



PyTorch Zero To All 6

# Logistic Regression

오수지



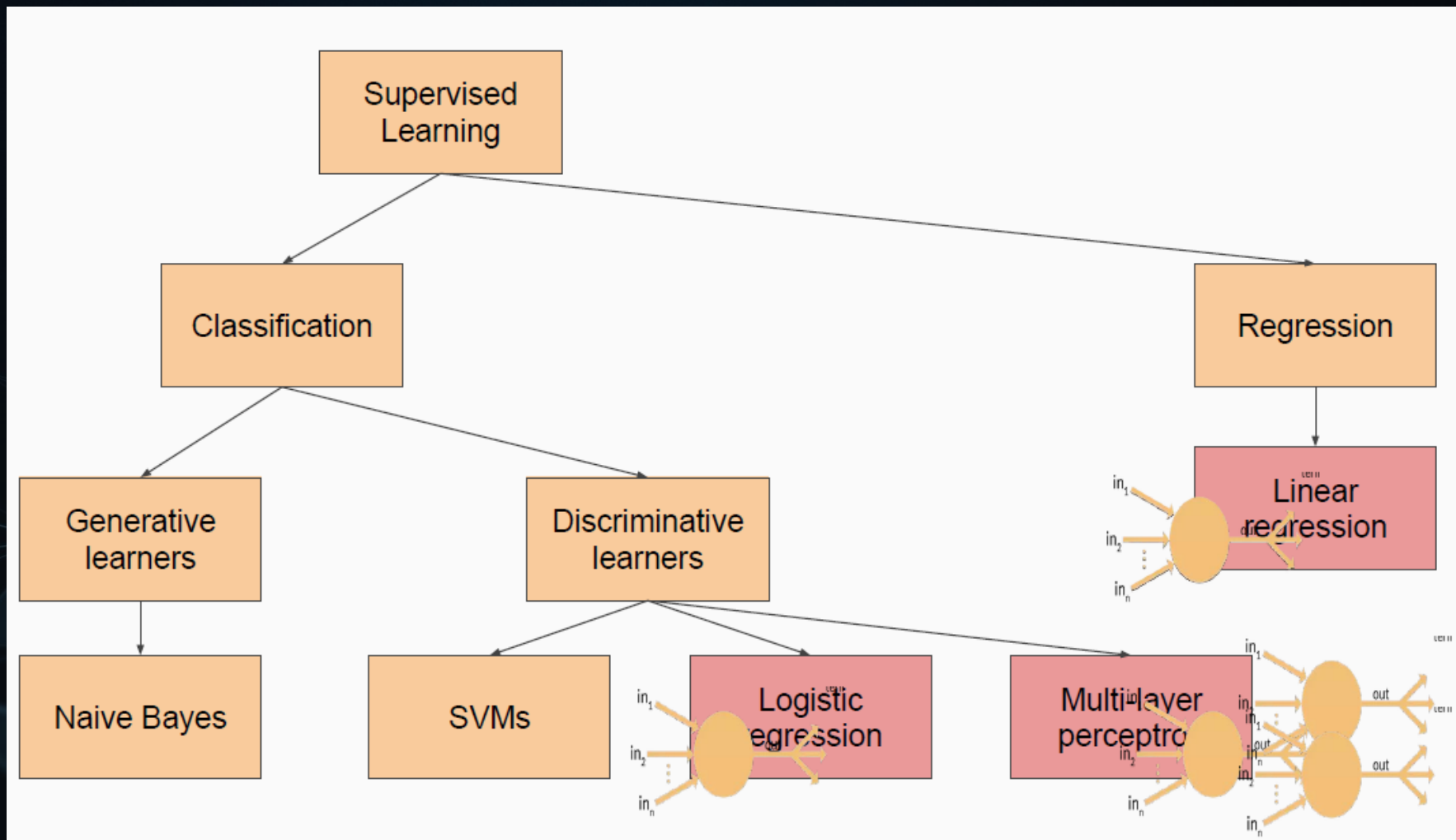
# Logistic Regression

Hours (x)	Points	fail/pass
1	2	0
2	4	0
3	6	1
4	?	?





# Logistic Regression





# Logistic Regression

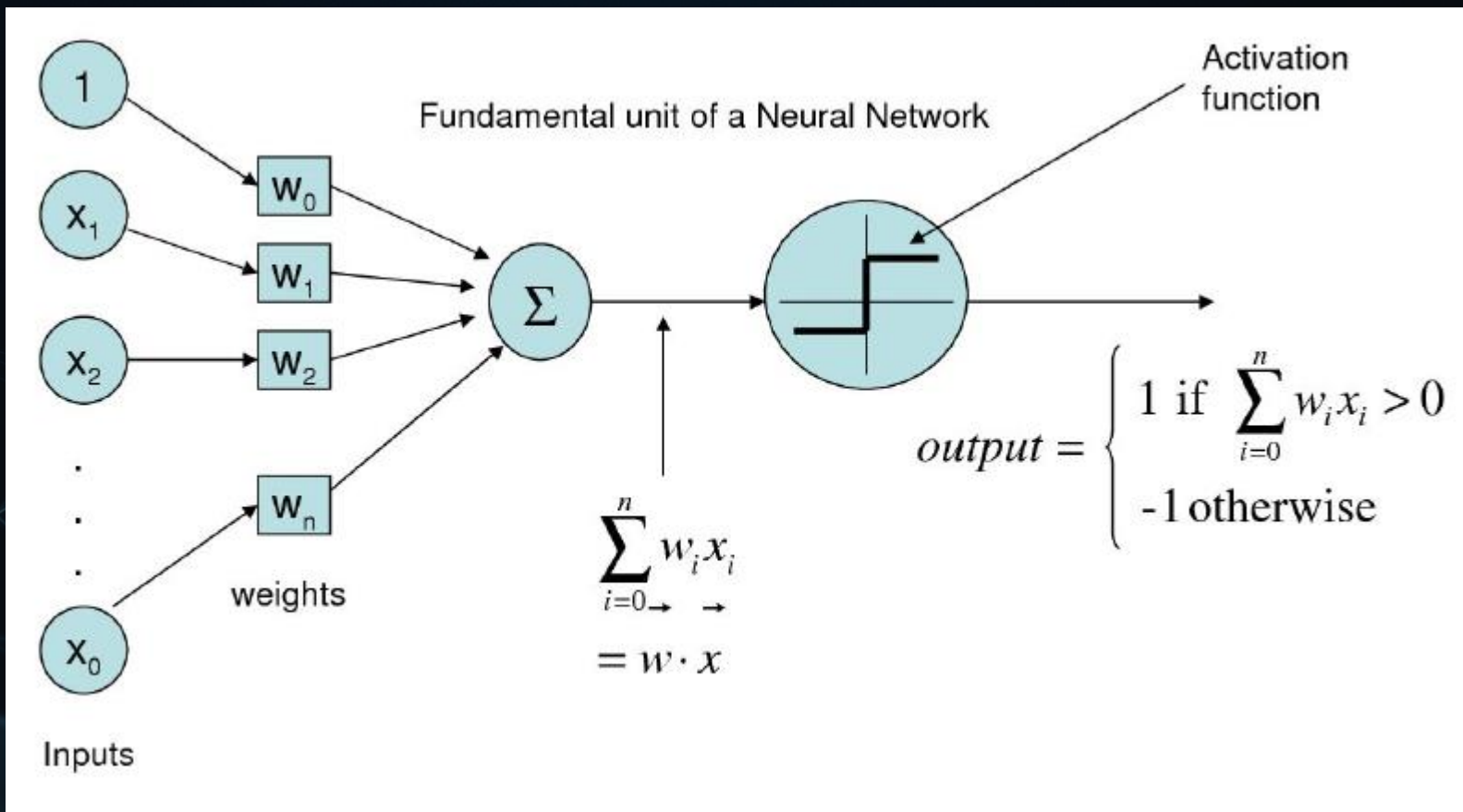
## Linear Regression

Sample	Temperature $x_1$	Rain/snow $x_2$	km of biking
s1	-22	10	2
s2	-24	12	0
s3	-15	20	3
s4	-8	13	4
s5	-5	40	3
s6	8	45	8
s7	-2	5	10
s8	5	5	12
s9	12	2	12
s10	8	5	14
s11	12	20	20
s12	13	15	20
s13	15	10	25
s14	18	5	20
s15	19	20	40

Sample	Temperature $x_1$	Rain/snow $x_2$	Bike / Drive
s1	-22	10	Drive
s2	-24	12	Drive
s3	-15	20	Drive
s4	-8	13	Drive
s5	-5	40	Drive
s6	8	45	Drive
s7	-2	5	Bike
s8	5	5	Bike
s9	12	2	Bike
s10	8	5	Bike
s11	12	20	Bike
s12	13	15	Bike
s13	15	10	Bike
s14	18	5	Bike
s15	19	20	Bike

## Logistic Regression

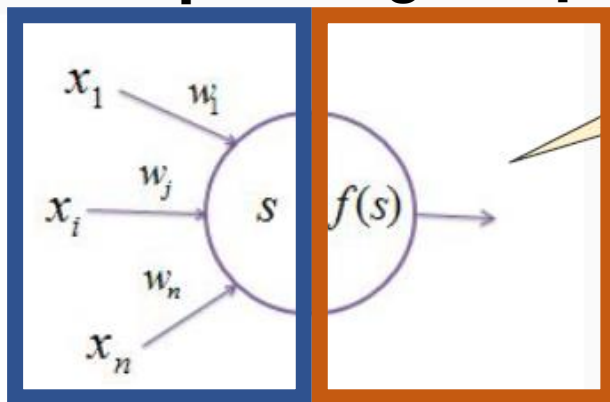
# Logistic Regression





# Logistic Regression

## Linear part Logistic part



For a linear model  
 $f(s) = s$

Summation

$$s = \sum w \cdot x$$

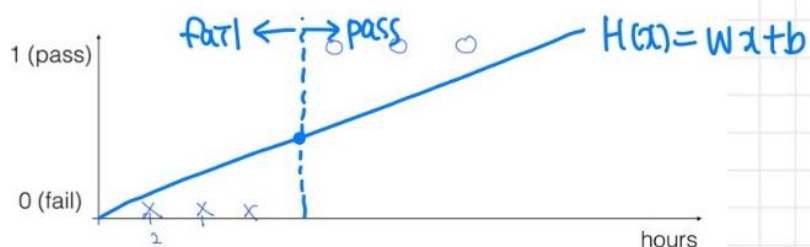
Transformation

$$f(s) = \frac{1}{1 + e^{-s}}$$

For a non-linear  
model,  $f(s)$  will apply  
a transformation

# Logistic Regression

- Classification을 Linear Regression으로 할 수 있을까?

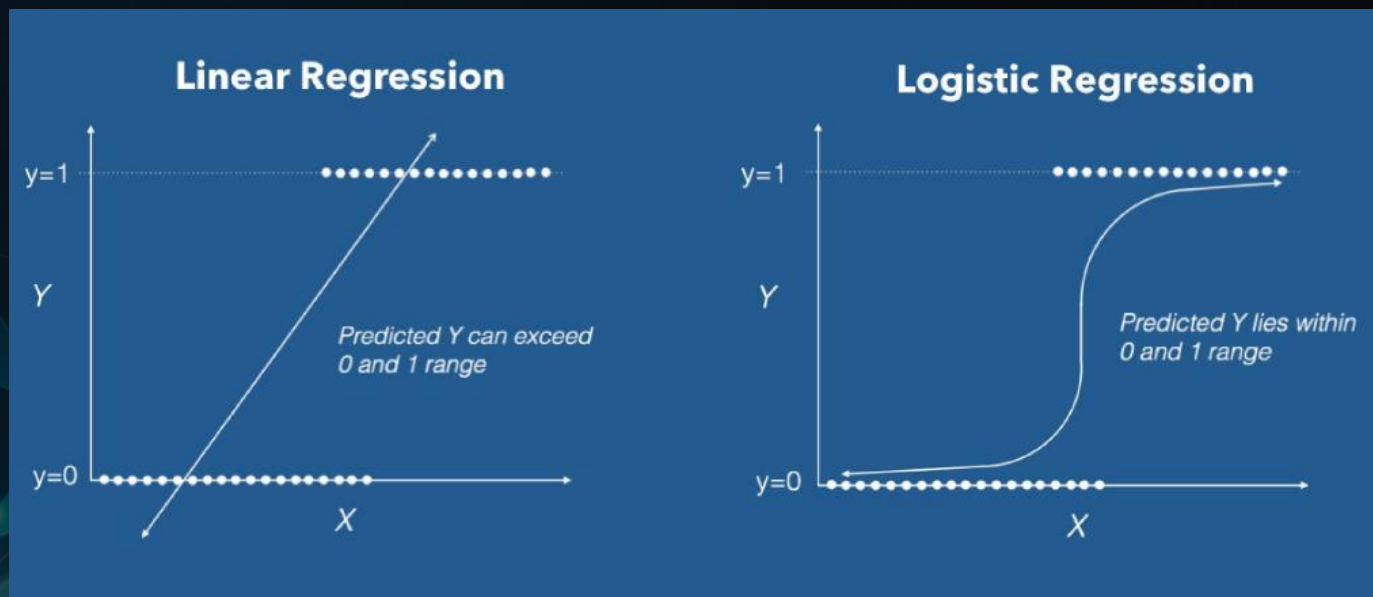


⇒ 가능한 함

⇒ BUT 예측값은 0 ~ 1 인데

Input 에 따라  $H(x)$ 는 1보다 크거나  
0보다 작은 값을 리턴함

↳  $H(x)$  값이 항상 0에서 1 사이에 있어야면?

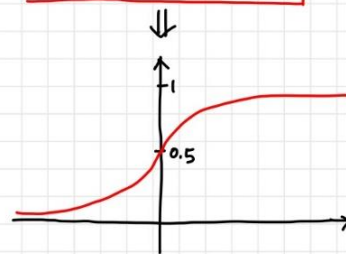




# Logistic Regression

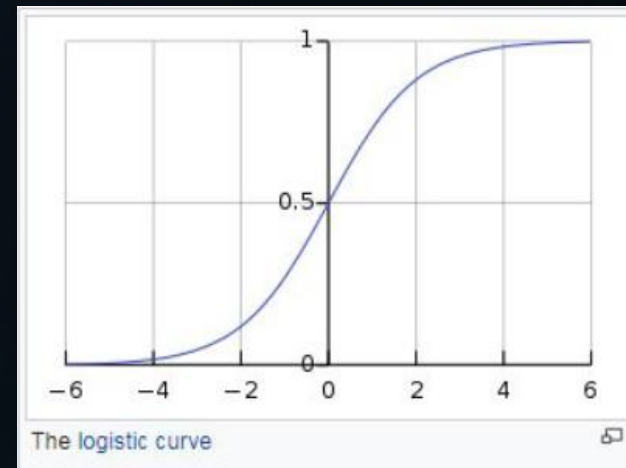
## • Sigmoid (Logistic) Function

$$g(z) = \frac{1}{1 + e^{-z}} \quad z = wx$$



$$H(x) = \frac{1}{1 + e^{-wx}} \quad H(x) \text{는 항상 } 0 \sim 1$$

↳ 이때의 cost function 모양은  
 $H(x) = wx + b$  일때와 달리





# Logistic Regression

- 이 경우를 위한 **NEW COST FUNCTION** !

$$\text{cost}(W) = \frac{1}{m} \sum c(H(x), y)$$

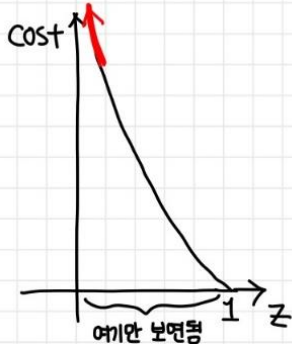
전체 cost의 평균

$$c(H(x), y) = \begin{cases} -\log(H(x)) & : y=1 \\ -\log(1-H(x)) & : y=0 \end{cases}$$

$$= -y \log(H(x)) - (1-y) \log(1-H(x))$$

# Logistic Regression

•  $y = 1$  (실제값이 1)



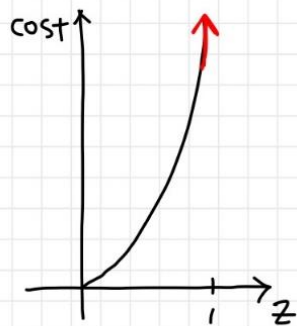
①  $H(x) = 1$  : 예측 성공

$\therefore \text{cost}(1) = 0$

②  $H(x) = 0$  : 예측 실패

$\therefore \text{cost}(0) = \infty$

•  $y = 0$  (실제값이 0)



①  $H(x) = 1$  : 예측 실패

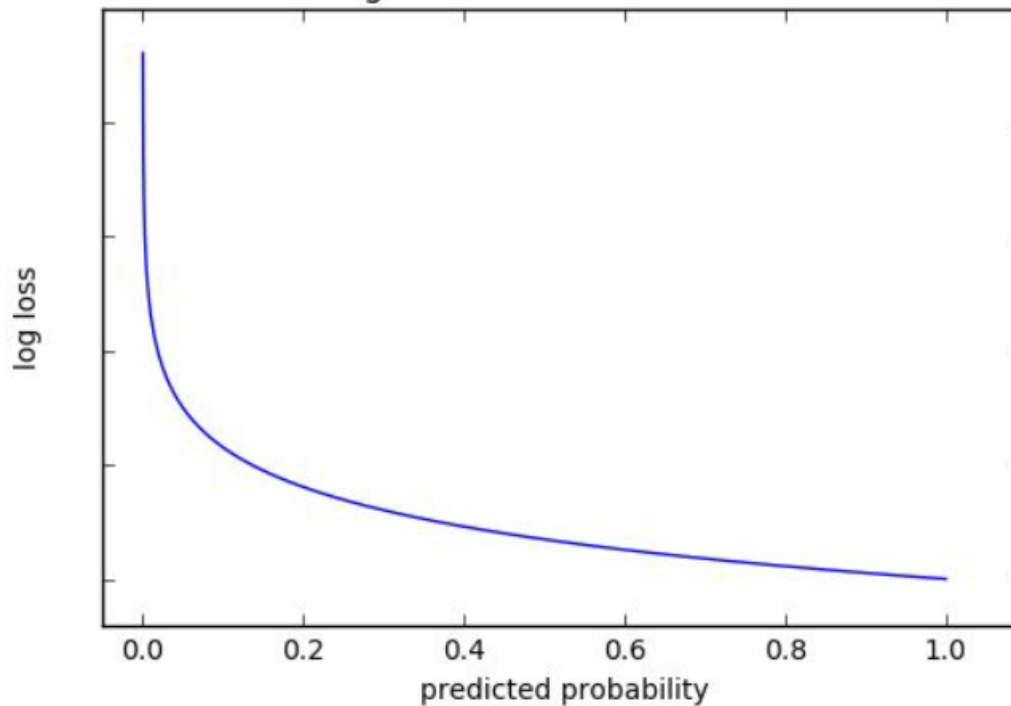
$\therefore \text{cost}(1) = \infty$

②  $H(x) = 0$  : 예측 성공

$\therefore \text{cost}(0) = 0$

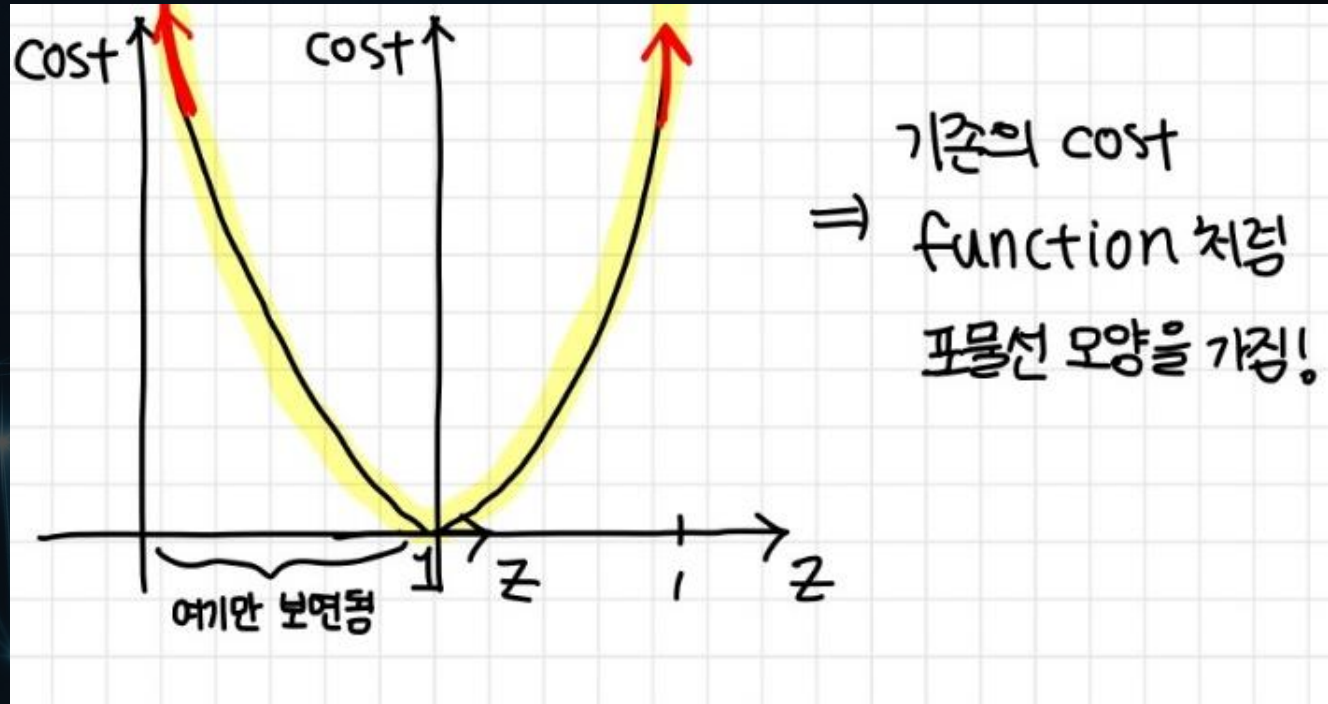
⇒ 예측 실패한 경우 cost가 커지고  
성공한 경우 cost가 0이 됨

Log Loss when true label = 1

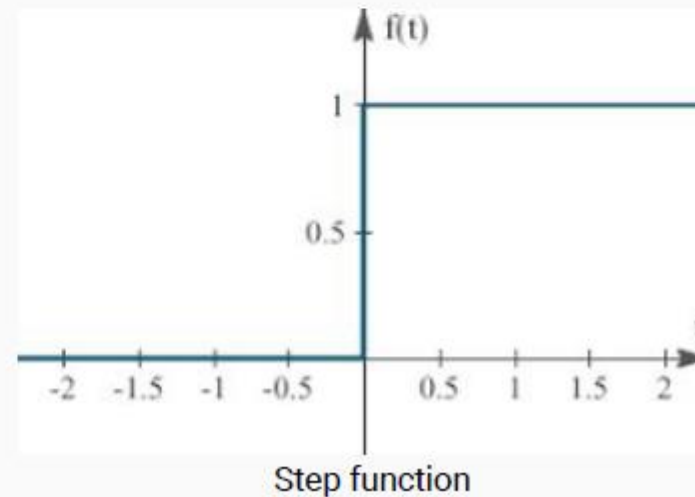
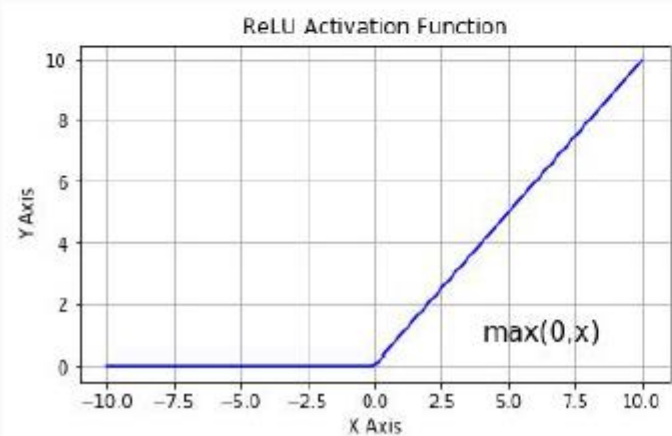
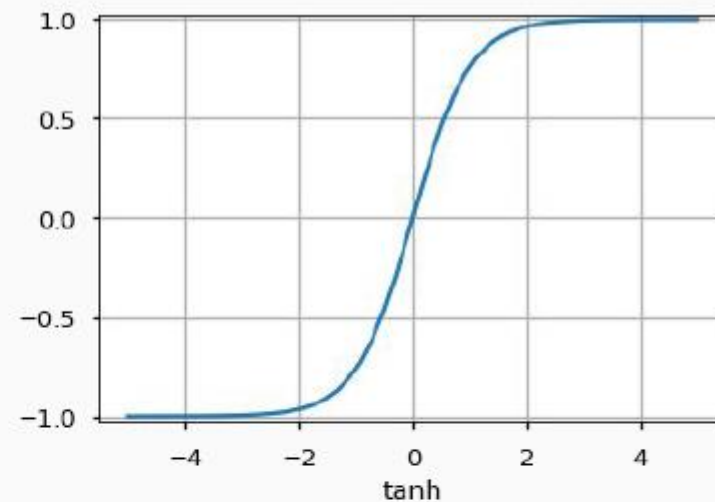
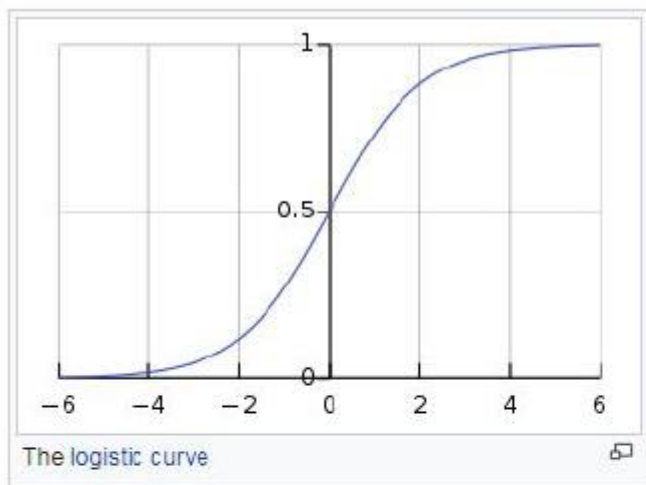




# Logistic Regression



# Logistic Regression







# Logistic Regression

```
1: procedure Logistic_regression_learner( $Xs, Y, Es, \eta$ )
2:   Inputs
3:      $Xs$ : set of input features,  $Xs = \{X_1, \dots, X_n\}$ 
4:      $Y$ : target feature
5:      $Es$ : set of training examples
6:      $\eta$ : learning rate
7:   Output
8:     function to make prediction on examples
9:   Local
10:     $w_0, \dots, w_n$ : real numbers
11:    initialize  $w_0, \dots, w_n$  randomly
12:    define  $pred(e) = \text{sigmoid}(\sum_i w_i * X_i(e))$ 
13:    repeat
14:      for each example  $e$  in  $Es$  in random order do
15:         $error := Y(e) - pred(e)$ 
16:         $update := \eta * error$ 
17:        for each  $i \in [0, n]$  do
18:           $w_i := w_i + update * X_i(e)$ 
19:    until termination
20:    return  $pred$ 
```



D-ai-ving

**감사합니다!**