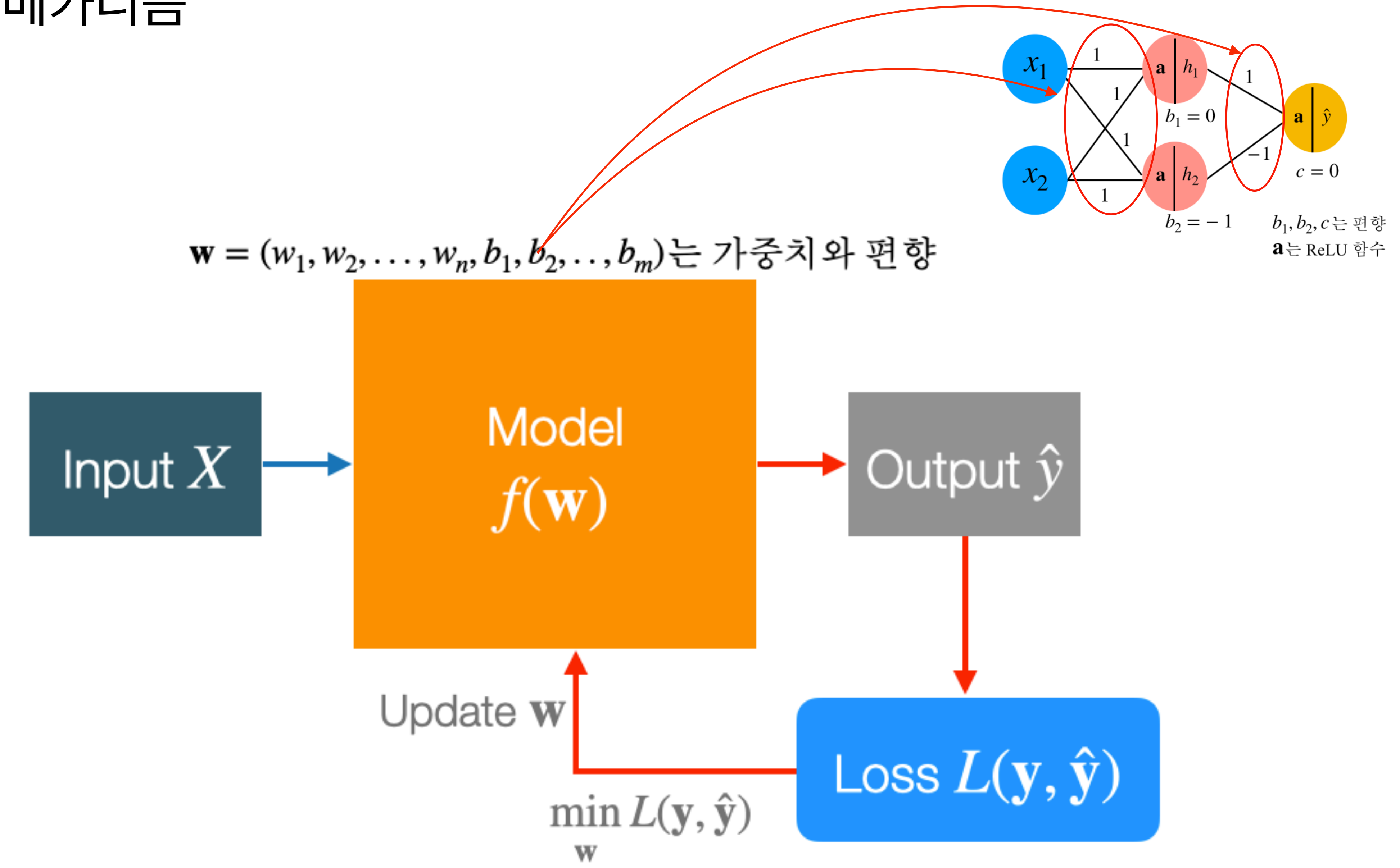
# 딥러닝 올인원

## 손실 함수
## 7강

딥러닝호형

# 손실 함수(Loss Function)

## 지도학습의 메카니즘



# 손실 함수(Loss Function)

## Classification의 대표적인 함수

알아두기 2.3.6 — **내적(inner product).**

두 벡터 $a=(a_1,a_2,...,a_n)$, $b=(b_1,b_2,...,b_n)$라고 하면

$$a \cdot b = a_1b_1 + a_2b_2 + ... + a_nb_n$$

내적은 $\cdot$라고 표기하며 각 성분의 곱의 합이다. 예를 들어 $a=(1,2,4), b=(0,1,2)$라고 하면 $a \cdot b = 1\cdot 0 + 2\cdot 1 + 4\cdot 2 = 10$이다.

알아두기 2.3.7 — **교차 엔트로피 함수(cross entropy function).**

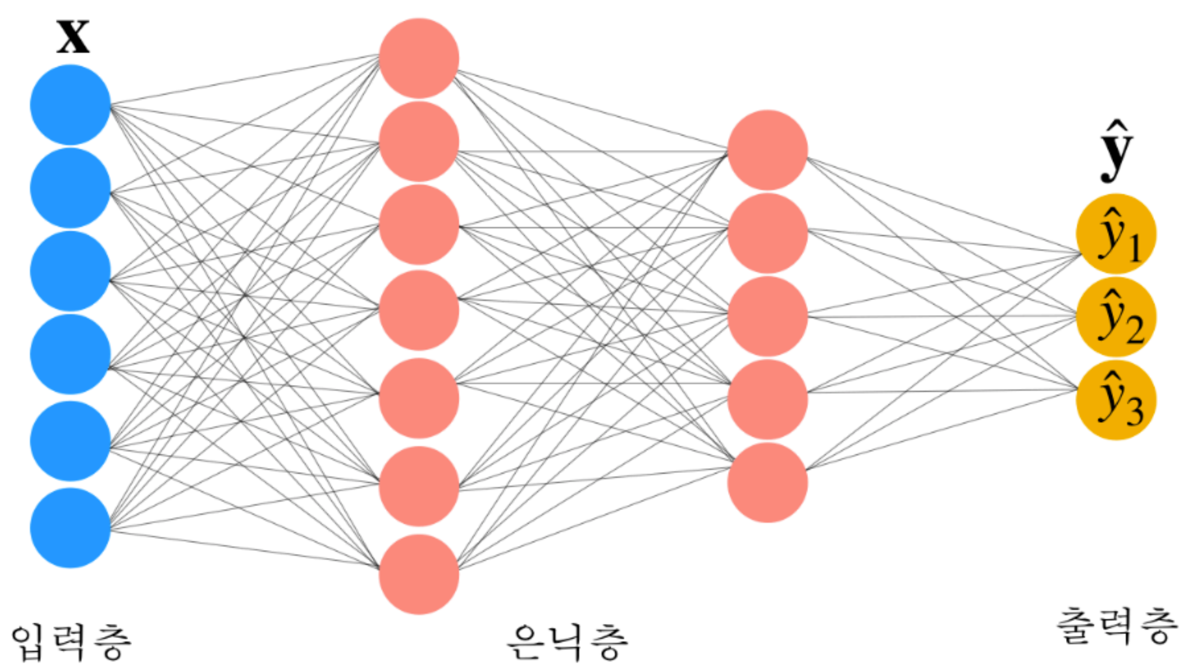
분류 문제를 위한 대표적인 손실 함수다.

$$L(\mathbf{y},\hat{\mathbf{y}}) = -\frac{1}{n}\sum_{i=1}^{n} \mathbf{y}_i \cdot log(softmax(\hat{\mathbf{y}})) \ (log\text{는 자연로그})$$

$\cdot$은 내적을 의미하며 실제값 $\mathbf{y}$는 원-핫 벡터다.

# 손실 함수(Loss Function)

강아지, 소, 호랑이 분류 문제



# 손실 함수(Loss Function)

## 강아지, 소, 호랑이 분류 문제

$$L(\mathbf{y},\hat{\mathbf{y}}) = -\frac{1}{n}\sum_{i=1}^{n}\mathbf{y}_i \cdot log(softmax(\hat{\mathbf{y}}))$$ ($log$는 자연로그)

| | $\hat{\mathbf{y}}$ |
|---|---|
| $\hat{\mathbf{y}}_1$ | (1,2,1) |
| $\hat{\mathbf{y}}_2$ | (1,4,2) |
| $\hat{\mathbf{y}}_3$ | (0,0,1) |
| $\hat{\mathbf{y}}_4$ | (1,0,2) |

$$softmax(\hat{y}) = \frac{e^{\hat{y}}}{\sum_{k=1}^{n} e^{\hat{y}_k}}$$

$$\hat{\mathbf{y}}_i = (\hat{y}_1, \hat{y}_2, \hat{y}_3)$$

| $softmax(\hat{\mathbf{y}})$ |
|---|
| (0.212,0.576,0.212) |
| (0.042, 0.844,0.114) |
| (0.212, 0.212, 0.576) |
| (0.245,0.090,0.665) |

| $log(softmax(\hat{\mathbf{y}}))$ |
|---|
| (-1.551,-0.552,-1.551) |
| (-3.170, -0.170,-2.172) |
| (-1.551, -1.551, -0.552) |
| (-1.406,-2.408,-0.408) |

●

| $\mathbf{y}$ | |
|---|---|
| (0,0,1) | $\mathbf{y}_1$ |
| (0,1,0) | $\mathbf{y}_2$ |
| (0,0,1) | $\mathbf{y}_3$ |
| (1,0,0) | $\mathbf{y}_4$ |

내적

| $log(softmax(\hat{\mathbf{y}})) \cdot \mathbf{y}$ |
|---|
| -1.551 |
| -0.170 |
| -0.552 |
| -1.406 |

$$-\frac{1}{4}\sum$$

0.9198

# 손실 함수(Loss Function)

## Classification의 대표적인 함수

알아두기 2.3.8 — 이진 교차 엔트로피 함수(**binary cross entropy function**).

이진 분류 문제를 위한 교차 엔트로피 함수다. 이진 분류의 특성을 이용해 알아두기 2.3.8에서 소개된 함수를 아래와 같이 간단한 식으로 유도할 수 있다.

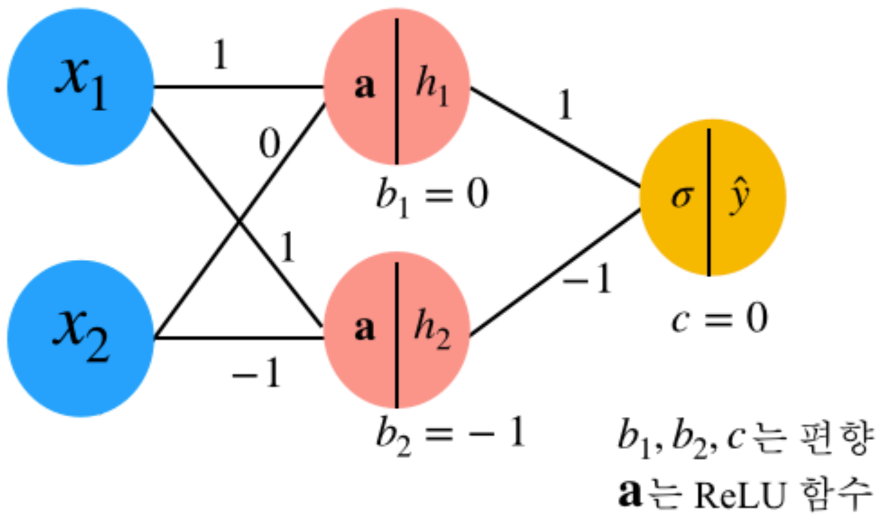
$$L(\mathbf{y}, \hat{\mathbf{y}}) = -\frac{1}{n}\sum_{i=1}^{n} y_i log(\hat{y}_i) + (1 - y_i) log(1 - \hat{y}_i)) \text{ } (log\text{는 자연로그})$$

이진 분류는 두 가지 상황을 분류하는 문제로서 원-핫 벡터로 표현하지 않아도 된다. 예를 들어 자동차 보유를 분류하는 문제라면 미보유는 0, 보유는 1으로 표현하여 벡터(e.g. (1,0),(0,1))를 사용하지 않고 표현할 수 있다. 따라서 예측을 하는 단계에서 소프트맥스를 사용하지 않고 일반적으로 시그모이드를 사용한다.

# 손실 함수(Loss Function)

## 차량 보유 유무 문제

| $x_1$ | $x_2$ | $y$ |
|---|---|---|
| 1 | 1 | 0 |
| 2 | 0 | 1 |
| 2 | 1 | 1 |




$$B = ReLU(XH^T + b) = ReLU\left(\begin{pmatrix}1 & 1\\ 2 & 0\\ 2 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & -1\end{pmatrix} + \begin{pmatrix}0 & -1\end{pmatrix}\right) = \begin{pmatrix}1 & 0\\ 2 & 1\\ 2 & 0\end{pmatrix}$$

$$\hat{\mathbf{y}} = \sigma(B\mathbf{v}^T + c) = \sigma\left(\begin{pmatrix}1 & 0\\ 2 & 1\\ 2 & 0\end{pmatrix}\begin{pmatrix}1\\ -1\end{pmatrix} + 0\right) = \begin{pmatrix}0.731\\ 0.731\\ 0.881\end{pmatrix} \rightarrow \begin{pmatrix}1\\ 1\\ 1\end{pmatrix}, \mathbf{y} = \begin{pmatrix}0\\ 1\\ 1\end{pmatrix}$$

$$L(\mathbf{y},\hat{\mathbf{y}}) = -\tfrac{1}{3}((1-0)\cdot log(1-0.731) + 1\cdot log(0.731) + 1\cdot log(0.881)) = 0.584$$