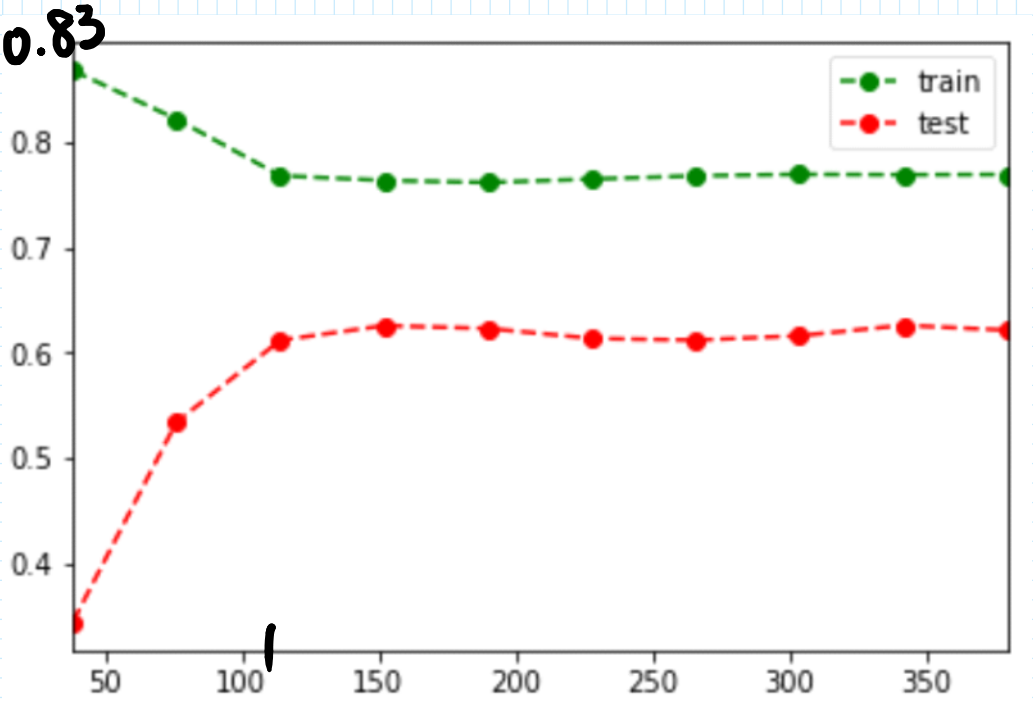
Day 3

2018년 12월 27일 목요일 오전 9:25

학습곡선



10% 30% # 샘플 100%

Ridge ✓ Lasso

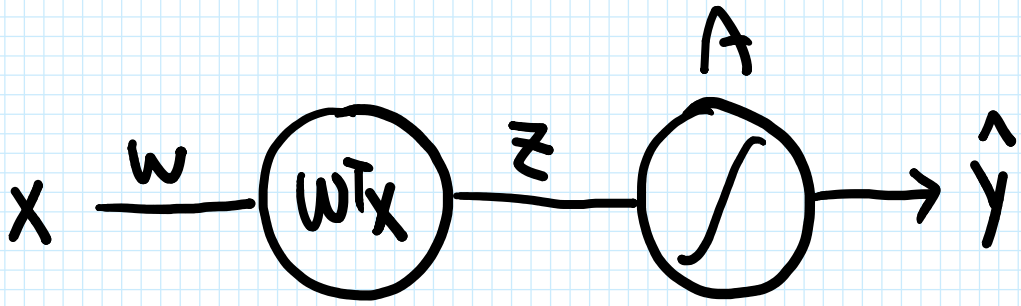
$+\alpha \|W\|^2$ $+\alpha \|W\|_1$

$=\alpha \sum W_i^2$ $=\alpha \sum |W_i|$

L2 L1

분류용 선형모델

Log. "Reg".



$$z = w_1 x_1 + \cdots + w_n x_n$$

$$= \sum_i w_i x_i$$

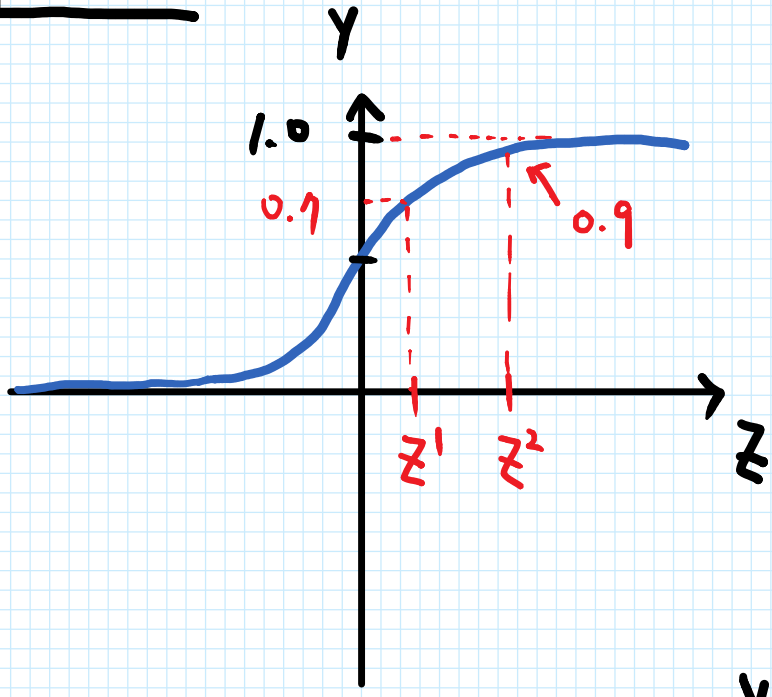
내적

$$= \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

$$= X w^T$$

$$= w^T x$$

Sigmoid



$z = \pm\infty$

$y = [0, 1]$

e.g. $y^1 = 0.6$

$y^2 = 0.9$

	y_1	y_2	$\hat{y}$
	60%	40%	y_1
	10%	90%	y_2

확률

```
예측확률 = logreg.predict_proba(X_test)
```

In [33]:

```
pd.DataFrame(예측확률)[:10]
```

.predict( )

Out[33]:

0.867

	y 0	y 1
0	8.674260e-01	0.132574
1	3.477141e-02	0.965229

$\hat{y}$

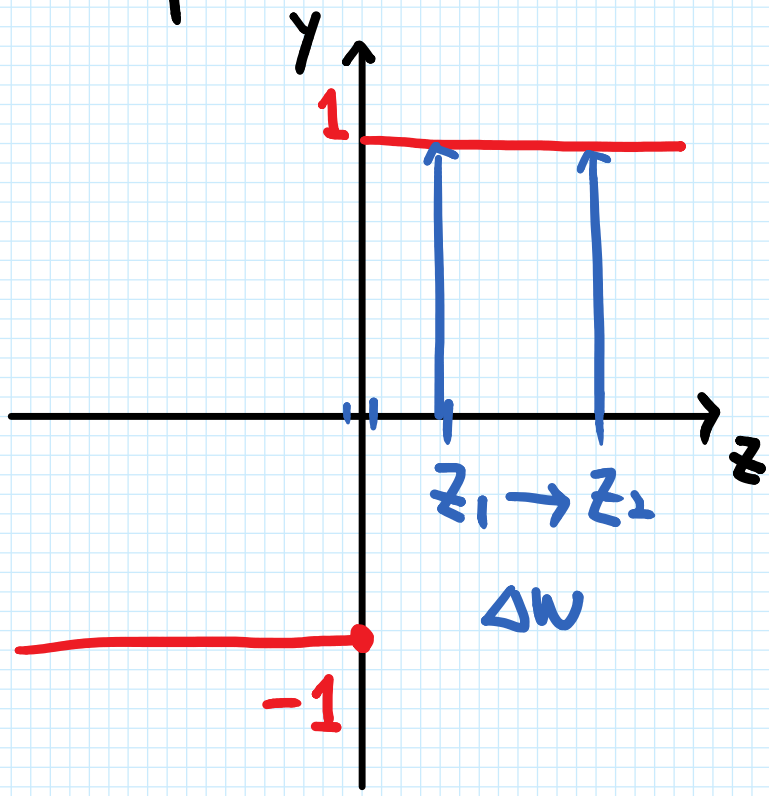
y_0

$\sum_i p(y_i) = 1.0$

$0 \le p(y_i) \le 1$

0	8.674260e-01	0.132574	y_0
1	3.477141e-02	0.965229	y_1
2	7.713378e-03	0.992287	y_1
3	2.684808e-02	0.973152	y_1

Perceptron



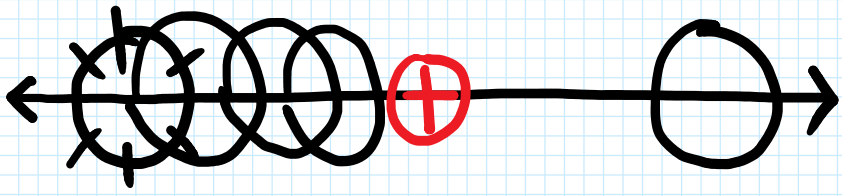
$Loss(y, \hat{y})$

$y=1$

$\sum (y^{(i)} - \hat{y}^{(i)})^2$

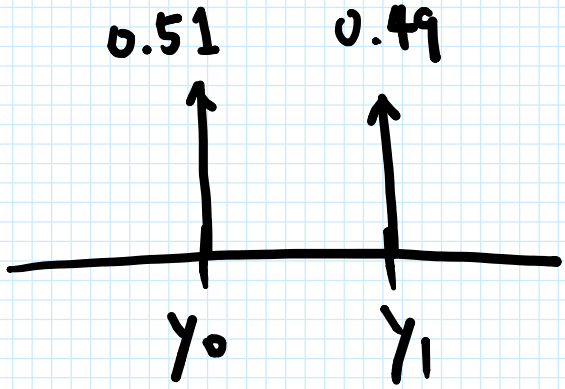
$y^{(1)} - \hat{y}^{(1)} = 1-1=0$

$y^{(2)} - \hat{y}^{(2)} = 1-1=0$



$\begin{bmatrix} x_1 & \cdots & x_{20} \end{bmatrix}$ ← $\begin{matrix} \hat{} & \hat{} \\ \vdots & \vdots \\ w & \end{matrix}$

$[x_1 \cdots x_{30}]$



$\hat{y} = y_0$

$$Loss(y, \hat{y}) + \underbrace{\text{벌칙(penalty)}}_{\text{= 규제 (Regularization)}}$$

$\alpha = \frac{1}{C}$

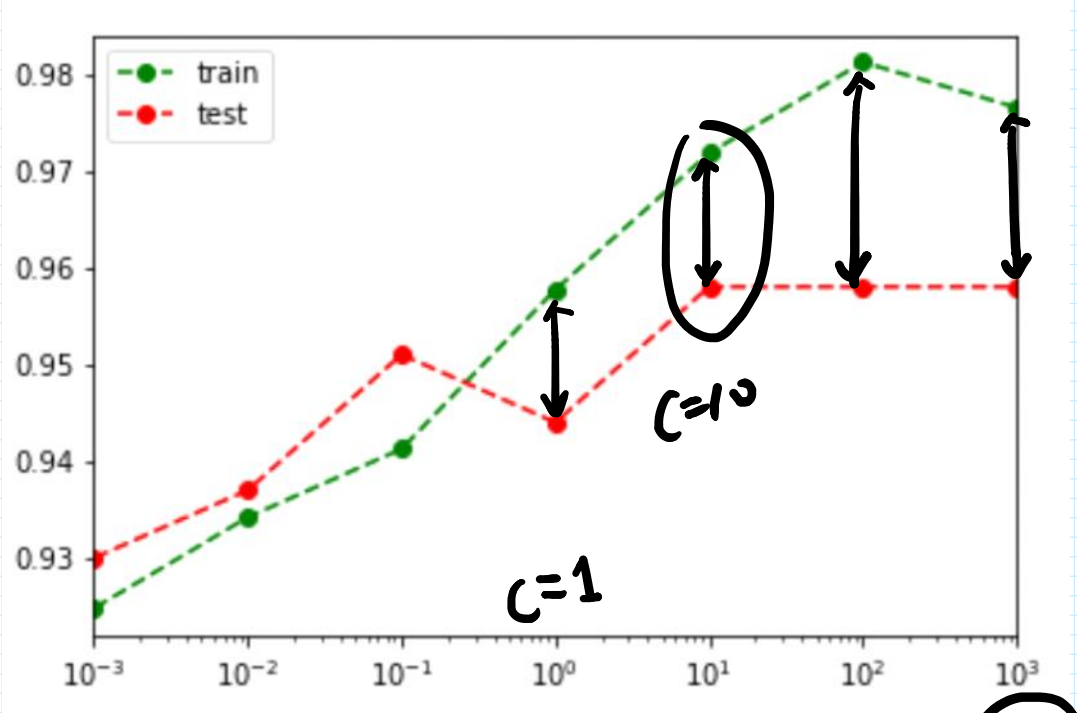
$\frac{1}{C}\|W\|^2$ 가중치 크기

L2 "기본"

C = 1/α

Grid Search

C \ γ	0.01	0.1	1.0
0.001			
0.01			
0.1		✓BEST	
1.0			
⋮			



C

C

ALT

Q. 텍스트 데이터 → 숫자 (Encoding)

e.g. BoW "단어뭉치" 1960s

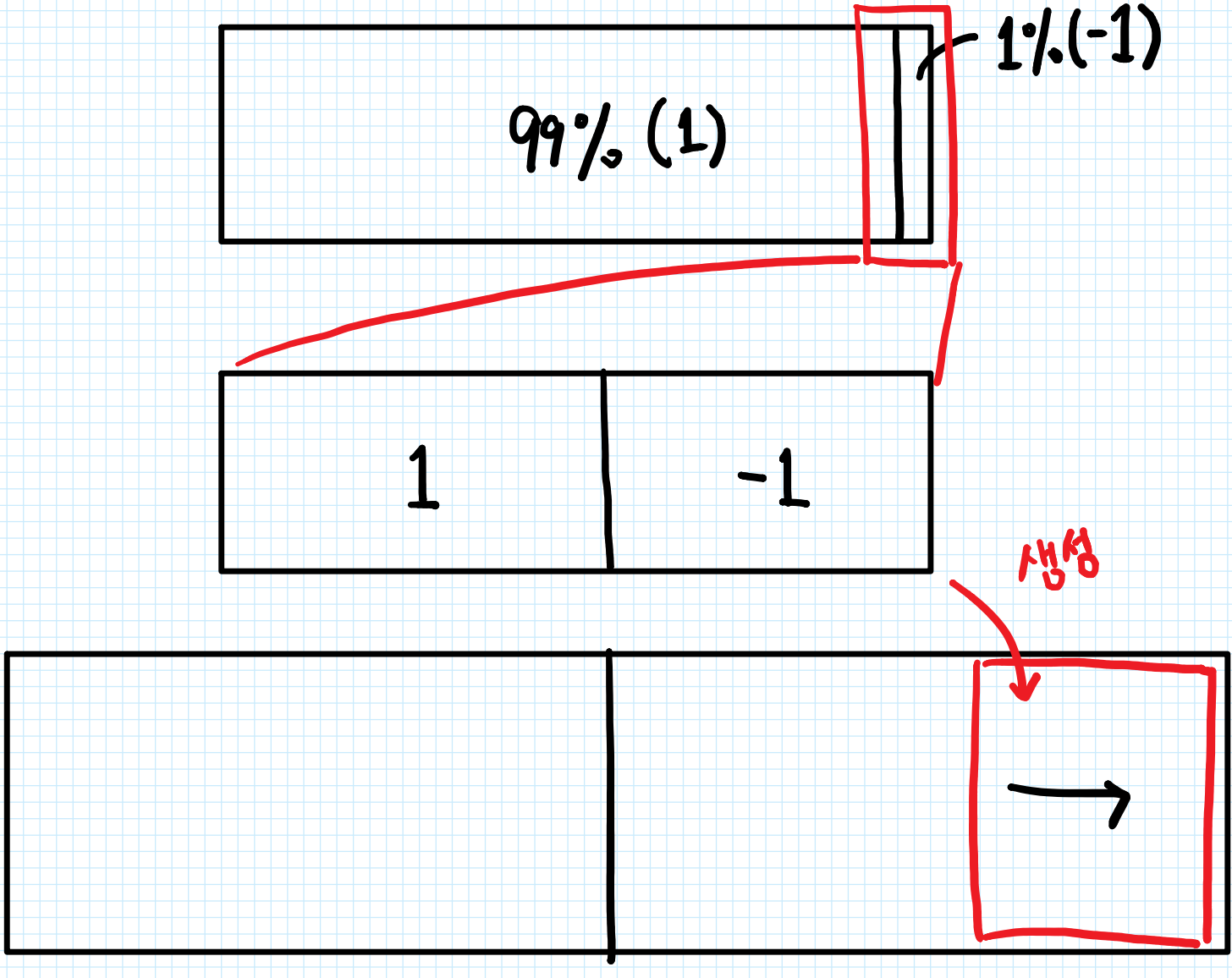
1. 파이썬 데이터 분석
2. 파이썬 머신러닝
3. 파이썬 딥러닝

BoW → { 파이썬, 데이터, 분석, 머신러닝, 딥러닝 }

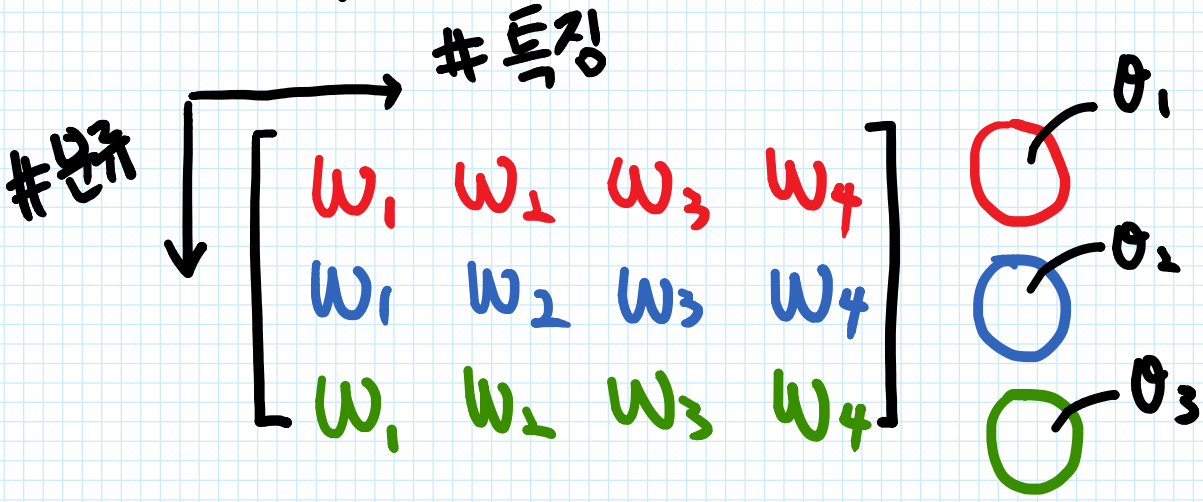
샘플 \ 특징	x_1 파이썬	x_2 데이터	x_3 분석	x_4 머신러닝	x_5 딥러닝	y
	1	1	1	0	0	
	1			1		

1	0	0	1	0
1	0	0	0	1

Q. 데이터 분포가 상당히 치우친 경우?



다중분류



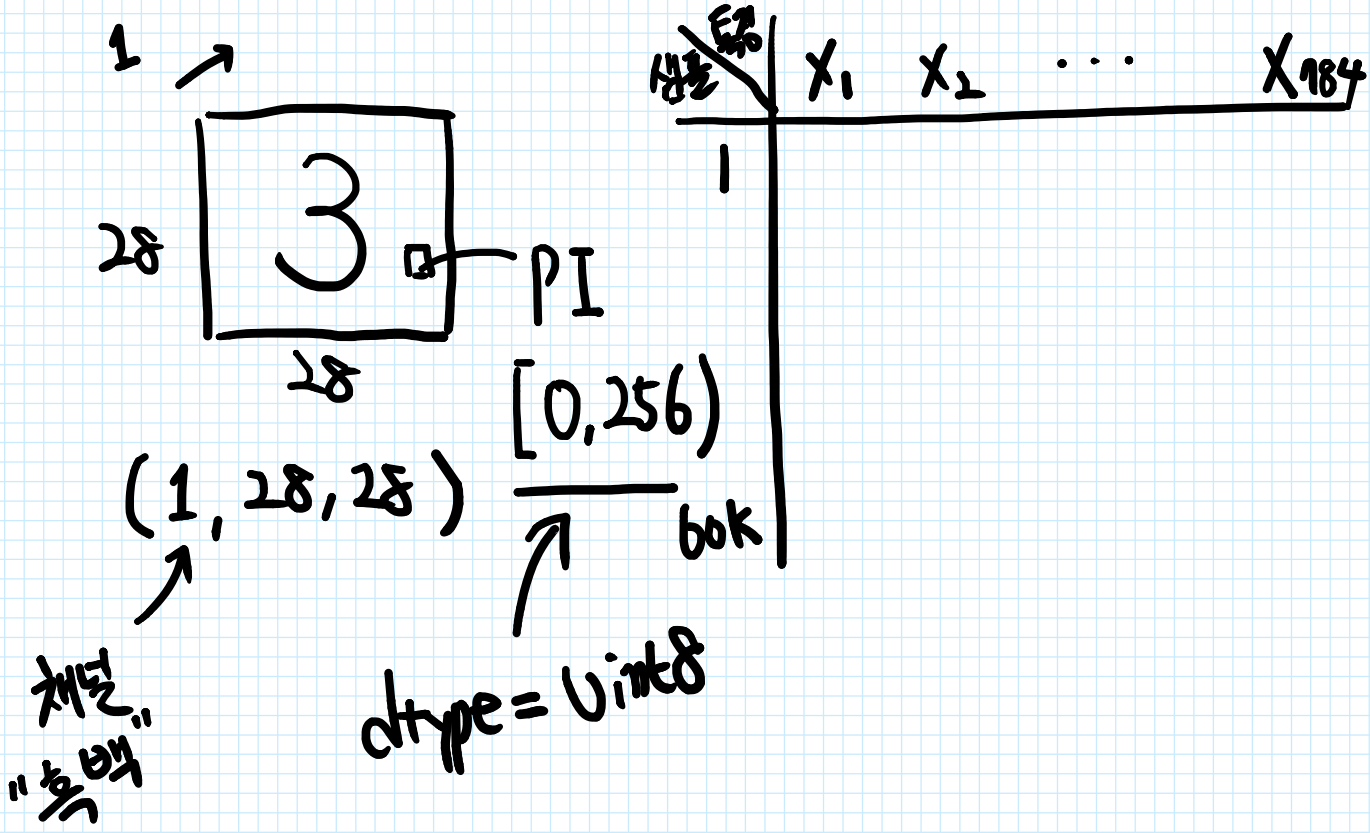
3×4

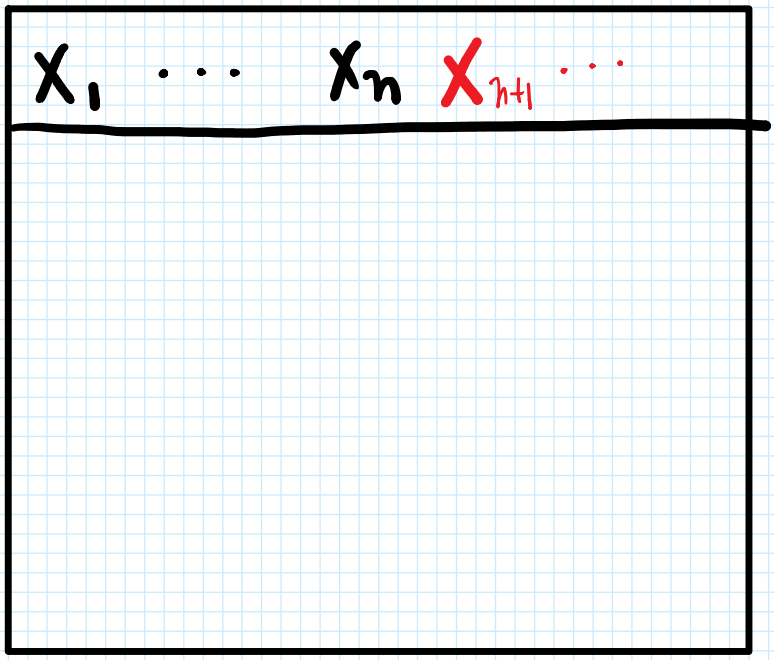
```
Sklearn.linear_model.
Log.Reg.(..., multi_class='ovr')
```

One Vs. Rest 알고리즘

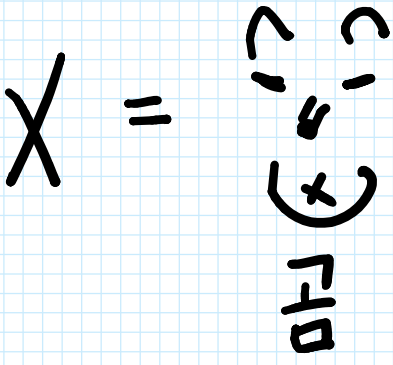
대용량 데이터 훈련

DataSet: MNIST





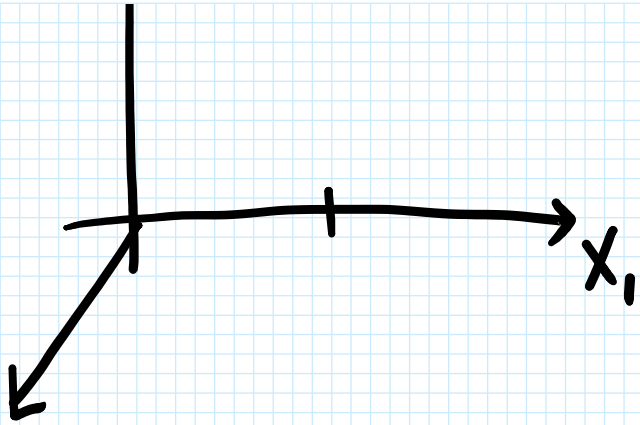
결정 트리
Desicism Tree

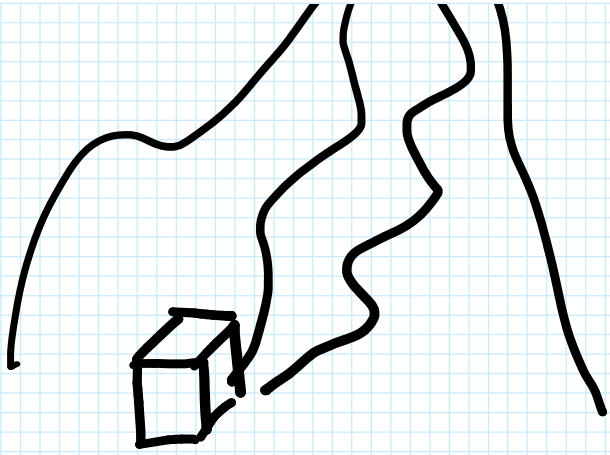


알고리즘

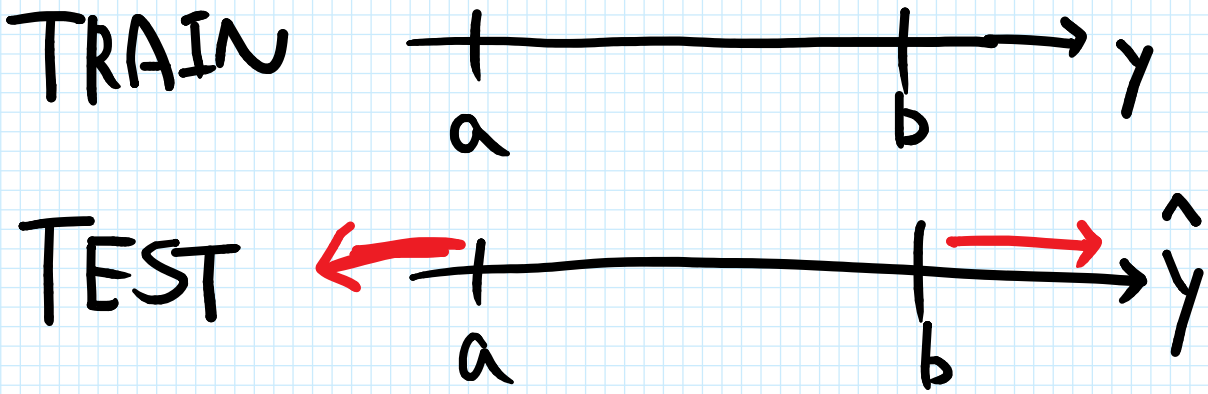
- if
- elif
- else

X_1 X_2 ... X_n 특성 선택

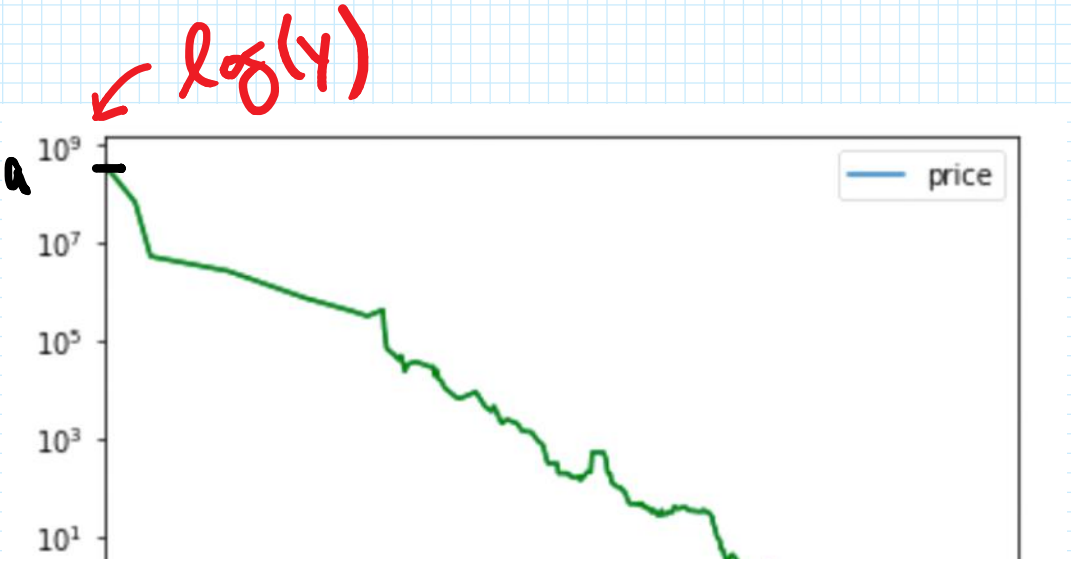


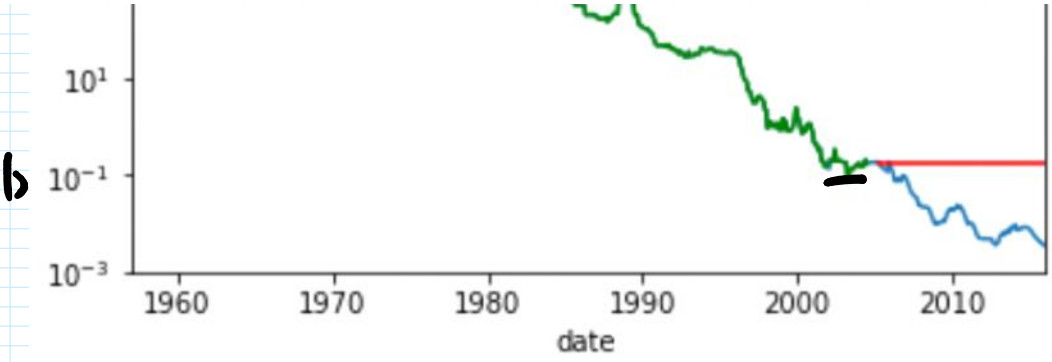


결정트리 회귀



오차 유의





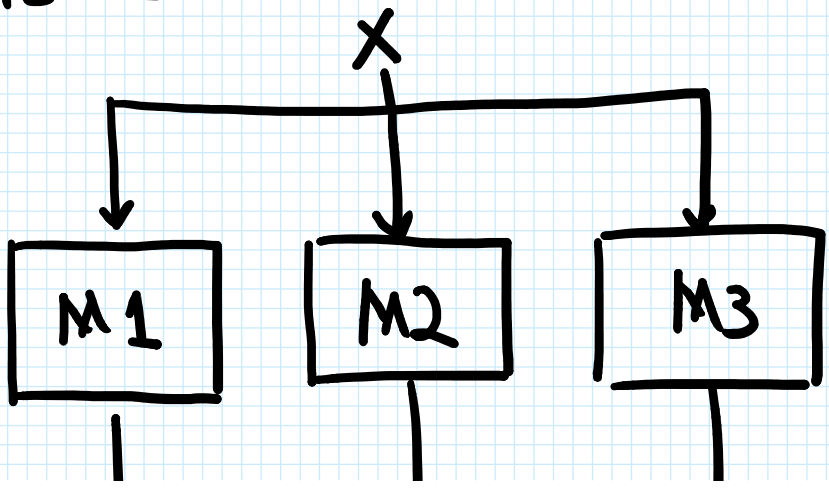
$\text{linreg}(X_{tr}, \log y_{tr})$

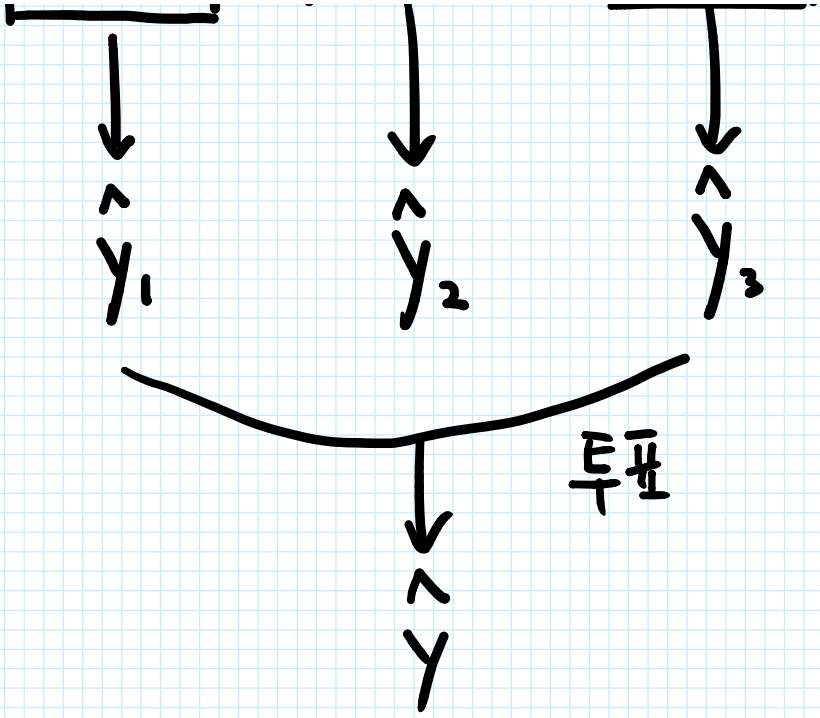
$.\text{predict}(X_{tr}) \rightarrow \log\hat{y}$

$e^{\log\hat{y}} \rightarrow \hat{y}$

조합 모델 ensemble

→ 과적합 개선





Random Forest