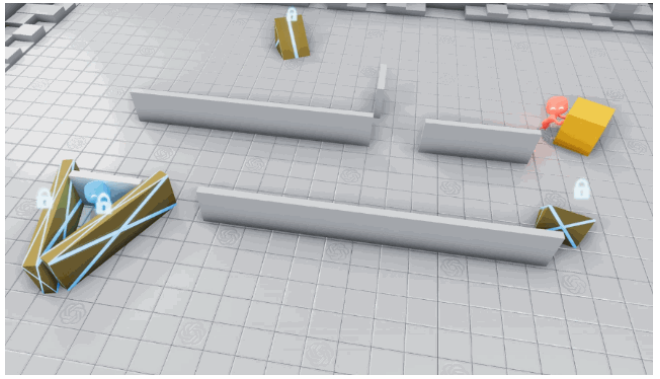
Define MDP → Calculating Bellman Equation → Get optimized Value / Policy function

Reinforcement Learning

OpenAI



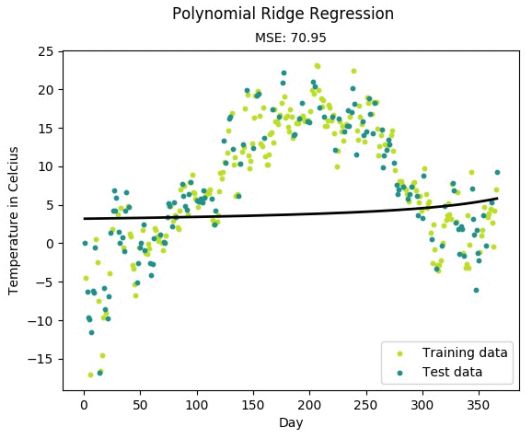




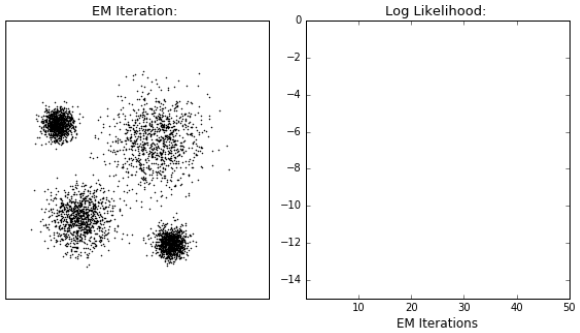
Guess [Action] [Reward] [State] !

✅ Over View

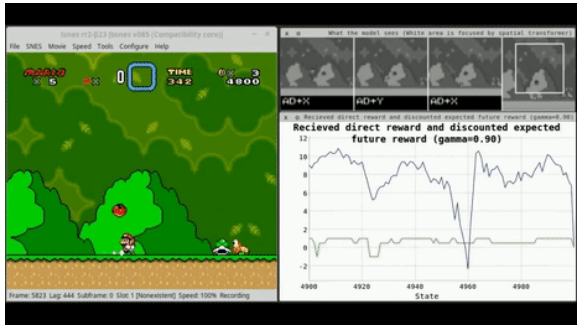
Supervised



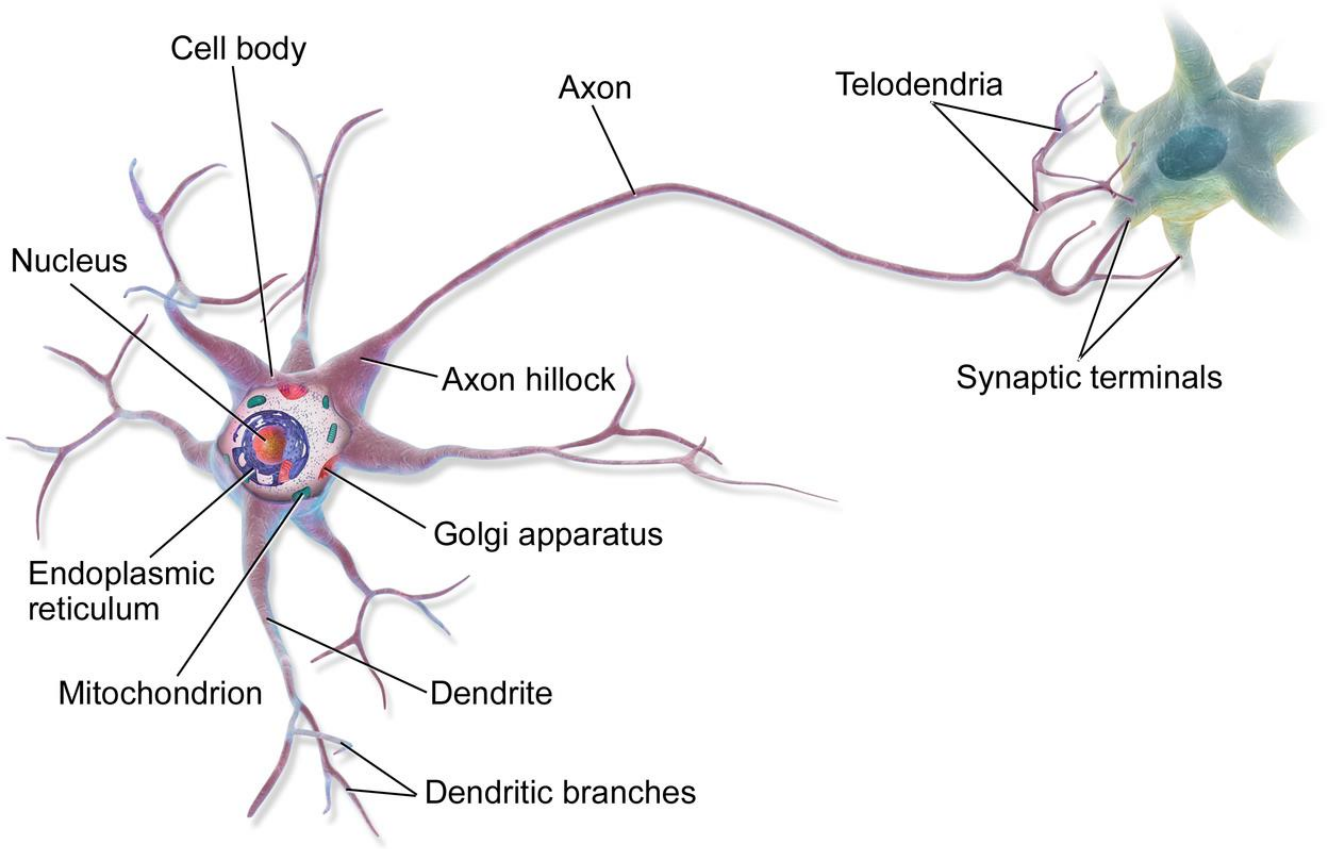
Unsupervised



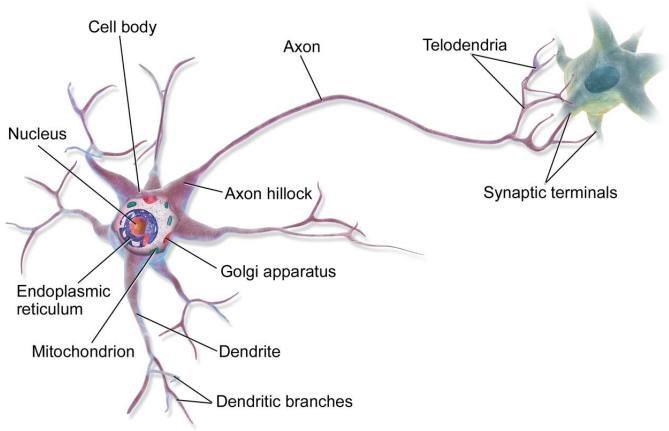
Reinforcement

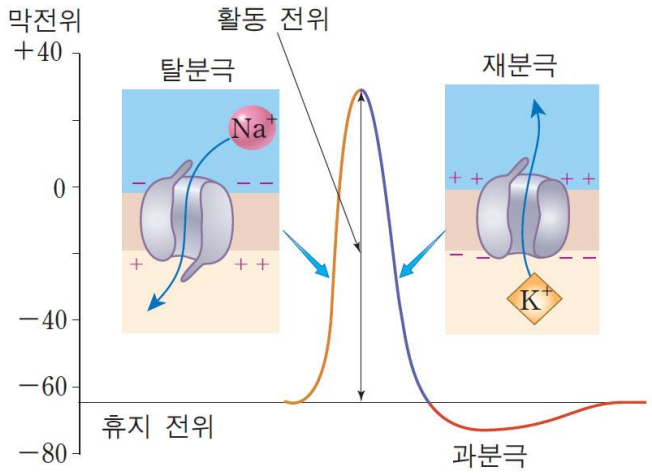


Neuro Morphic



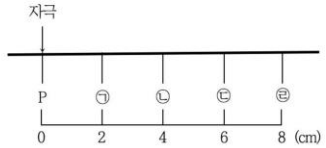
Neuro Morphic





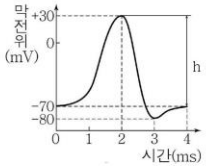
[20. 자극의 전달 #2]

20. 그림은 어떤 민말이집 신경 X의 축삭돌기 일부와 P지점으로부터 ㉠~㉣까지의 거리를, 표는 이 축삭돌기의 지점 P에 역치 이상의 자극을 1회 주고 일정 시간이 지난 후 시점 t_1과 t_2일 때 각각 네 지점 ㉠~㉣에서 측정한 막전위를 나타낸 것이다. (가)~(라)는 ㉠~㉣ 중 하나이다.



시점	측정한 막전위(mV)			
	(가)	(나)	(다)	(라)
t_1	-80	+20	+15	-75
t_2	+30	-40	+10	-20

○ X의 ㉠~㉣에서 활동 전위가 발생했을 때, 각 지점에서의 막전위 변화는 다음 그래프와 같다.

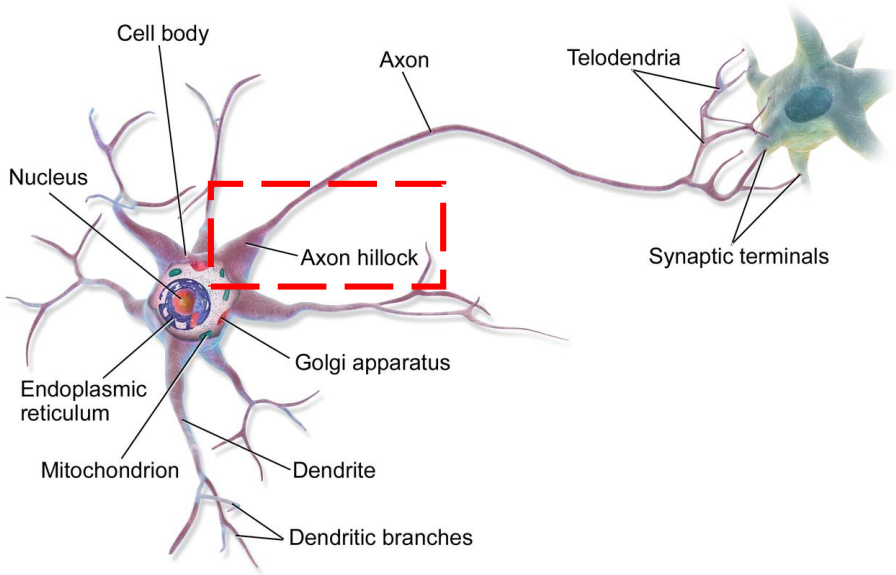


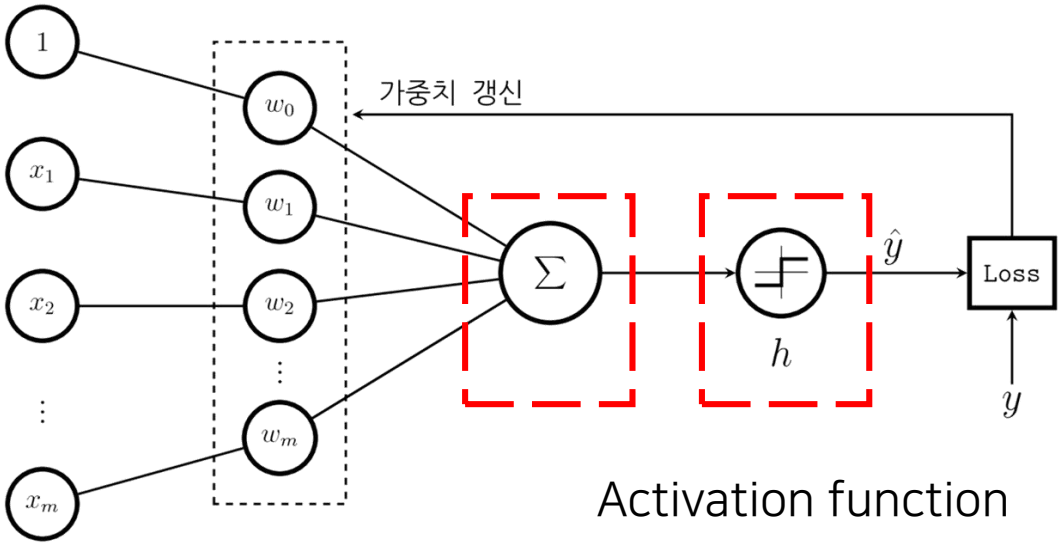
이에 대한 설명으로 옳은 것만을 〈보기〉에서 있는 대로 고른 것은? (단, 흥분의 전도는 1회만 일어났고, 휴지 전위는 −70mV이다.)

〈보 기〉

ㄱ. t_1은 t_2보다 빠른 시점이다.
ㄴ. (나)는 ㉣이다.
ㄷ. X에서 흥분의 전도 속도가 2cm/ms이면, t_1은 5ms이다.
ㄹ. X에서 흥분의 전도 속도가 2cm/ms이고 자극을 주고 경과된 시간이 6ms일 때, ㉠에서 세포막을 통한 Na^+의 이동에 ATP가 사용된다.
ㅁ. t_1일 때 ㉢에서 탈분극이 일어나고 있다.
ㅂ. t_2일 때 ㉠에서 K^+이 K^+ 통로를 통해 세포 밖으로 유출된다.
ㅅ. Na^+의 막 투과도는 t_1일 때 ㉡에서가 t_2일 때 ㉢에서보다 높다.
ㅇ. t_1일 때 $\frac{|㉡\text{에서의 막전위}|}{|㉠\text{에서의 막전위}|}$는 1보다 크다.
ㅈ. t_2일 때 ㉣에서 세포막 안쪽이 양(+)전하를 띤다.
ㅊ. 이 자극보다 세기가 큰 자극을 주면 h값이 커진다.

Neuro Morphic



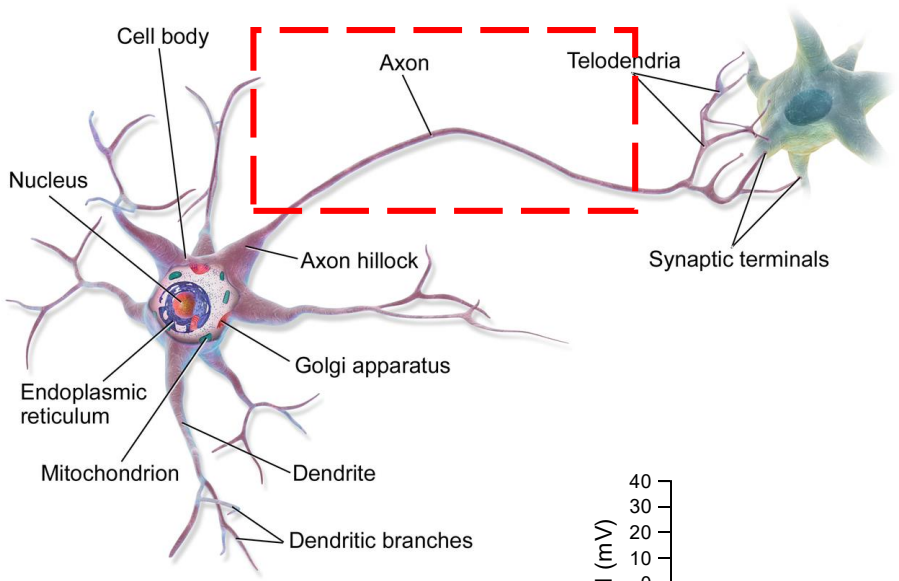


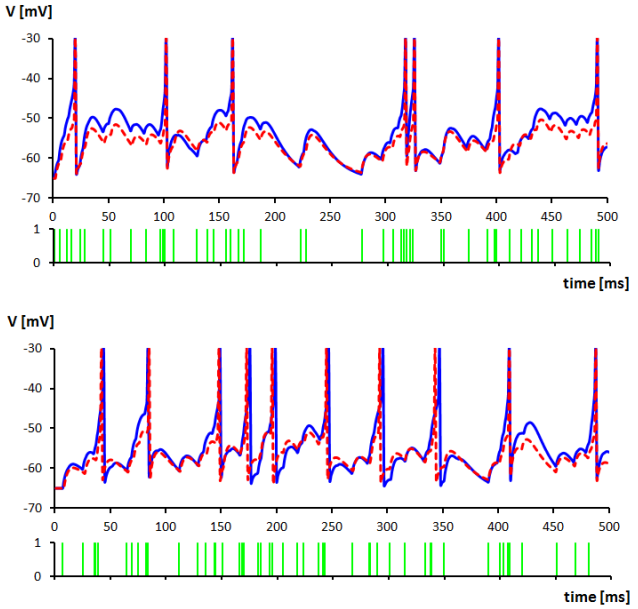
Activation function

LTP: Long-term potentiation → Weight Update

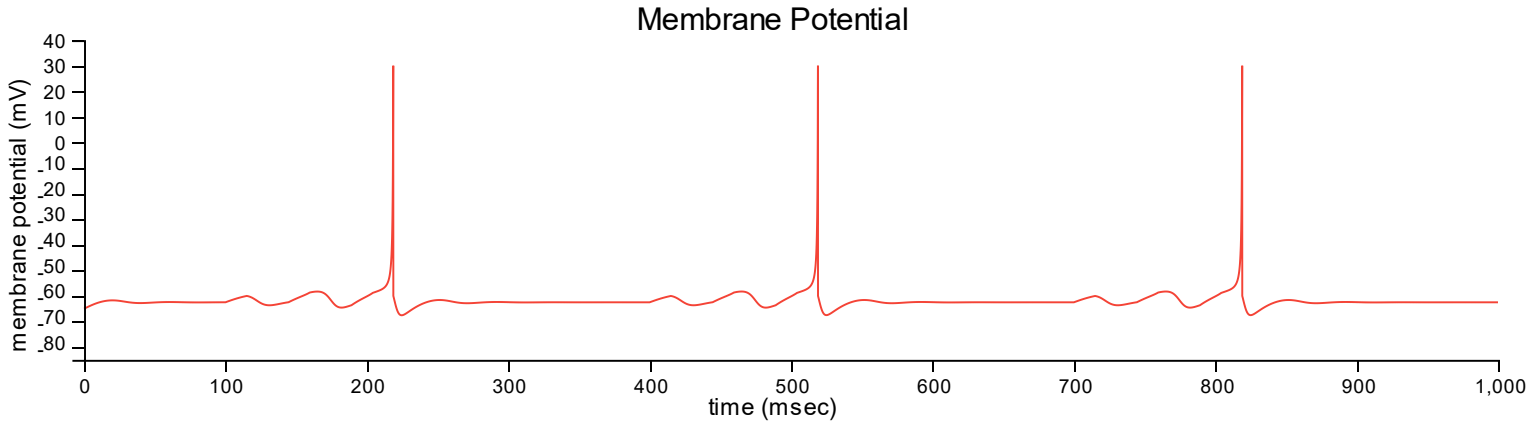
Back-Propagation!!

Neuro Morphic

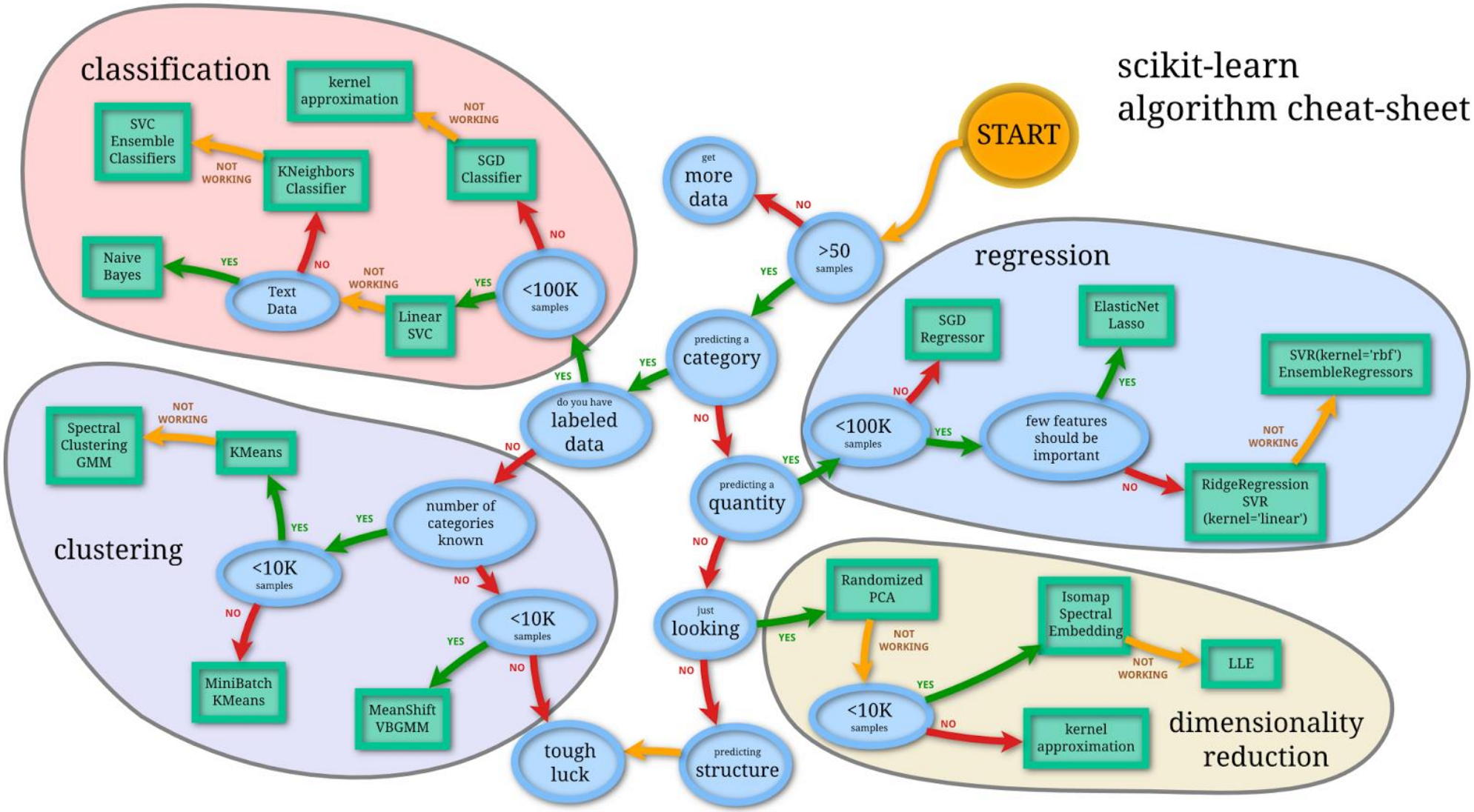




Spiking Neural Net
Evolvable Neural Unit
...

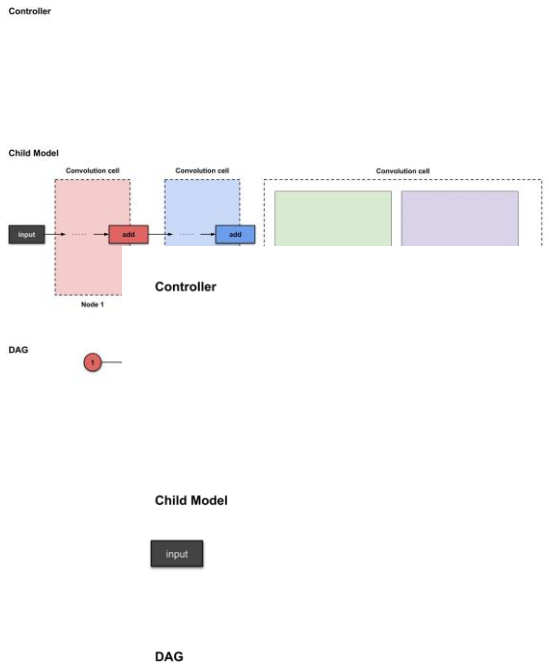


And More…


scikit-learn
algorithm cheat-sheet
START
classification
kernel approximation
NOT WORKING
SVC Ensemble Classifiers
NOT WORKING
KNeighbors Classifier
SGD Classifier
NO
Naive Bayes
YES
NO
NOT WORKING
YES
Text Data
NOT WORKING
Linear SVC
YES
<100K samples
NO
get more data
NO
>50 samples
YES
regression
SGD Regressor
ElasticNet Lasso
SVR(kernel='rbf') EnsembleRegressors
predicting a category
YES
NO
<100K samples
NO
YES
few features should be important
YES
NOT WORKING
NO
RidgeRegression SVR (kernel='linear')
do you have labeled data
YES
NO
predicting a quantity
YES
NO
clustering
Spectral Clustering GMM
NOT WORKING
KMeans
number of categories known
YES
YES
<10K samples
NO
NO
<10K samples
YES
NO
MiniBatch KMeans
MeanShift VBGMM
just looking
YES
NO
Randomized PCA
NOT WORKING
Isomap Spectral Embedding
NOT WORKING
LLE
YES
<10K samples
NO
kernel approximation
dimensionality reduction
tough luck
predicting structure

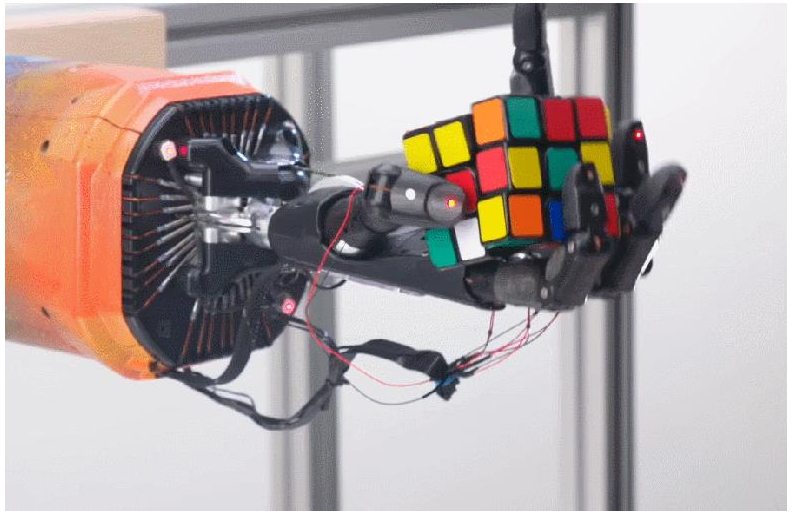
More…

GAN

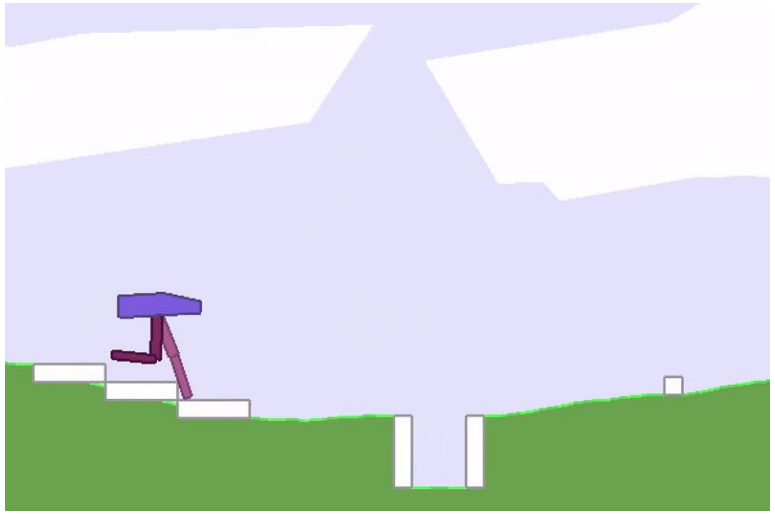

Controller
Child Model
Convolution cell
Convolution cell
Convolution cell
input
add
add
Node 1
Controller
DAG
Child Model
input
DAG

NAS

More…



Robotics



EA

THANKS

BAEK SU WHAN (백수환)

Email. zhsjzhsj@gmail.com

Phone. +82 010 2447 5811

KwangWoon Univ Computer Engineering

BML Lab.