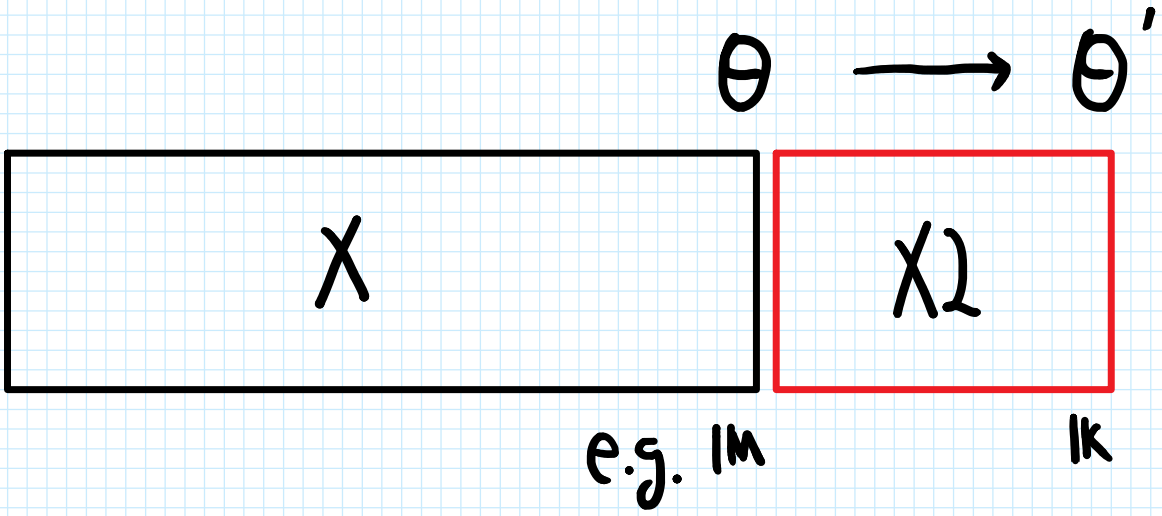
Day 4

2018년 12월 27일 목요일 오후 5:14



.fit(X)

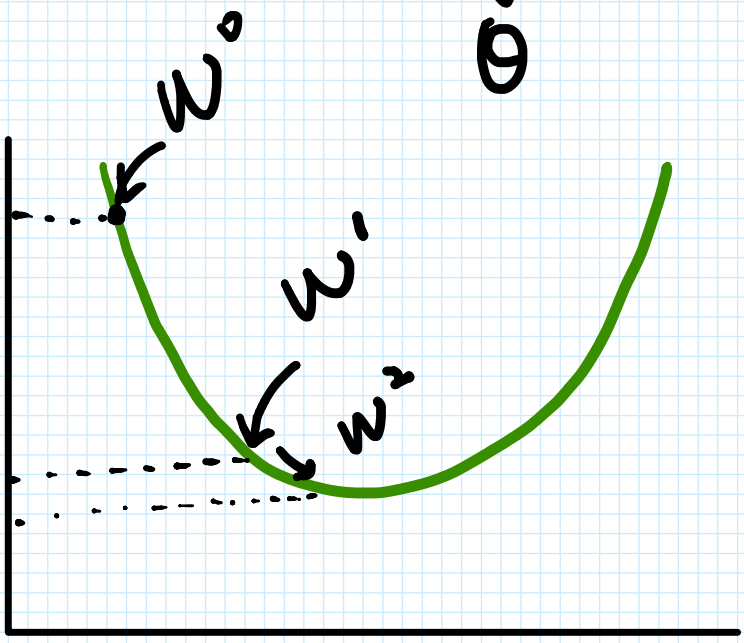
↓

θ

.fit(X+X2)

.partial-fit(X2)

배치 단위



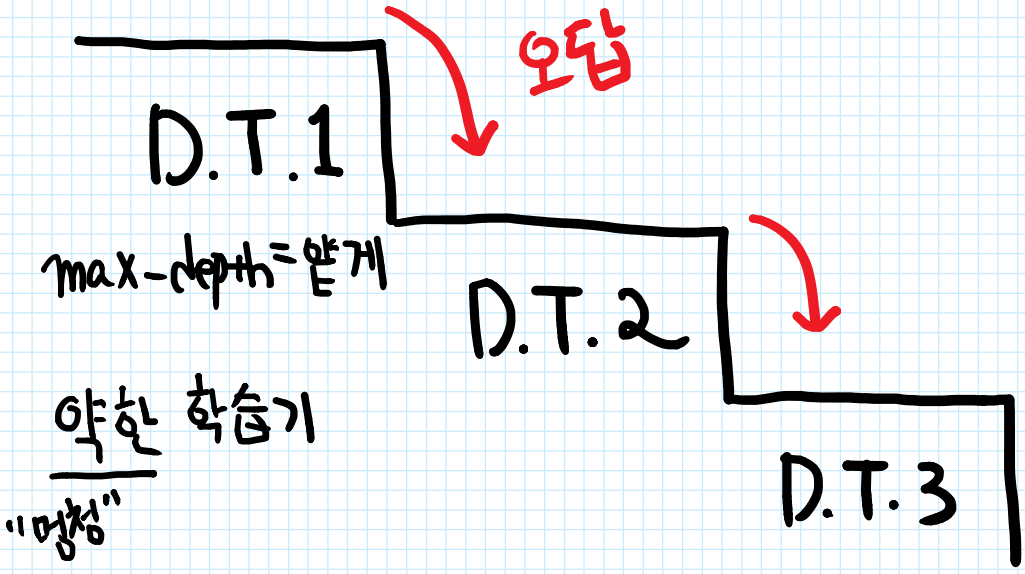
e.g. SGD ⟨ Clf. / Rg.

앙상블 모델 (이어서) Sklearn.ensemble

1) Random Forest(n_estimators)

2) G.B.R.T

Adaptive Boost
"오답 강조"



약한 학습기
"멍청"

Sklearn.ensemble

GradientBoostingClassifier(. . .

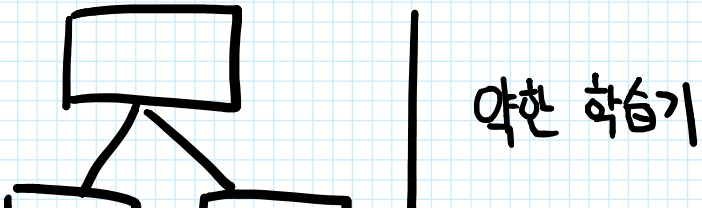
n_estimators = → 과소

max_depth=5,

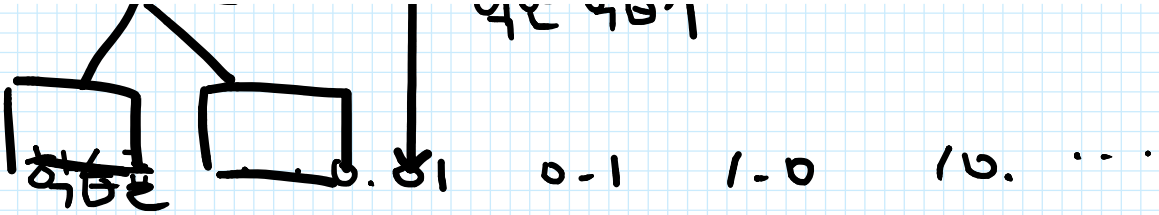
learning_rate=0.1, . . .)

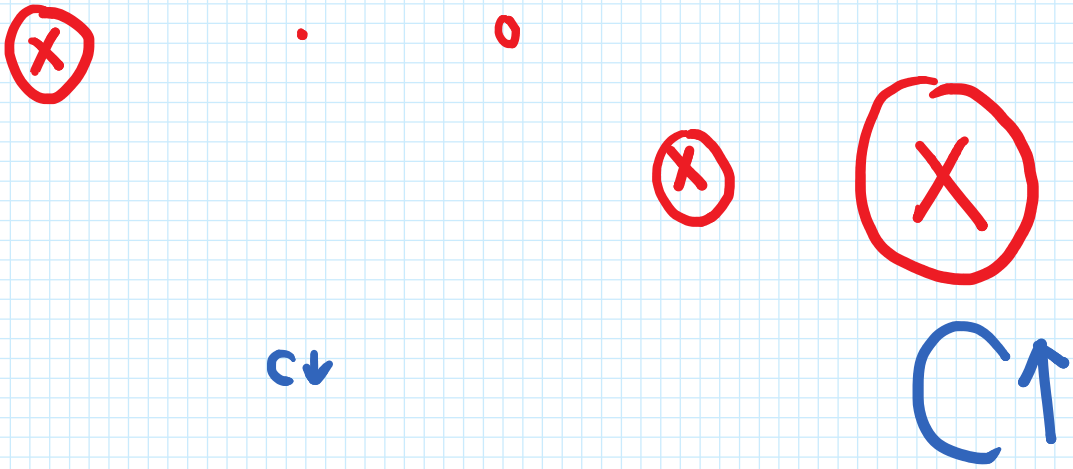
깊이

와차강조



약한 학습기





그리드 탐색

max_depth \ LR	0.01	0.1	1.	10
1				
2				
3				
4				
5				

sklearn.model_selection.

GridSearchCV(

GridSearchCV(

estimator = model( )

param_grid= 변수조합

CV = None

).fit(X, y)

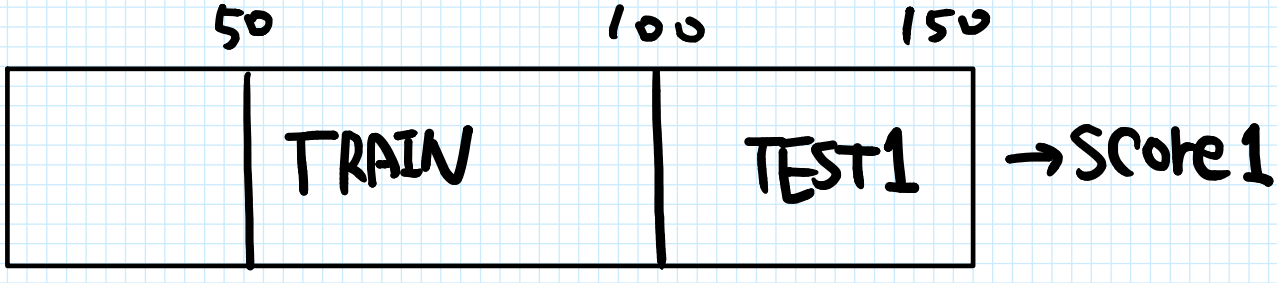
.best_params_

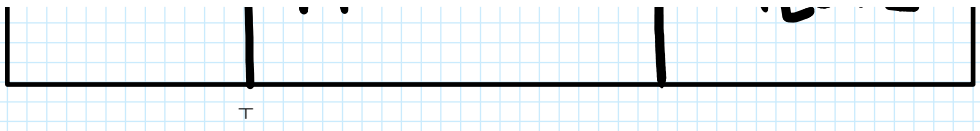
.best_estimator_

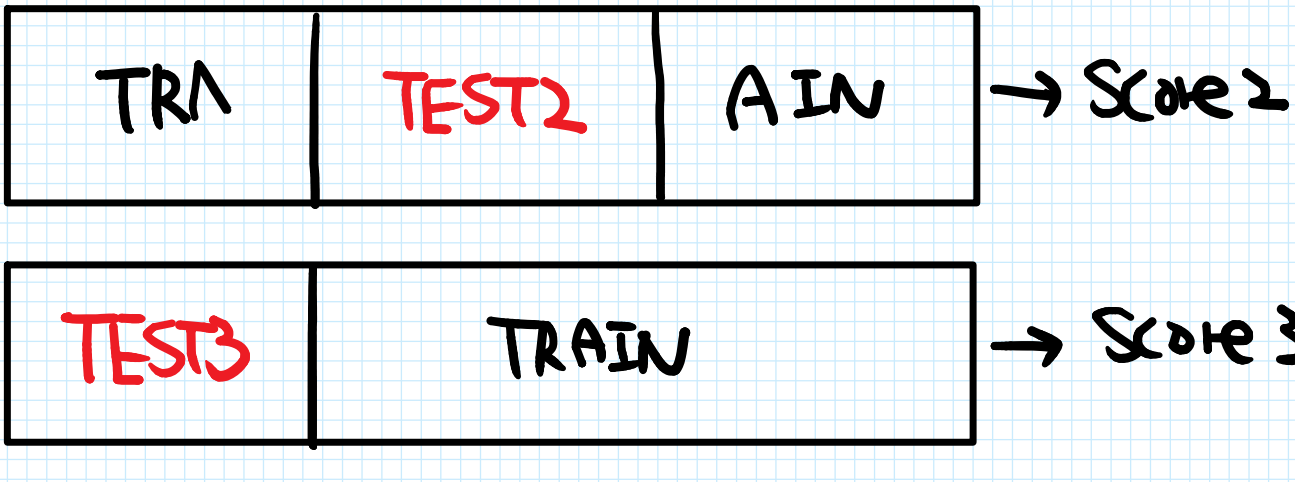
.best_score_

교차검증 C.V.

Cross Validation







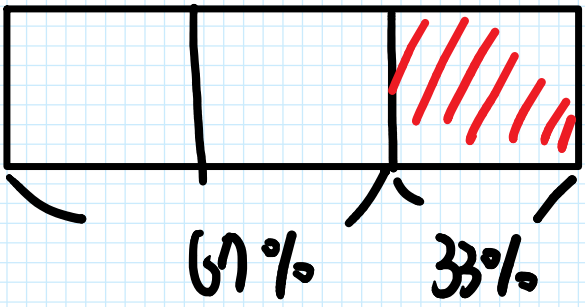
평균 → Score

모델 평가

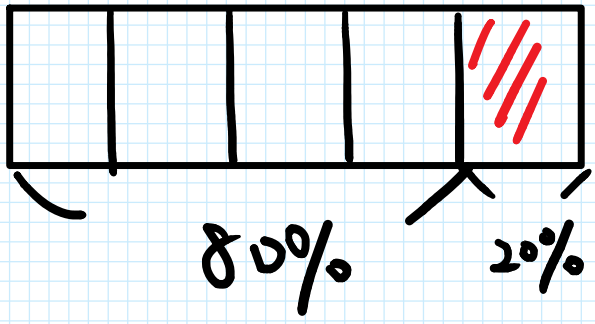
Sklearn.model_selection.

Cross_val_score(CV=3

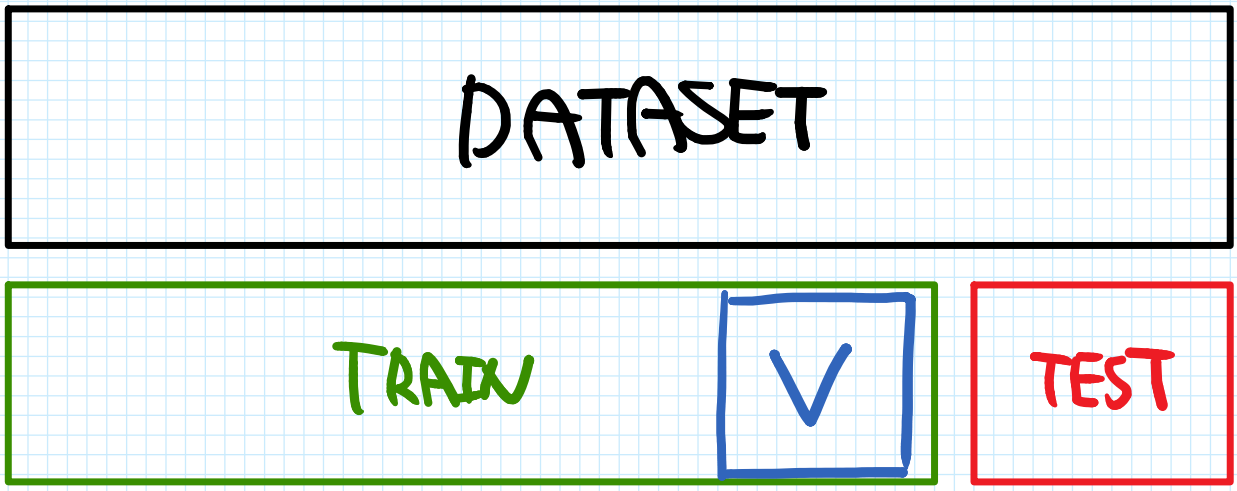
CV=3



CV=5

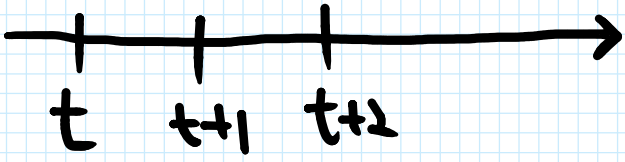


확인 데이터 ≠ 샘플 데이터



Q. AlphaGo

Q. 시계열 데이터?



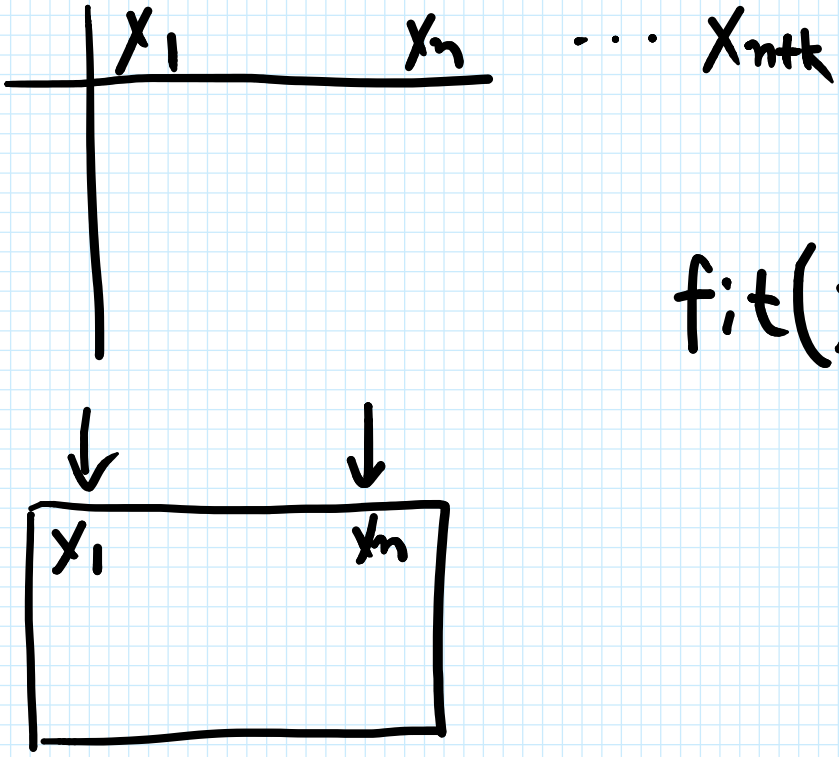
Recurrent Neural Network

Q. 학습을 "설명" 하기

"Explainable AI"

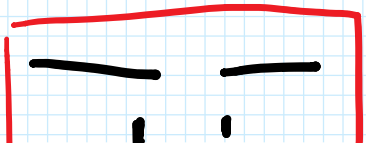
직관적 지능

Q. 특징 변화시?

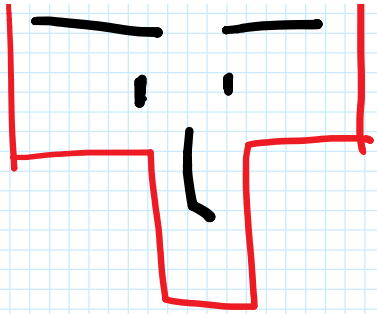


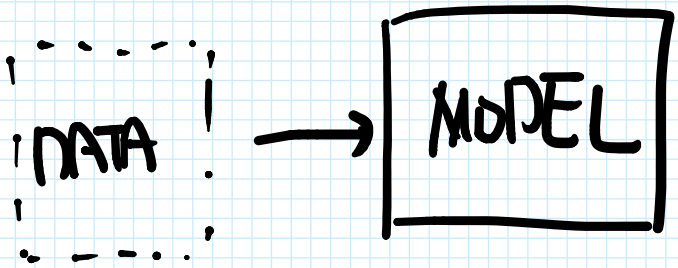
$fit(X')$

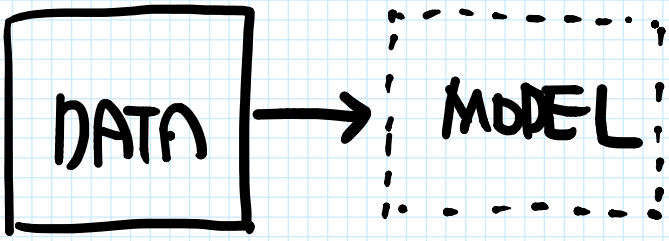
Q. 특징 추출



$X \rightarrow T(X)$

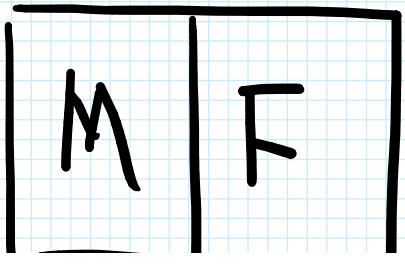






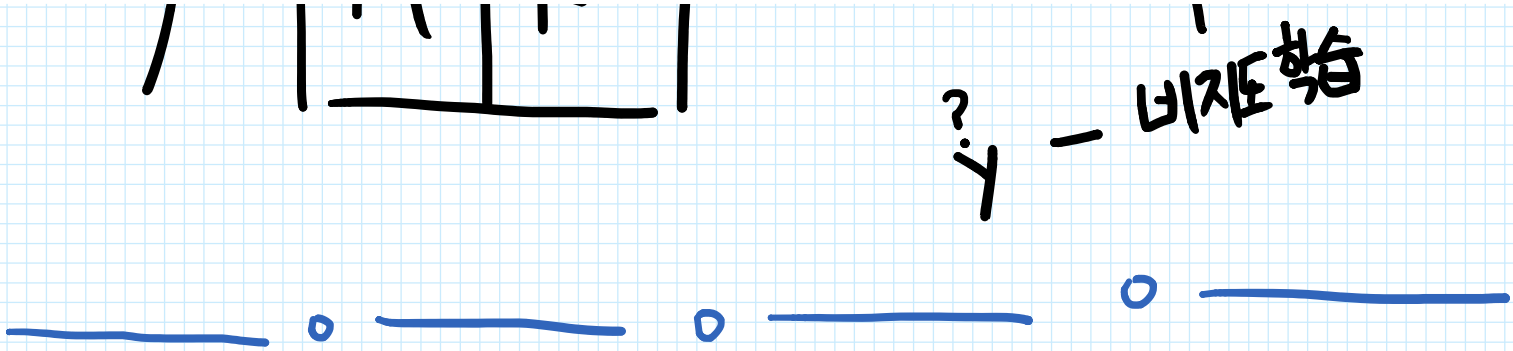
Q. "뇌"데이터

y



.fit(X, ?)

↑

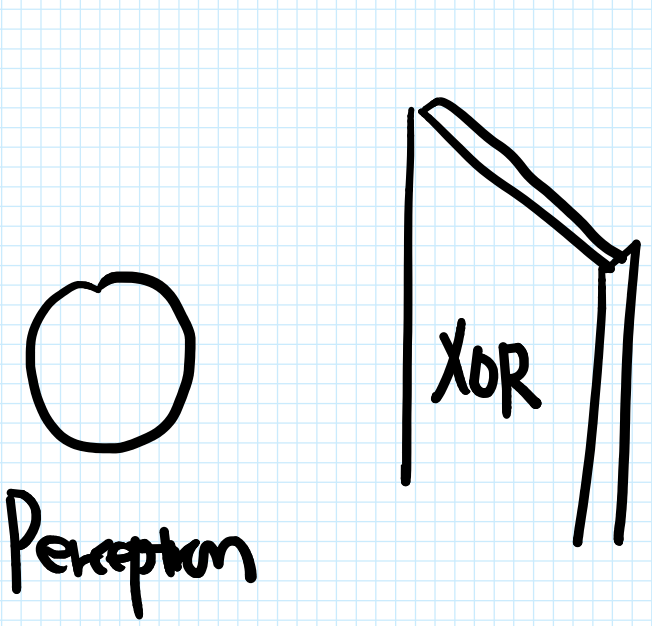


Kernel Support Vector Machine

비선형

1995

S.V.M.



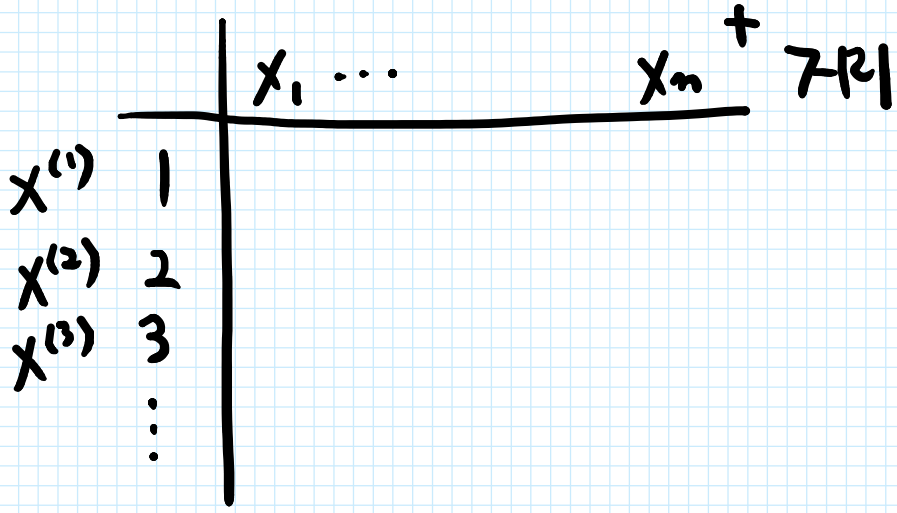
Perception

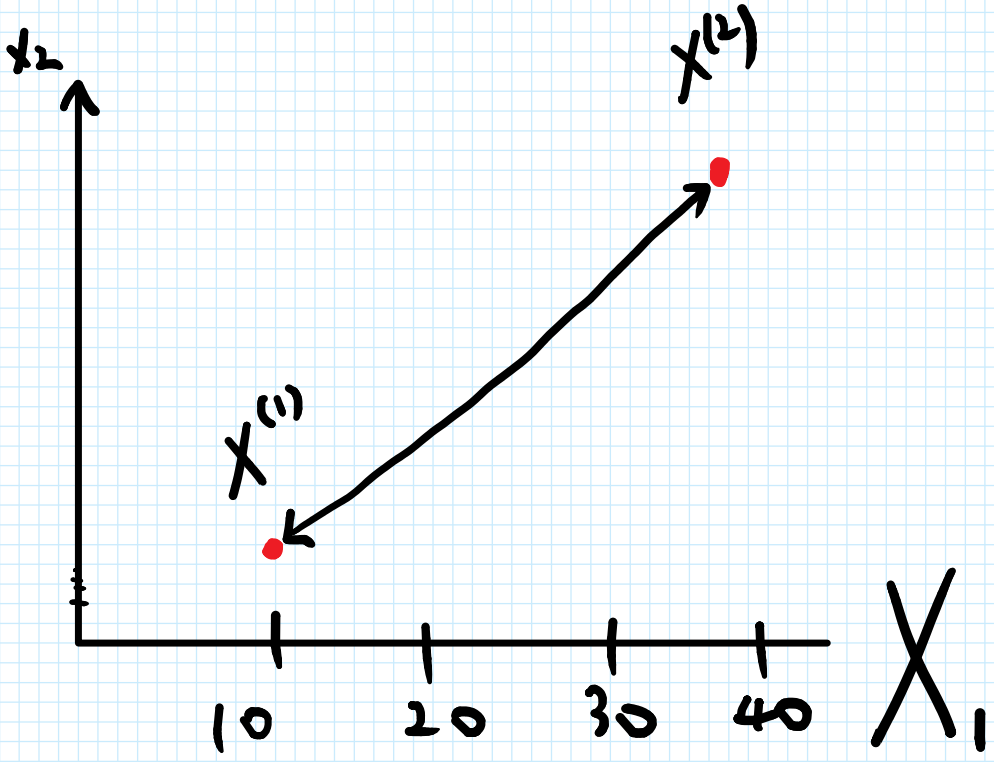
SVM 1995

Kernel "trick"

$$\|X^{(i)} - X^{(j)}\|^2$$

거리





1) 변환

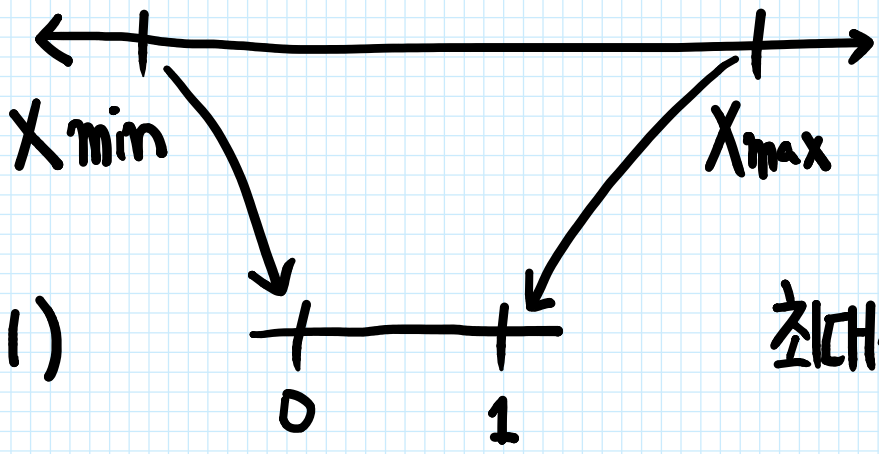
$$X' \leftarrow T(X)$$

2) 근접

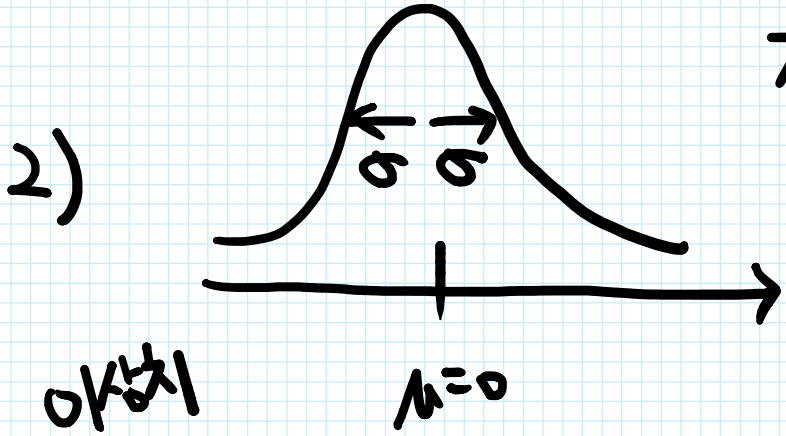
2) 관점

? y ⇐ 추정

단위조정 Scaling



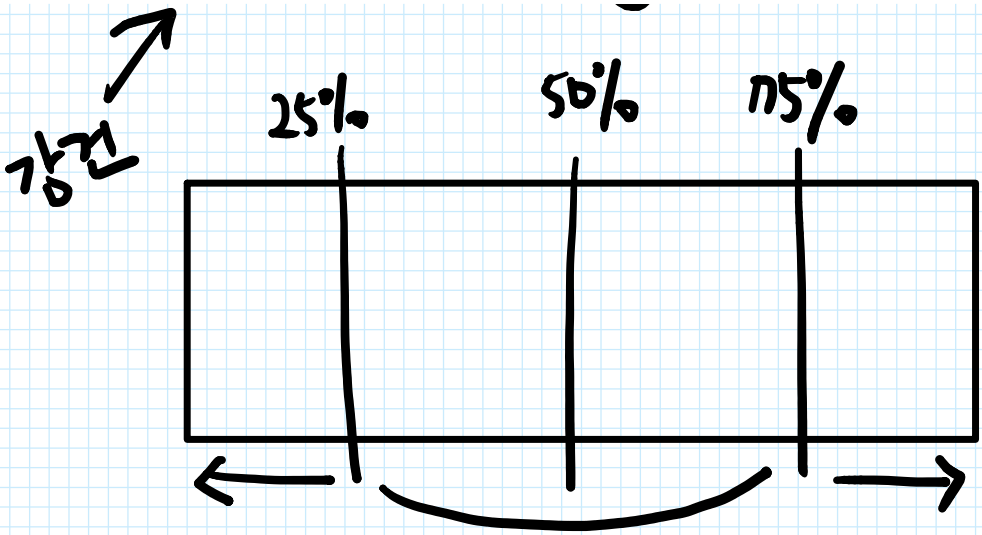
최대최소(MinMax)



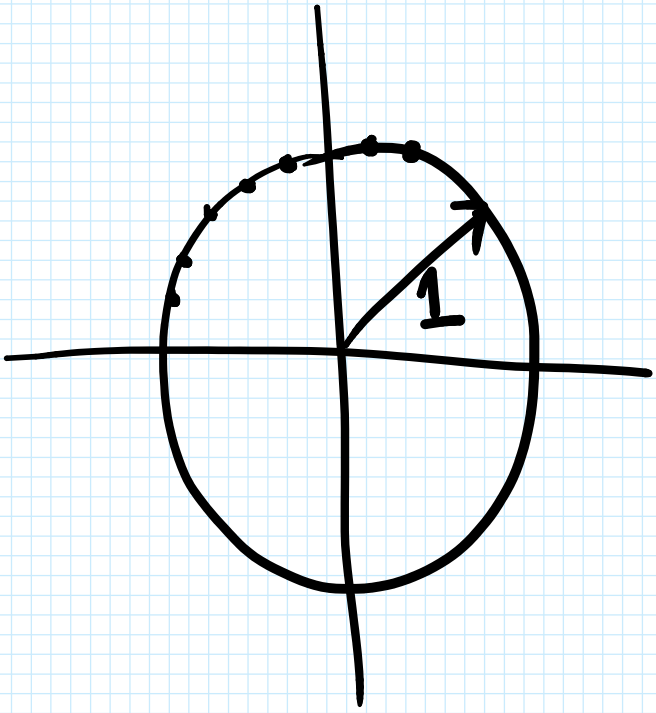
정규화(Standard)

아웃라이어

3) Robust Scaling

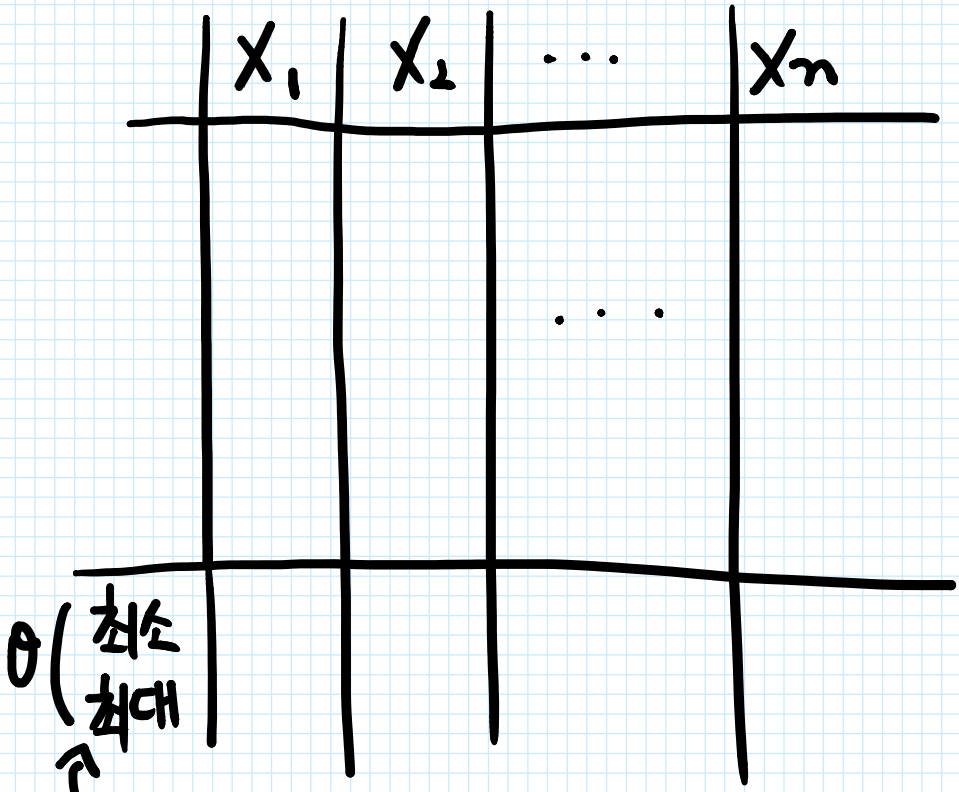


4)



Sklearn. preprocessing.

... Scaler( )



$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Scaler.fit(X)

$X_{new} = [x_1 \quad \cdots \quad x_n]$

X_{new_scaled}

sklearn.svm.

Log. Reg.

P.T.R

SVC/R(C=1.0, kernel='linear'

'rbf'

커널방식규정

gamma=

커널범위가정

'rbf'

gamma=

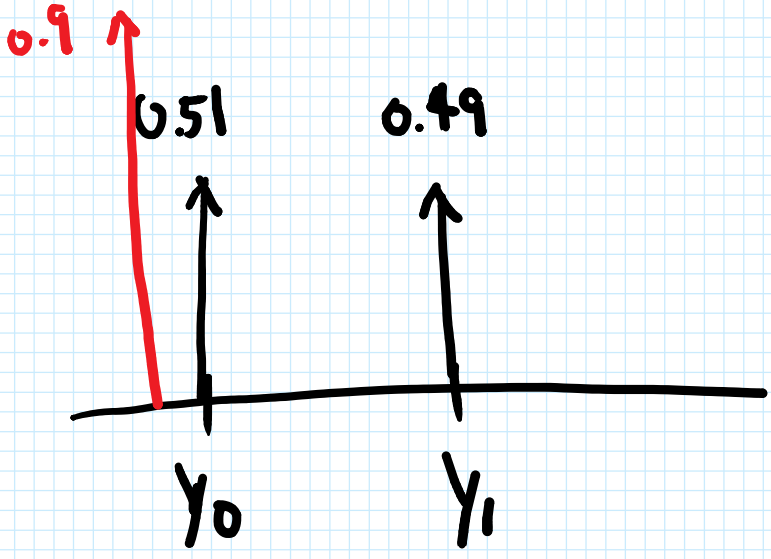
Sklearn.

model.predict(X) → label "$\hat{y}$"

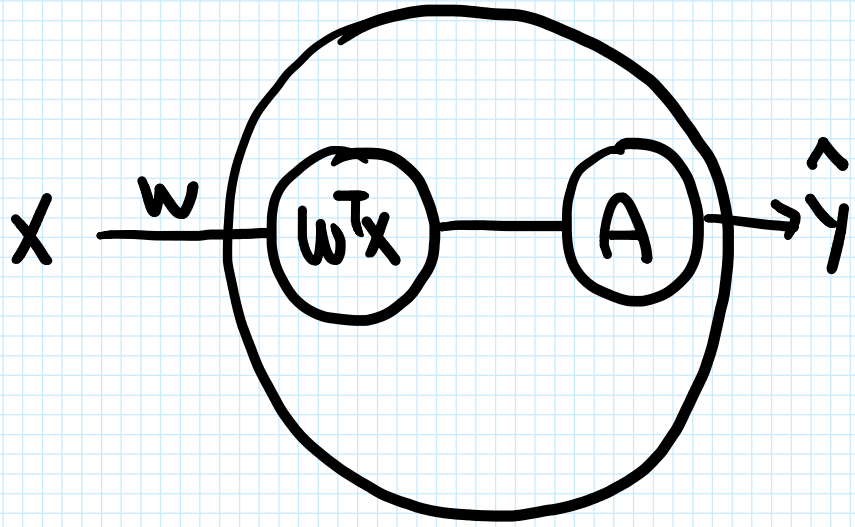
↑ proba

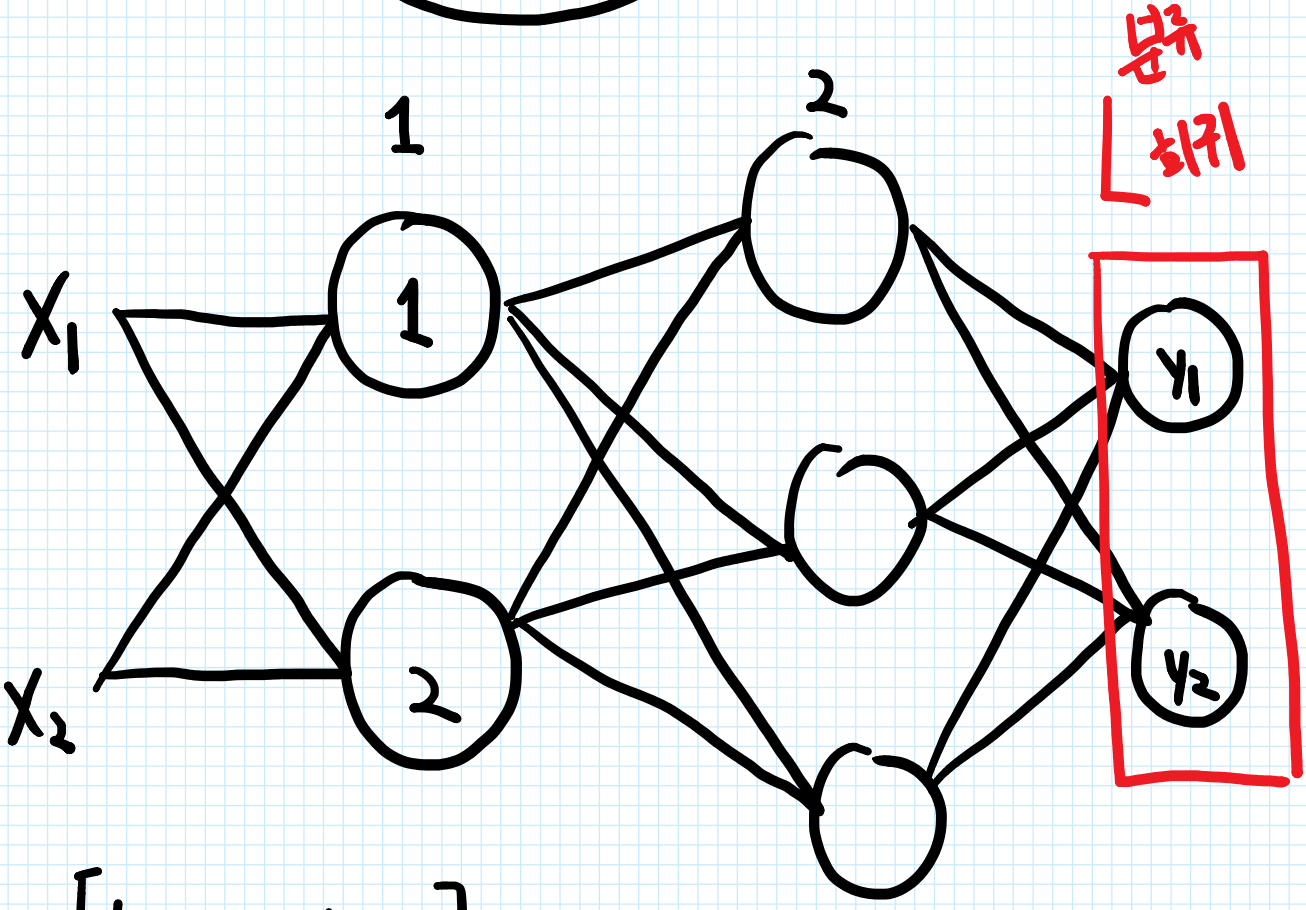
tensorflow(keras)

model.predict(X) → $[P(y_1), P(y_2), \cdots, P(y_3)]$



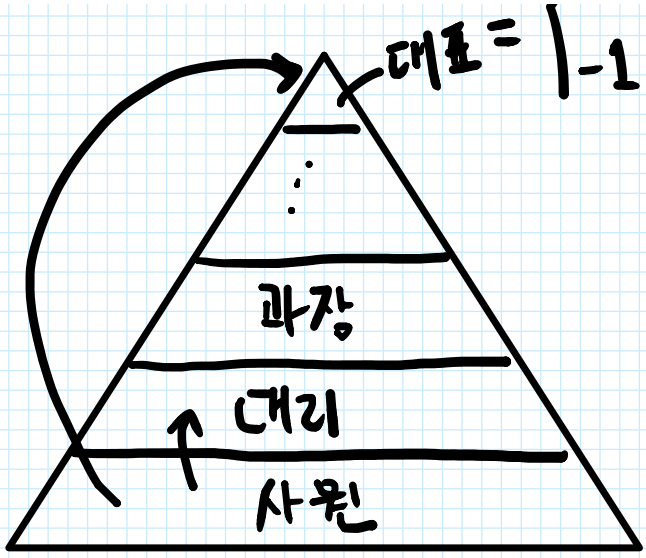
신경망 (다층 퍼셉트론) "딥러닝"

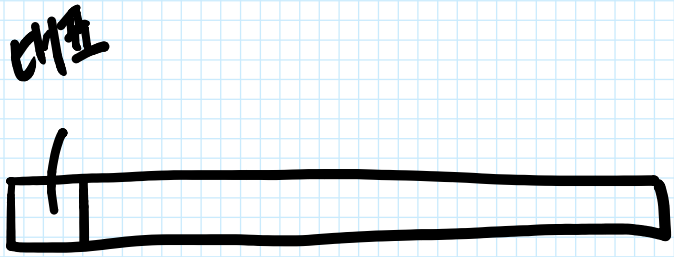




$$W=\begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \end{bmatrix}$$

→ 대표 = $\begin{cases} 1 \\ -1 \end{cases}$





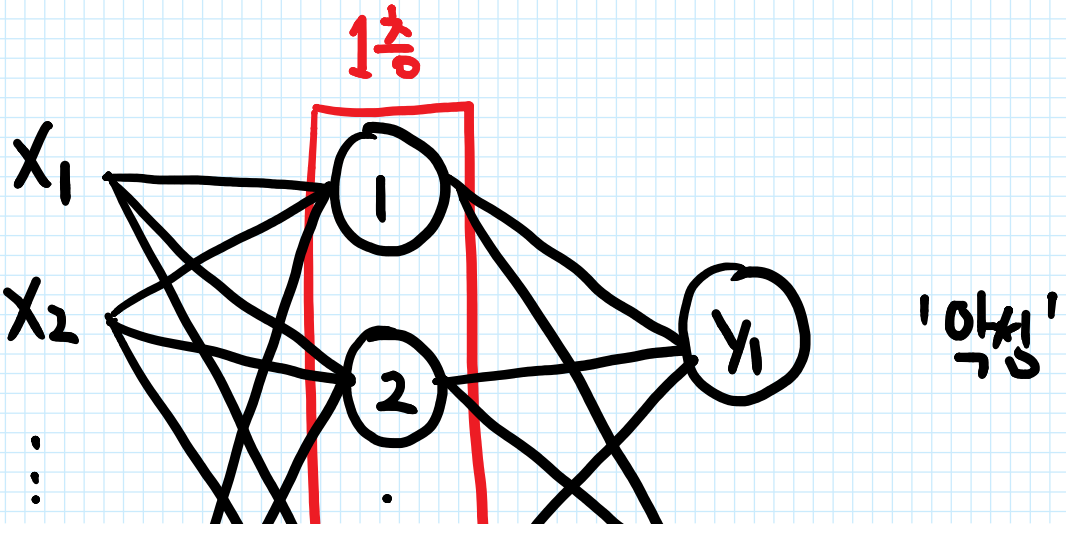
sklearn.neural_network

MLPClassifier/Regressor(

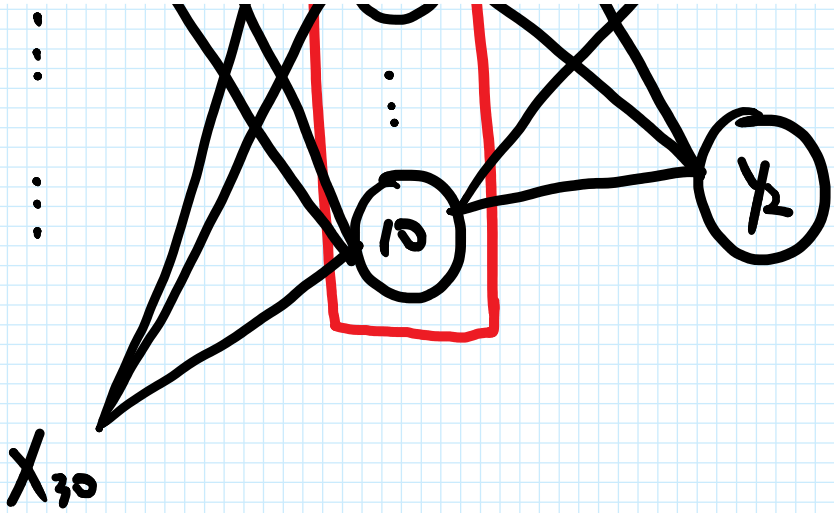
hidden_layer_size = [n1, n2, ...]

)

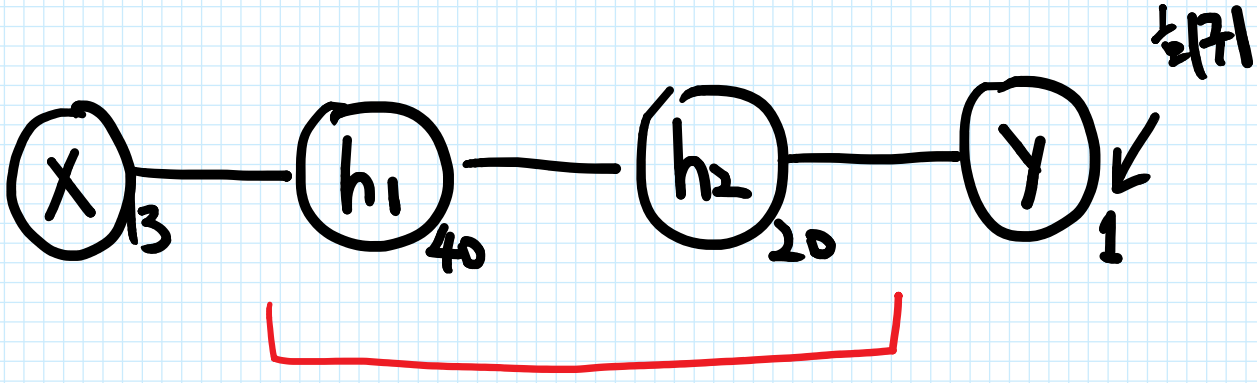
e.g. hidden_layer_size = [10]



'양성'



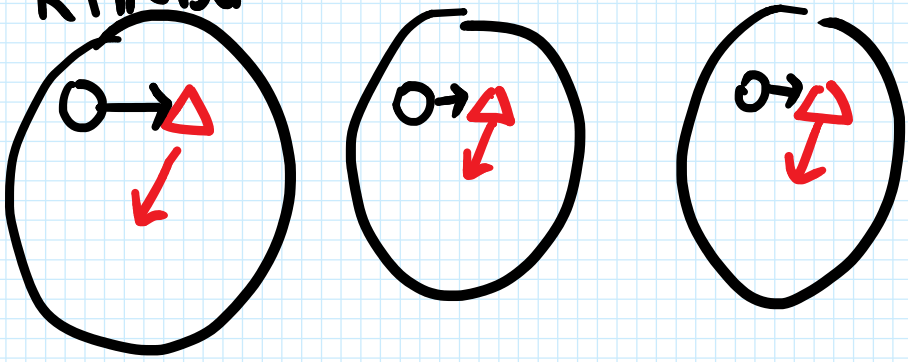
hidden_layer_size=[40,20]



군집

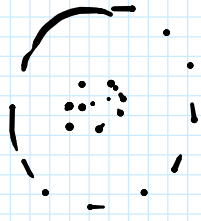
K-Means

1) RANDOM



DBSCAN

eps : 점들 거리



ARI

x y = [0, 1, 2]

y = [2, 0, 1]