

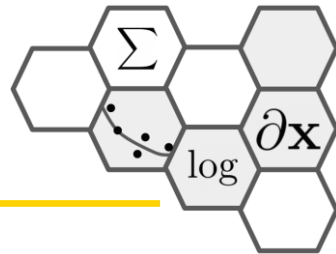
LG U+ Why Not SW 캠프 6기 Python 데이터 분석 I

Softmax Regression (N-class)

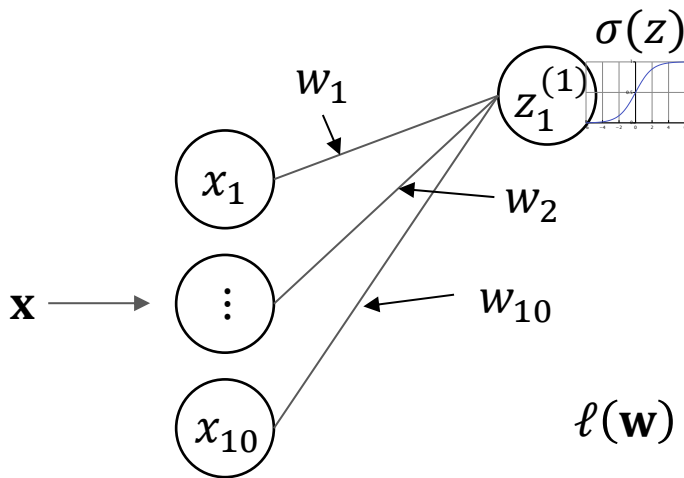
조준우

metamath@gmail.com

Logistic Regression



- 로지스틱 회귀



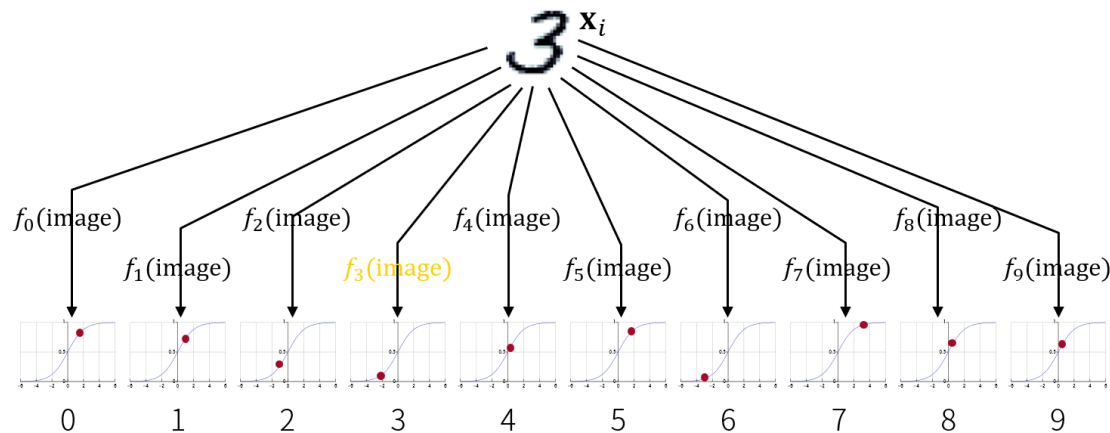
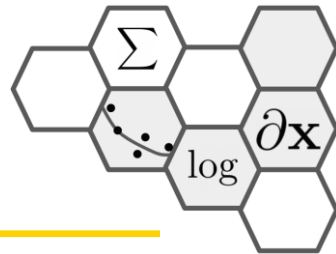
$$\ell(\mathbf{w}) = \frac{1}{2} \sum_{i=1}^N (y_i - \sigma(z_i))^2$$

$$\ell(\mathbf{w}) = - \sum_{i=1}^N y_i \log \sigma(z_i) + (1 - y_i) \log(1 - \sigma(z_i))$$

$$f_1: \mathbb{R}^{10} \rightarrow \mathbb{R}$$

$$f_2: \mathbb{R} \rightarrow \mathbb{R}$$

10진 분류 손실함수: BCE??



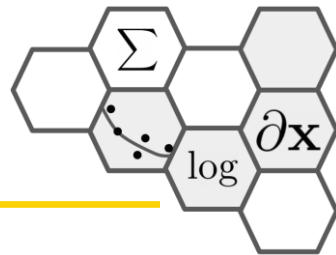
$$\ell(\mathbf{w}) = - \sum_{i=1}^N y_i \log \sigma(z_i) + (1 - y_i) \log(1 - \sigma(z_i))$$

linear function

logistic sigmoid

i		
0		2
1		0
2		3
3		5
4		4
	\vdots	\vdots

2진 분류 손실함수: BCE



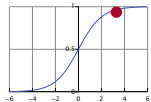
- $f_3(\text{image}) \approx 1, 3$ 인가? 3이 아닌가?

Binary Cross Entropy

$$\ell(\mathbf{w}) = - \sum_{i=1}^N y_i \log \sigma(z_i) + (1 - y_i) \log(1 - \sigma(z_i))$$

0 or 1

$f_3(\text{image})$



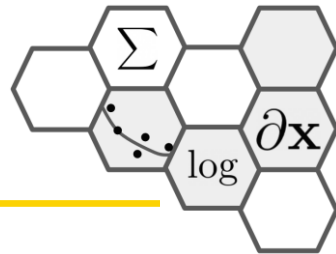
3

$\log \sigma(z_{i3})$

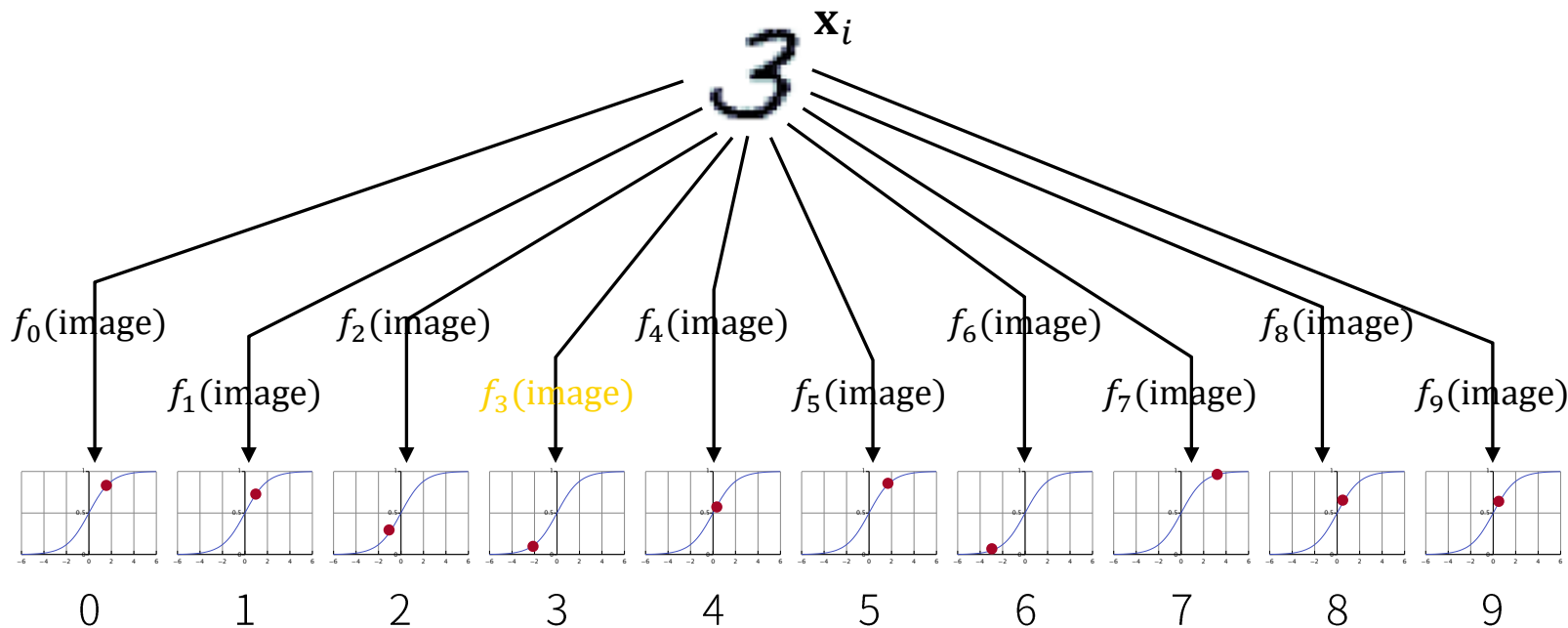
\mathbf{x}_i

i	\mathbf{x}		y
0		2	0
1		0	0
2		3	1
3		5	0
4		4	0
	\vdots		

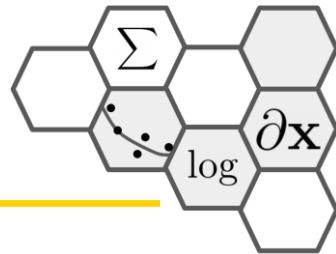
여러 함수를 사용한 분류기



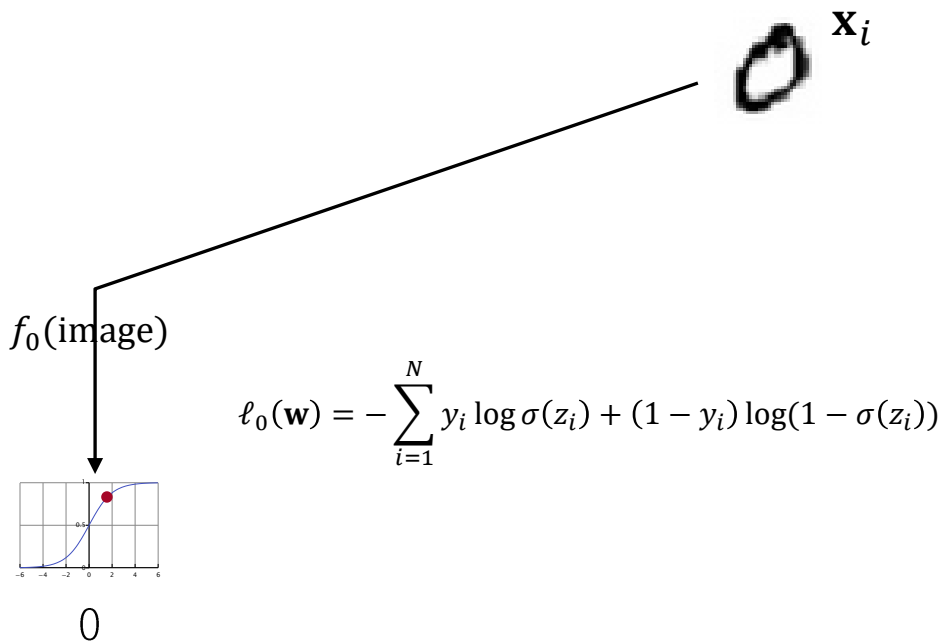
- $f_3(\text{image}) \approx 1$, $f_{i \neq 3}(\text{image}) \approx 0$



여러 함수를 사용한 분류기

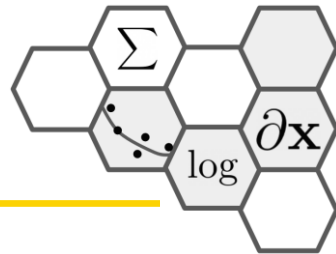


- $f_3(\text{image}) \approx 1, f_{i \neq 3}(\text{image}) \approx 0$



i	\mathbf{x}		y
0		2	0
1		0	1
2		3	0
3		5	0
4		4	0
		\vdots	

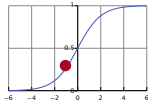
여러 함수를 사용한 분류기



- $f_3(\text{image}) \approx 1, f_{i \neq 3}(\text{image}) \approx 0$

$$\ell_2(\mathbf{w}) = - \sum_{i=1}^N y_i \log \sigma(z_i) + (1 - y_i) \log(1 - \sigma(z_i))$$

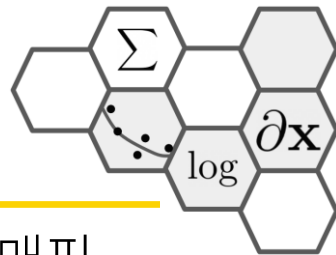
$f_2(\text{image})$



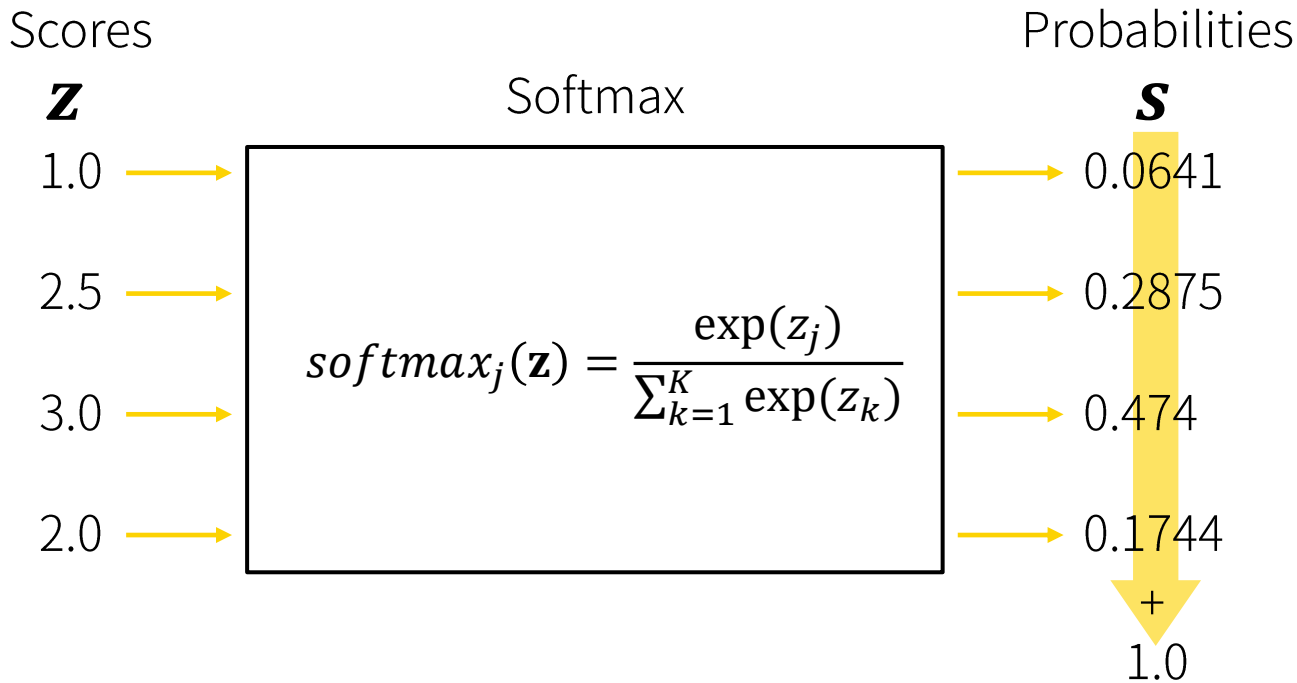
2

i	\mathbf{x}		y
0	<input type="text"/>	2	1
1	<input type="text"/>	0	0
2	<input type="text"/>	3	0
3	<input type="text"/>	5	0
4	<input type="text"/>	4	0
	\vdots		

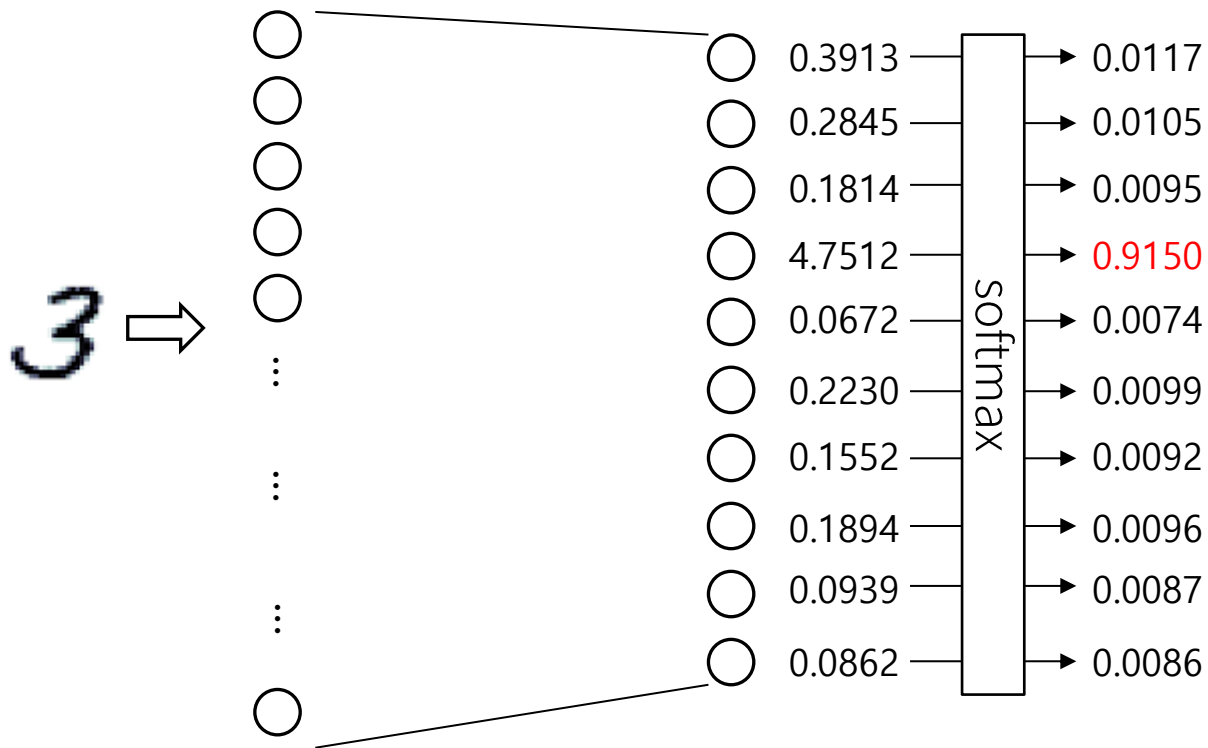
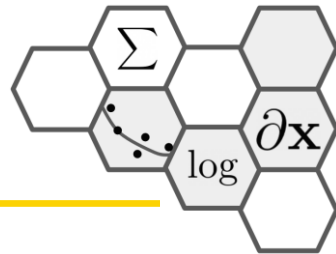
소프트맥스 함수



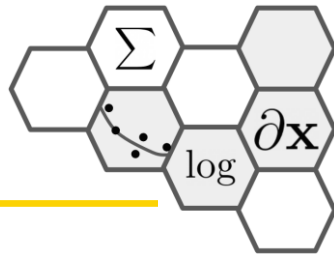
- $\text{softmax}: \mathbb{R}^K \rightarrow [0,1]^K$: 실수 k 개가 0에서 1사이의 숫자 k 개로 매핑



소프트맥스 함수 활용



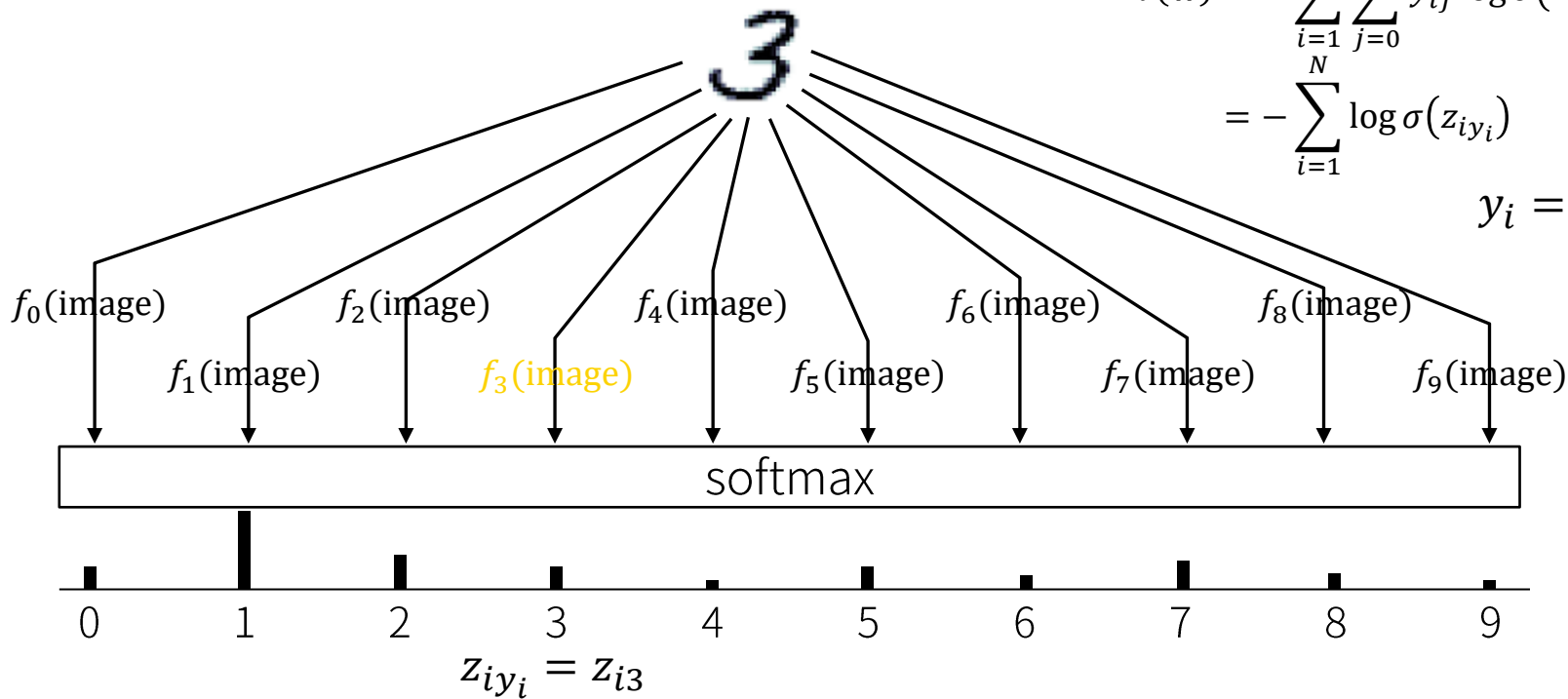
다클래스 분류 손실함수: CE



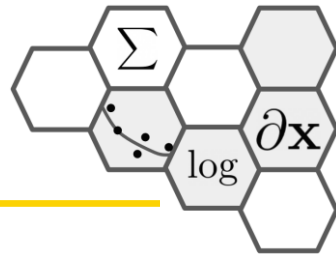
- BAD : $f_3(\text{image})$ 함수가 1에 가까운 값을 출력해야 함

$$\begin{aligned}\ell(\mathbf{w}) &= - \sum_{i=1}^N \sum_{j=0}^{K-1} y_{ij} \log \sigma(z_{ij}) \\ &= - \sum_{i=1}^N \log \sigma(z_{iy_i})\end{aligned}$$

$$y_i = 3$$



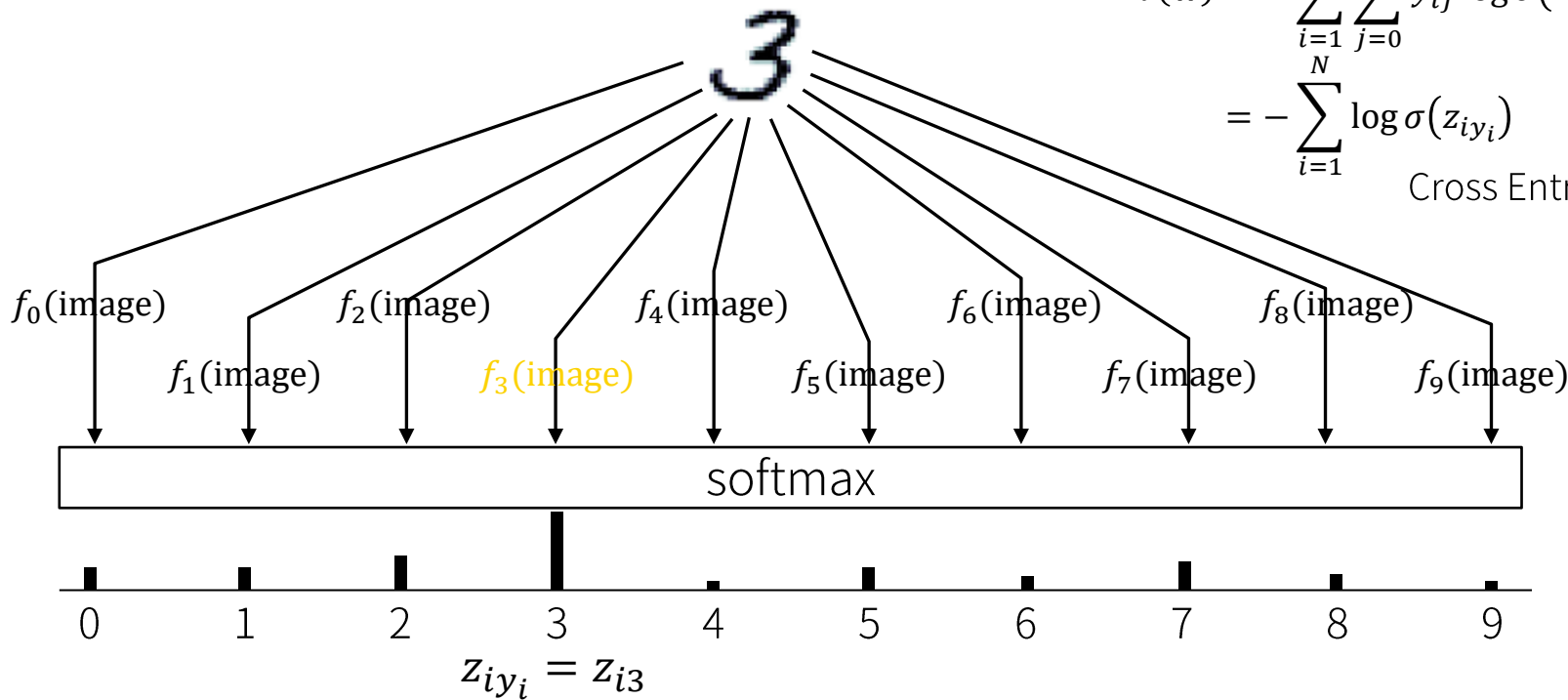
소프트맥스 함수 활용



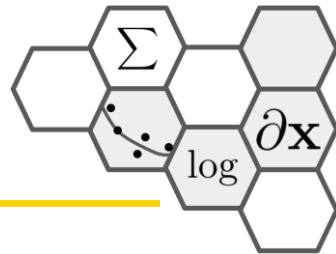
- BAD : $f_3(\text{image})$ 함수가 1에 가까운 값을 출력해야 함

$$\begin{aligned}\ell(\mathbf{w}) &= - \sum_{i=1}^N \sum_{j=0}^{K-1} y_{ij} \log \sigma(z_{ij}) \\ &= - \sum_{i=1}^N \log \sigma(z_{iy_i})\end{aligned}$$

Cross Entropy



Summary



Regression: Mean Squared Error, target y : real number

2-Classification: Binary Cross Entropy, target y : 0, 1

n-Classification: Cross Entropy, target y : 0, 1, 2, \dots

