Idea Factory Intensive Program #2

딥러닝 홀로서기 #2

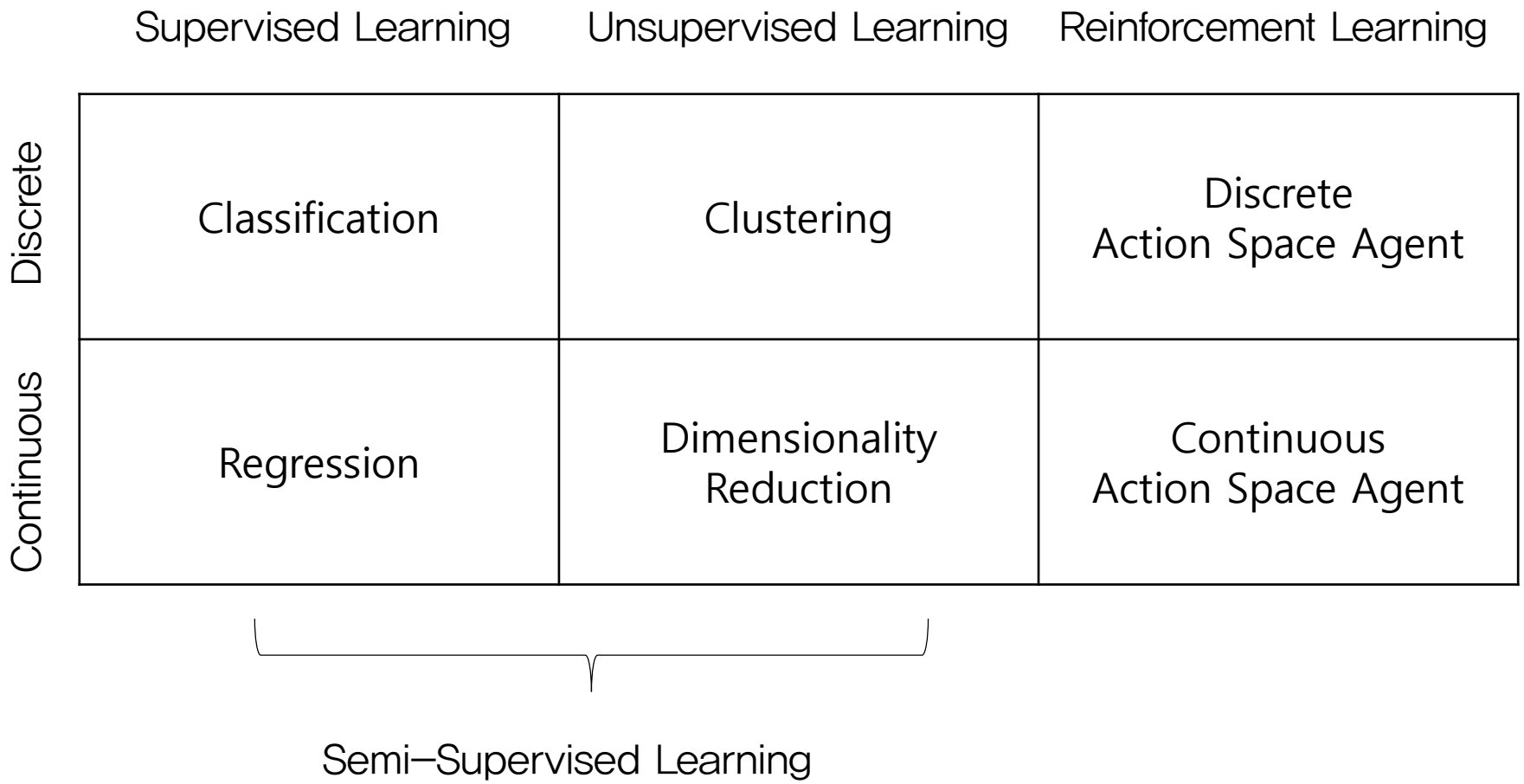
이론강의/PyTorch실습/코드리뷰

딥러닝(Deep Learning)에 관심이 있는 학생 발굴을 통한
딥러닝의 이론적 배경 강의 및 오픈소스 딥러닝 라이브러리 PyTorch를 활용한 실습

Topics to learn today

1. **Review from last lecture**
   - Problems of ML / Linear Regression
   - Linear Regression with Pytorch
2. **Binary/Multinomial Classification Problem**
   - with Logistic Regression
   - Multinomial Classification with Pytorch
3. **History of Deep Learning**
   - from simple perceptron to CNN
4. **Solving XOR Problem with MLP**
   - Feed forward / Backpropagation
   - Solving Regression and Classification Problem with MLP

Review from Last Lecture



Review from Last Lecture

Hypothesis

Model

Cost

Loss

Optimization

Review from Last Lecture

Hypothesis

Model

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

Cost

Loss

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

Optimization

$$W := W - \alpha \frac{\partial}{\partial W} cost(W)$$

Review from Last Lecture

Hypothesis

Model

$$H(X) = XW$$

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_{11}w_1 + x_{12}w_2 + x_{13}w_3 \\ x_{21}w_1 + x_{22}w_2 + x_{23}w_3 \\ x_{31}w_1 + x_{32}w_2 + x_{33}w_3 \\ x_{41}w_1 + x_{42}w_2 + x_{43}w_3 \\ x_{51}w_1 + x_{52}w_2 + x_{53}w_3 \end{pmatrix}$$

Cost

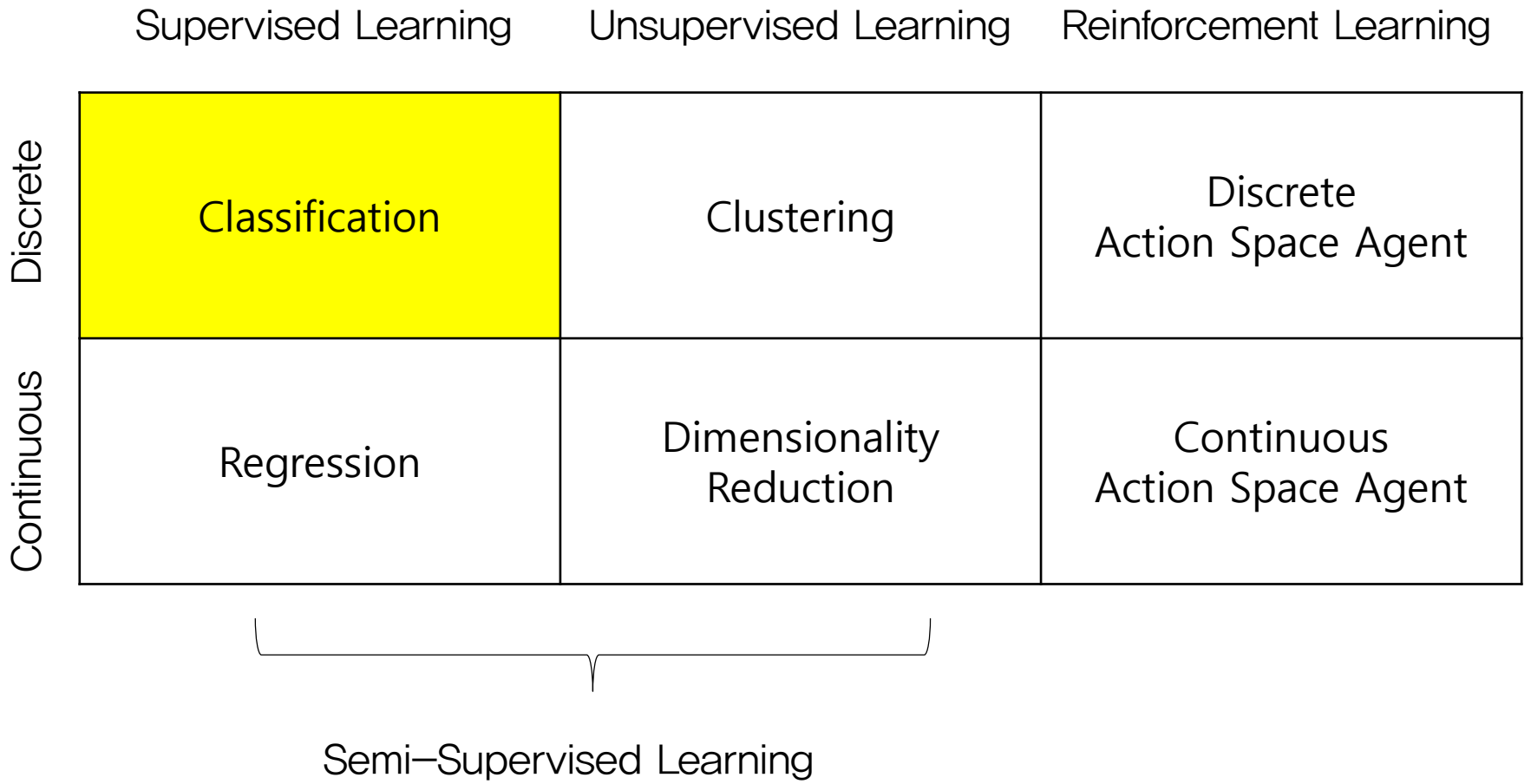
Loss

$$cost(W, b) = \frac{1}{m} \sum_{i=1}^{m} (H(x_1^{(i)}, x_2^{(i)}, ..., x_n^{(i)}) - y^{(i)})^2$$

Optimization

Binary Classification

Binary Classification



Example of Binary Classification

Dog or Cat?

Show this facebook feed to user or not?

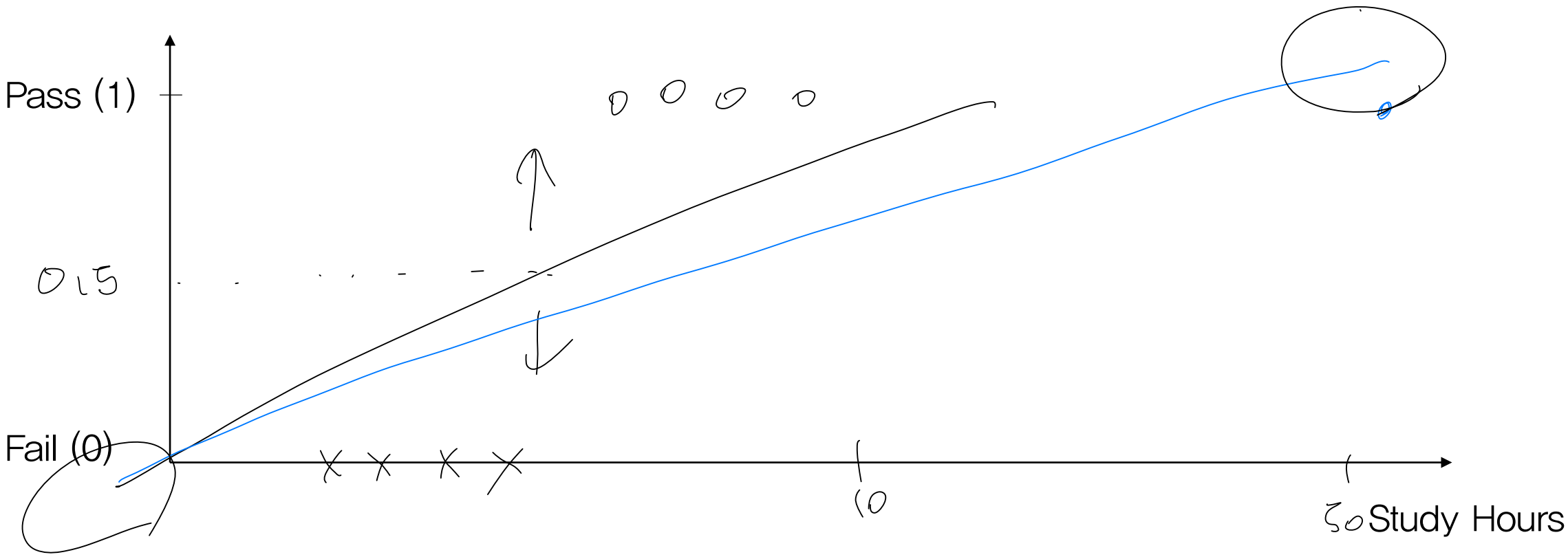
0, 1 encoding

Dog(0), Cat(1)

Show the feed (0)
Do not show the feed (1)

Binary Classification Hypothesis

Pass(1)/Fail(0) based on study hours

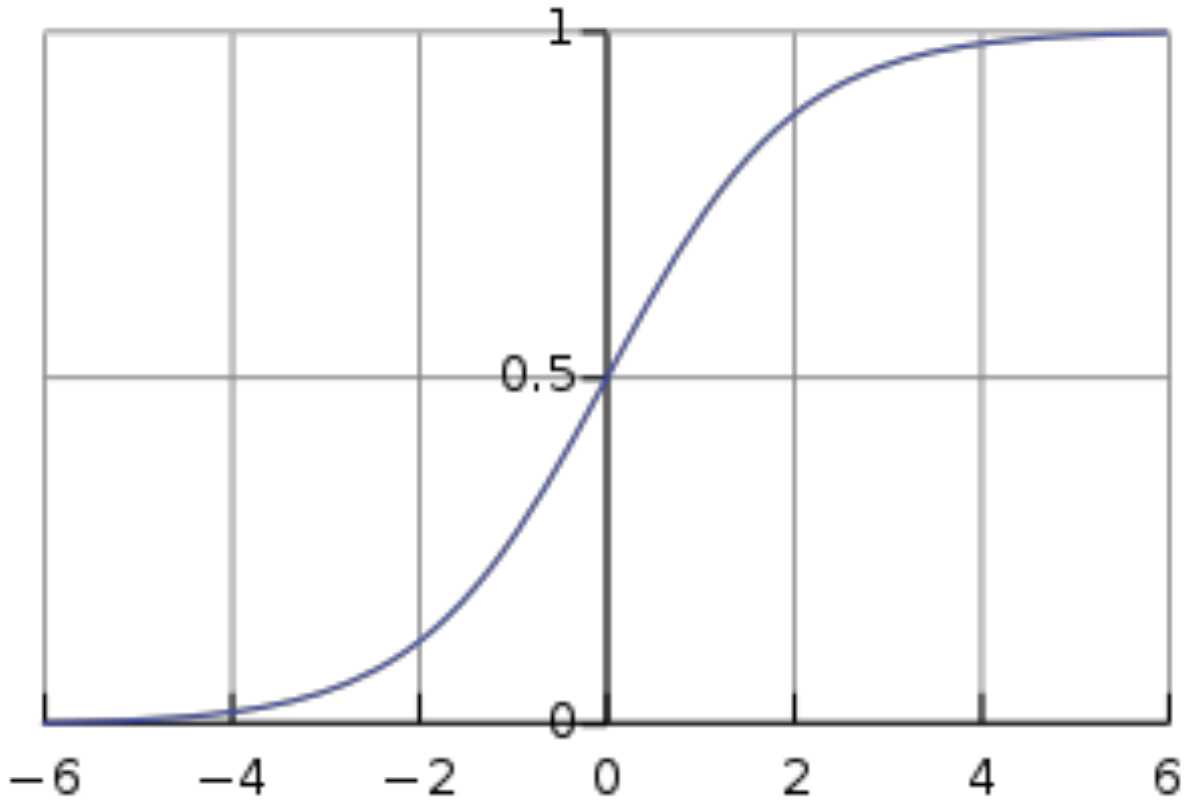


Logistic Hypothesis

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

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

Sigmoid (Logistic) Function



$$G(z) = \frac{1}{1 + e^{-z}}$$

ex) $G(\infty) = \frac{1}{1+e^{-\infty}} = 1$

$G(-\infty) = \frac{1}{1+e^{\infty}} = 0$

Logistic Hypothesis

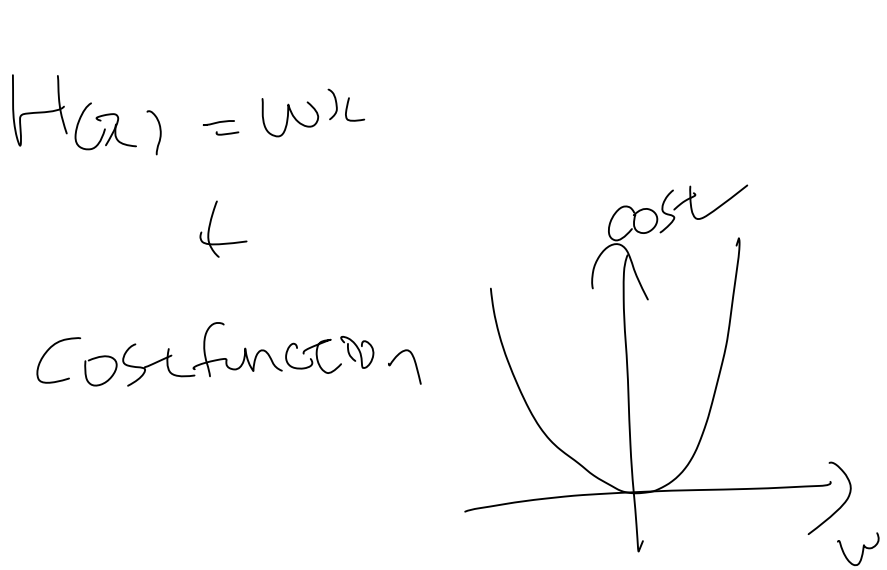
$$H(X) = \frac{1}{1 + e^{-WX}}$$

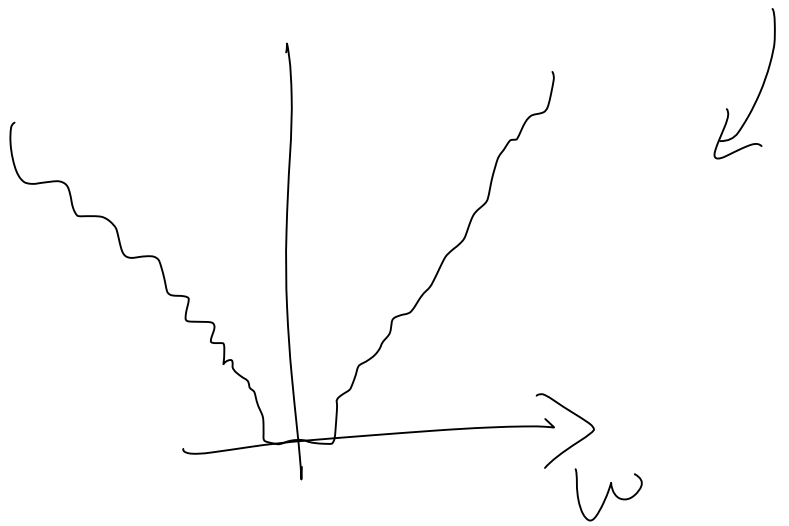
Binary Classification
Cost

Cost Fuction

Can we use the cost function of linear regression?

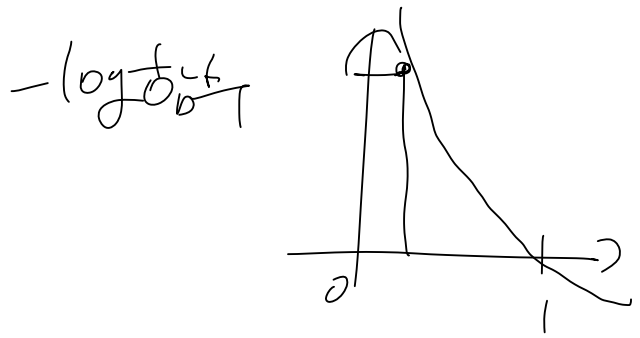
$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2 \qquad H(X) = \frac{1}{1+e^{WX}}$$





Cross Entropy : 예측과 실제값의 차이

Difference between two probability distribution



$$H(P,Q) = -\sum P(x) log(Q(x))$$

실제확률 예측확률

Cost Function

$$cost(W) = \frac{1}{m} \sum_{i=1}^{n} c(H(x^{(i)}), y^{(i)})$$

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

Cost Function

$$cost(W) = \frac{1}{m}\sum_{i=1}^{n} c(H(x^{(i)}), y^{(i)})$$

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

$$c(H(x), y) = -ylog(H(x)) - (1 - y)log(1 - H(x))$$

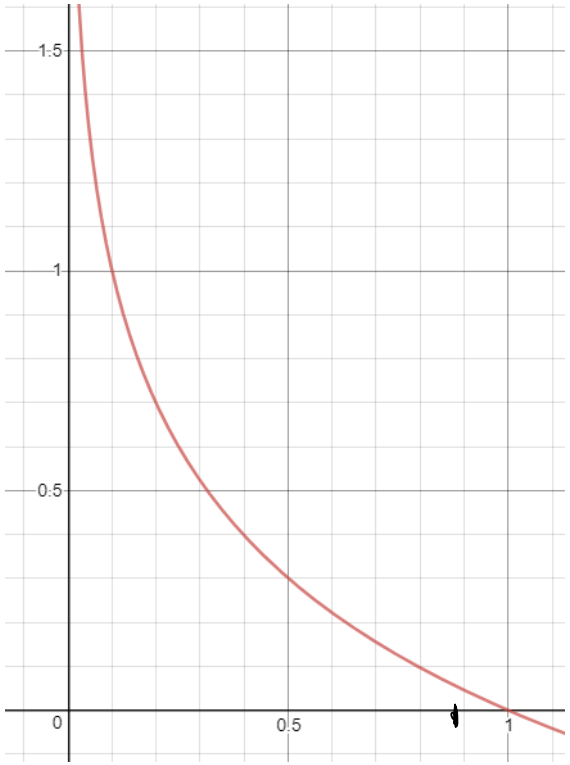
Cost Function

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

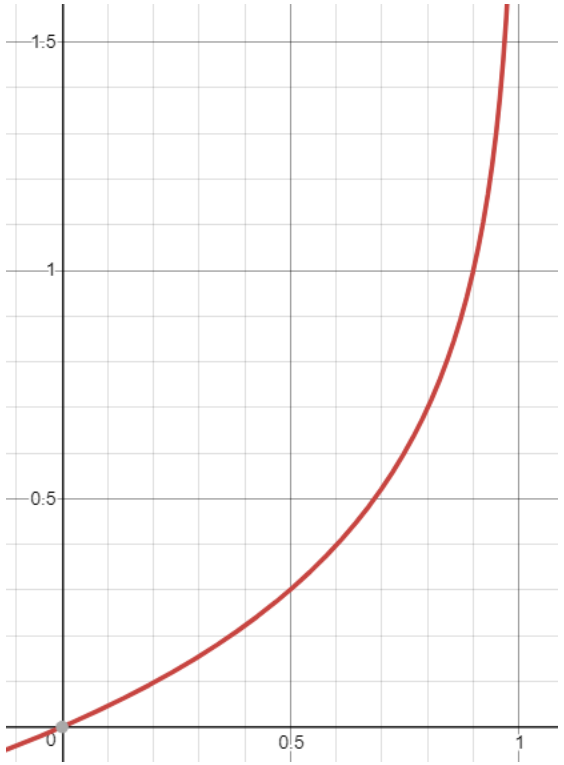
Pass

fail

$-\log x$

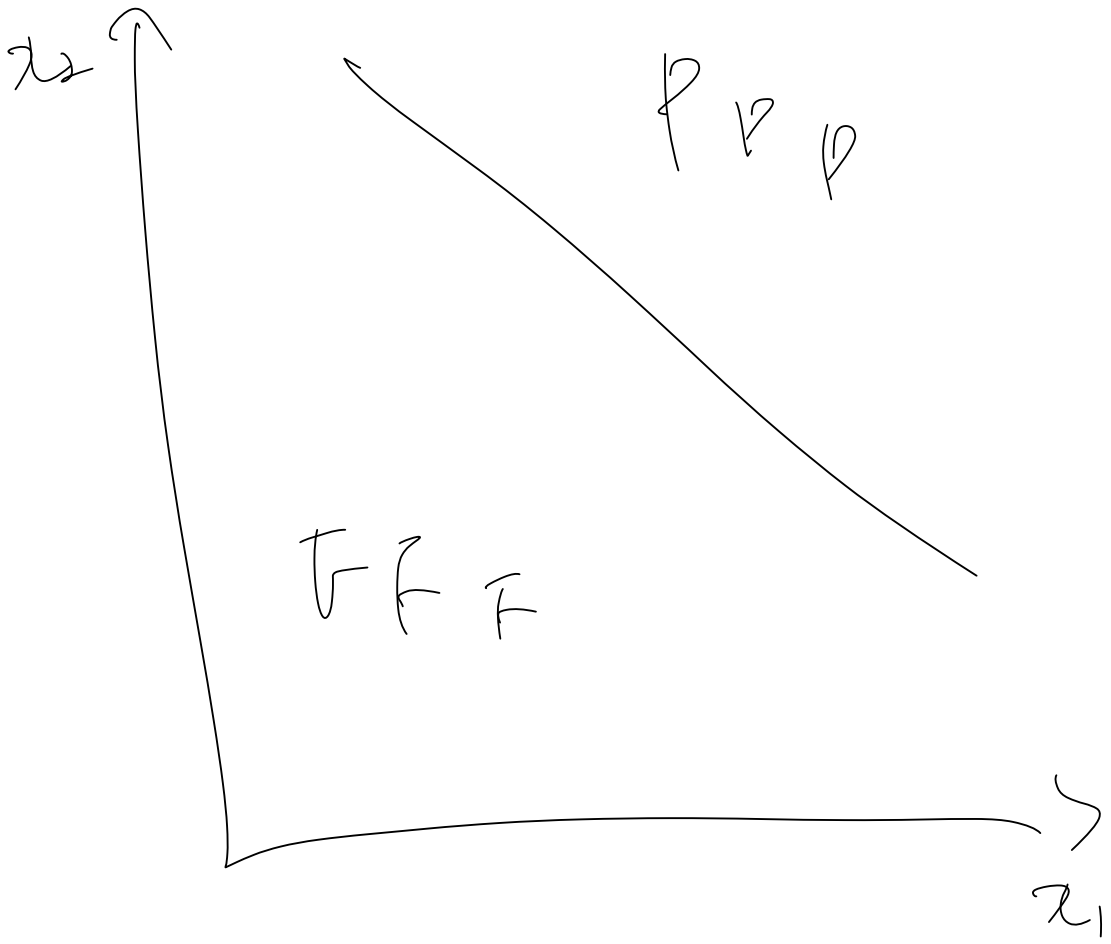


$-\log(1-x)$



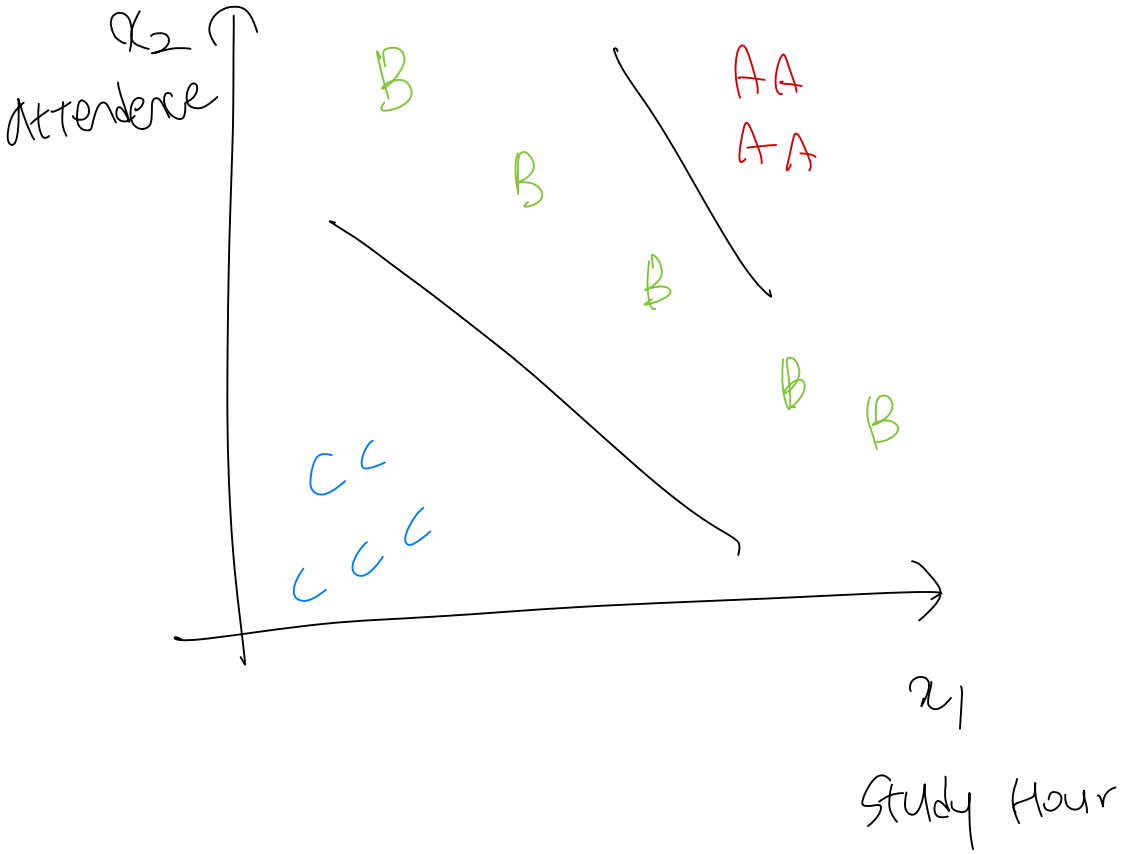
Multinomial Classification

Binary Classification



Goal -

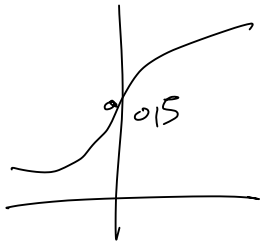
Multinomial Classification



Sigmoid $(Wx+b) = 0.5$

$\Leftrightarrow$ 0

$W_1x_1 + W_2x_2 + b = 0$



$\Rightarrow$ 3개로 분류하기 어렵다 X

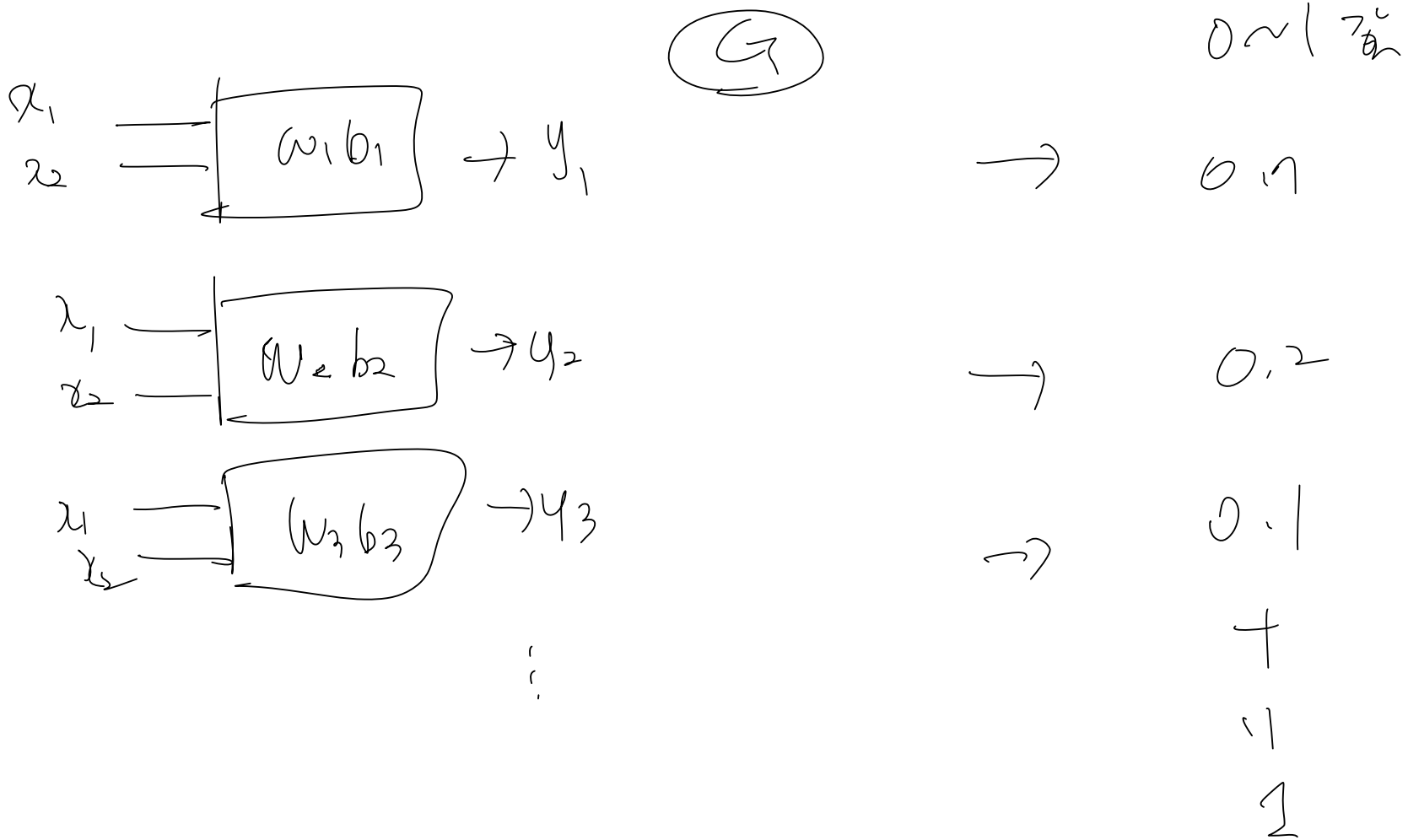
Multinomial Classification Hypothesis

Multinomial Classification

$$H(X) = XW + b$$

Multinomial Classification

$$H(X) = G(XW + b)$$



Softmax function

$$Softmax(\hat{y}_i) = \frac{e^{\hat{y}_i}}{\sum_j e^{\hat{y}_j}}$$

$$= \frac{e^{y_1}}{e^{y_1} + e^{y_2} + e^{y_3}}$$: 세 개의 출력 값

Softmax function

Multinomial Classification Cost

Cost function

pred_y	true_y
S(y)	y
0.7	1
0.2	0
0.1	0

$$Cost(W, b) = \sum_i y_i log(S(WX + b))$$

$$\hat{y} = WX + b$$

Binary and Multinomial Classification

	Binary	Multinomial
Hypothesis	$Sigmoid(WX)$	$Softmax(WX)$
Cost	$-ylog(H(x)) - (1-y)log(1-H(x))$	$\sum -ylog(H(X))$