

Chapter 02. 가장 단순한 신경망을 통해 작동 원리 이해하기

STEP2. 다중 분류 문제

다중 클래스 분류 Multi-Class Classification



“고양이가 맞니?”
이진 분류

“네!”



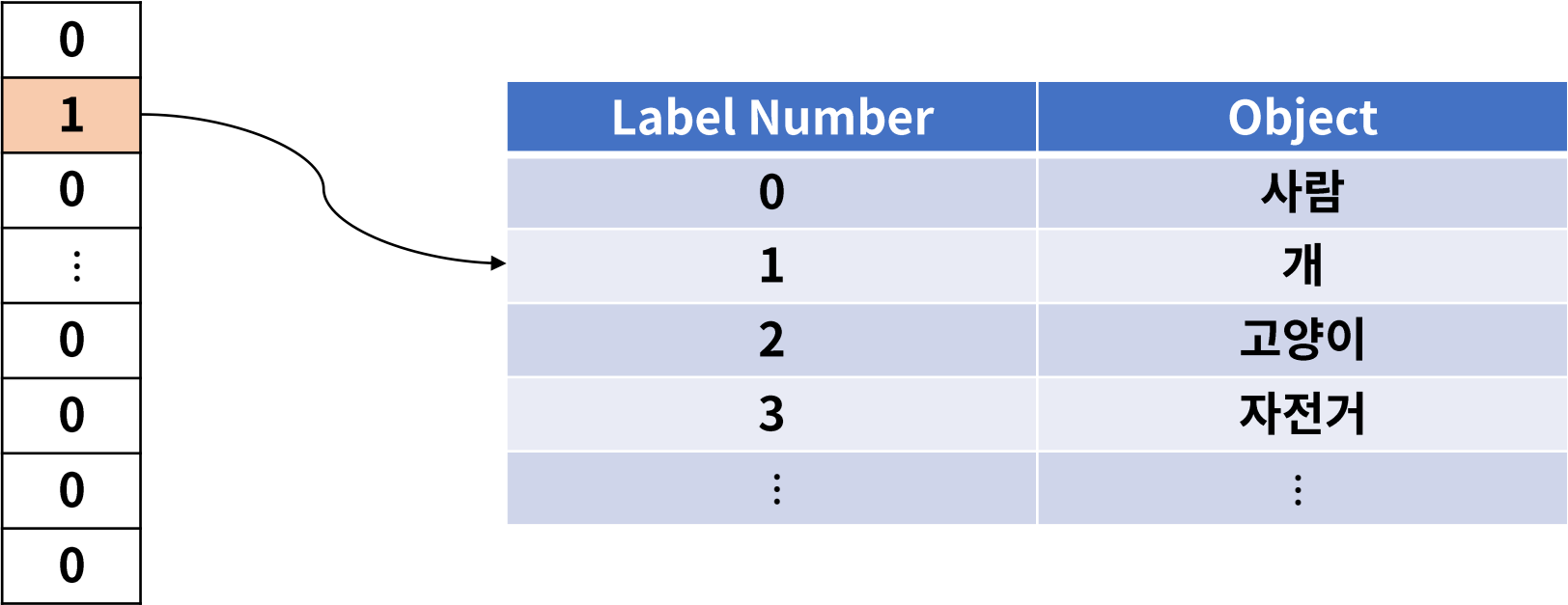
“이게 뭐니?”
다중 분류

“개!”

이진 분류에서는 ‘어떤 물체’인지 표현할 필요가 없으나, 다중 분류에서는 ‘어떤 물체’인지 표현해야 한다.

정답을 어떻게 표현할 것인가? 원-핫 인코딩

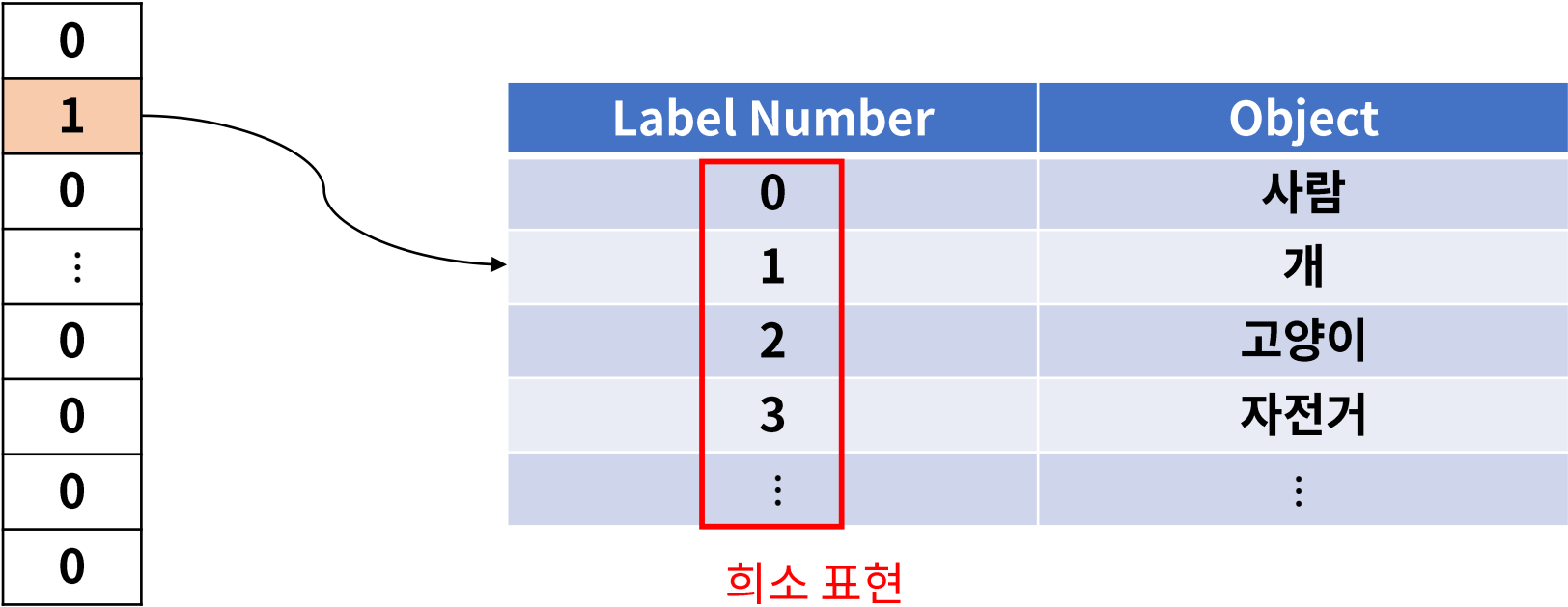
One-Hot Encoding : 한 개의 값만 1이고, 나머지 값은 0인 벡터로 표현하는 기법



원 핫 인코딩된 벡터는 미리 정해 둔 Table을 이용해 어떤 물체인지 알 수 있다.

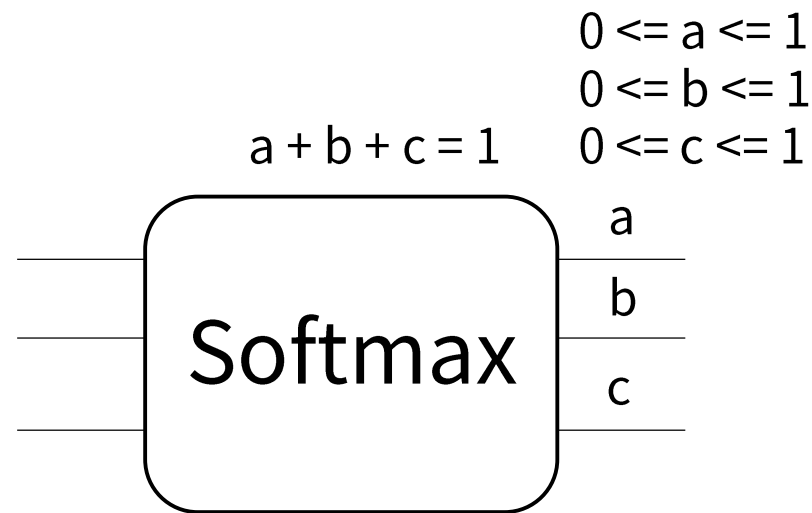
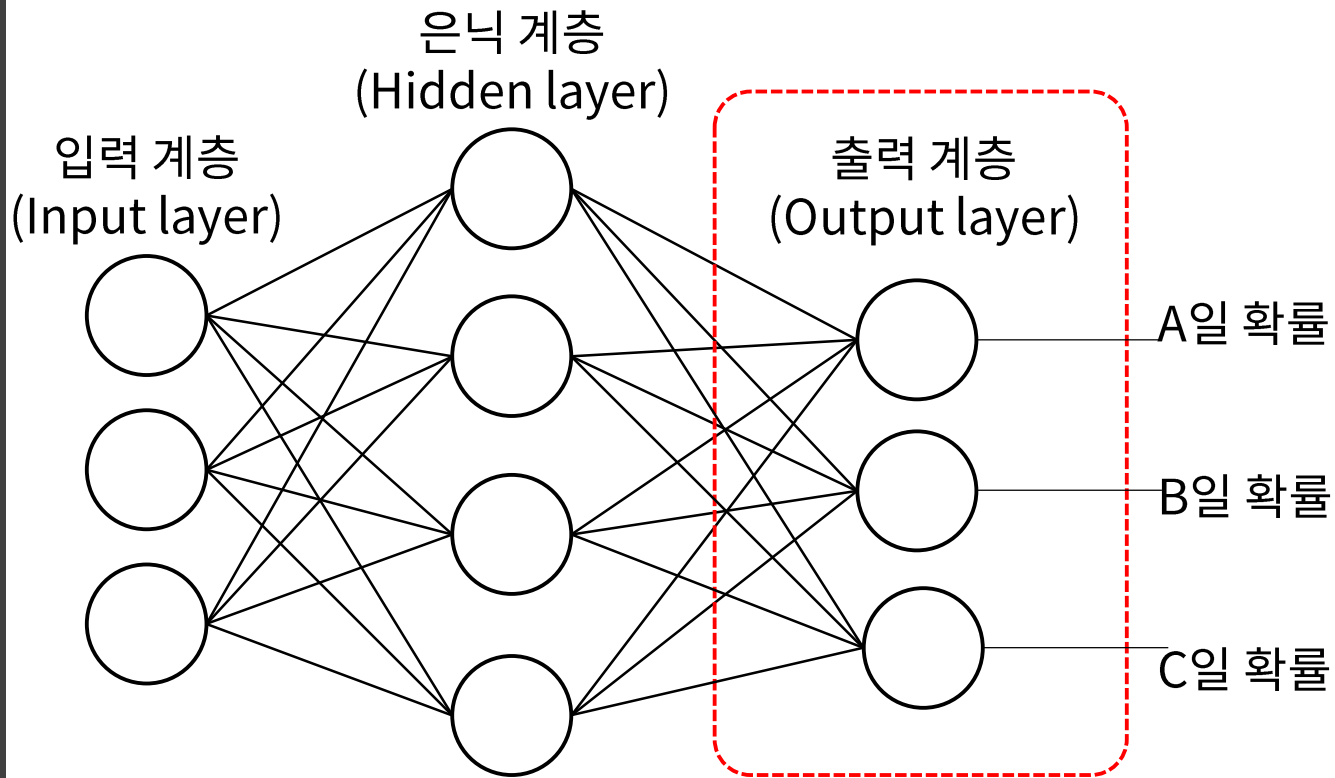
원-핫 인코딩의 희소 표현

희소 벡터(Sparse Vector): 대부분의 값이 0이고 크기가 있는 값이 희소하게 나타나는 벡터



희소 표현을 이용해 벡터 전체를 표기하지 않고, 숫자 하나로 표현할 수 있다.

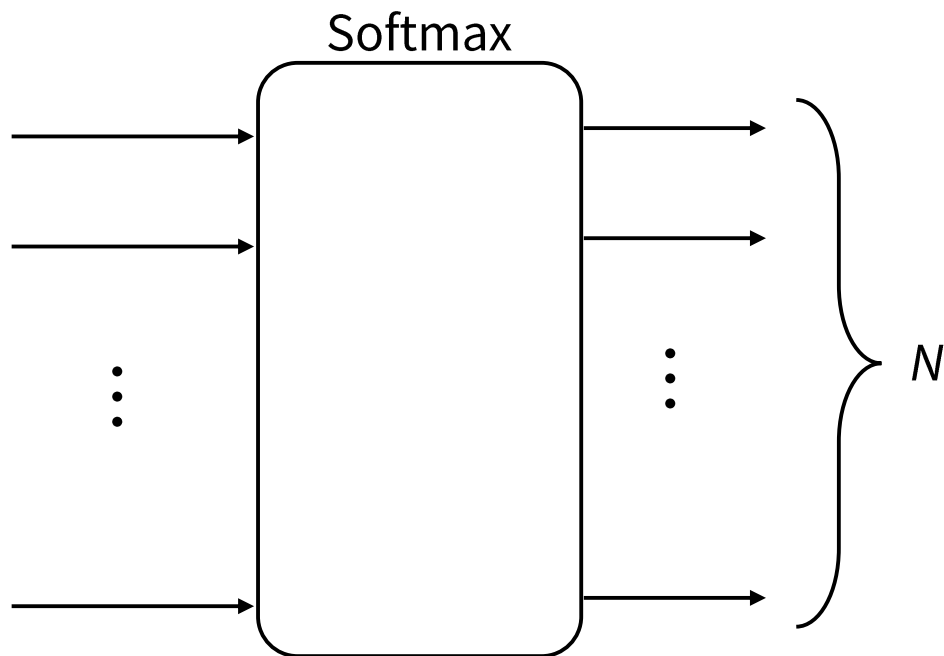
얇은 신경망을 이용한 다중 클래스 분류



한가지 확률만 100%이고 나머지는 0%인 것이 정답 (= 원-핫 인코딩)

실제 알고리즘 출력을 확률로 변환하기 위해 Softmax 함수를 사용한다.

어떻게 출력을 계산할 것인가? Softmax Function

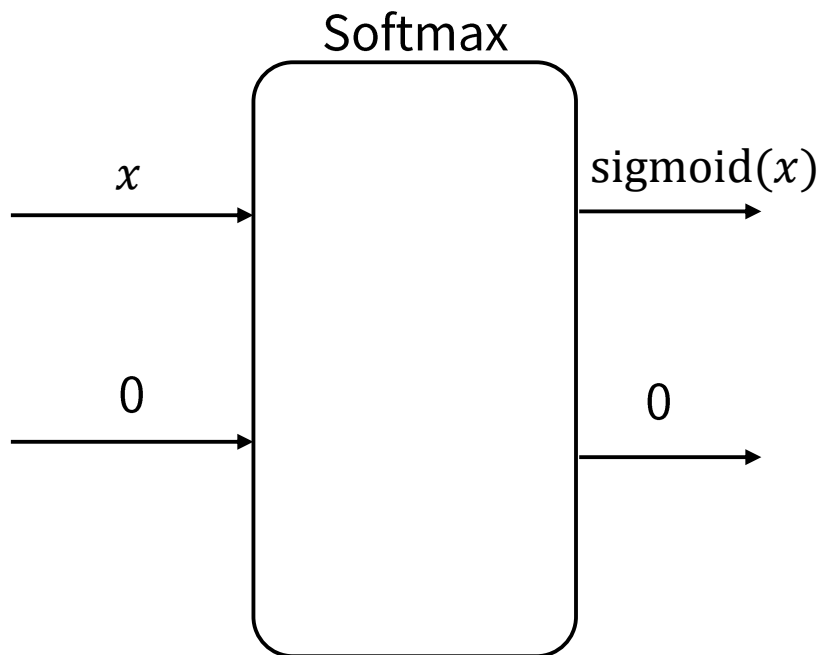


$$\text{softmax}(\mathbf{x})_i = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

- 각 입력의 지수함수를 정규화한 것
- 각 출력은 0~1 사이의 값을 가짐
- 모든 출력의 합이 반드시 1이 됨
- 여러 경우의 수 중 한가지에 속할 ‘확률’을 표현

Softmax는 최종 출력 단에서 **N가지 범주로 분류하는 Multi-class classification**에 쓰임

Softmax vs. Sigmoid



$$\text{softmax}([x, 0])_0 = \frac{e^x}{e^x + e^0} = \frac{1}{1 + e^{-x}} \quad \text{sigmoid}(x)$$

Sigmoid는 하나의 입력을 0으로 강제한 2-Class Softmax 함수와 동일하다.
2가지 클래스를 구분하기 위해 1개의 입력을 받는다는 점에 주목.

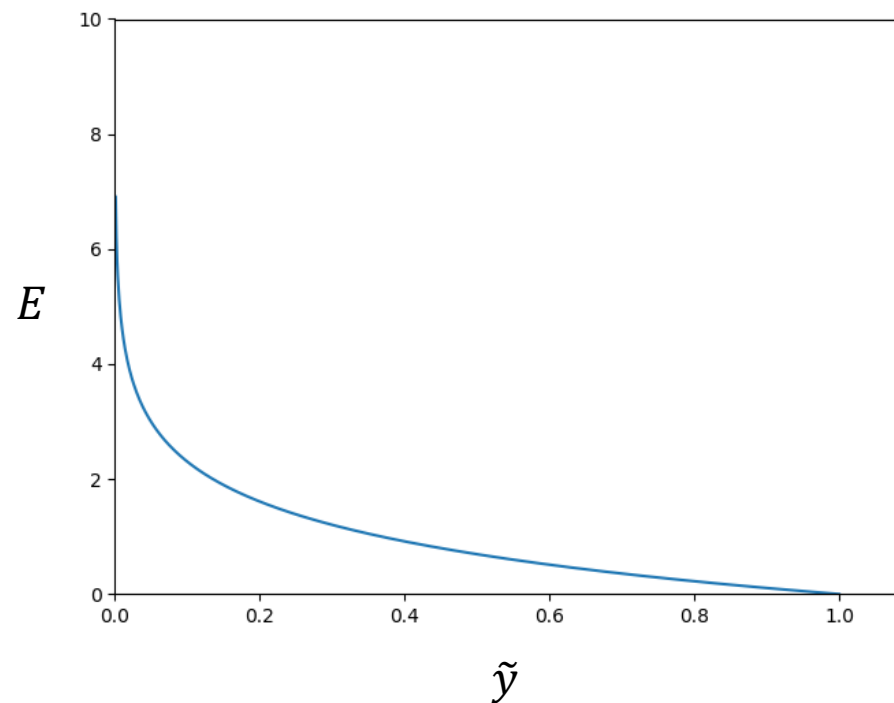
정답과 출력을 어떻게 비교할까? 교차 엔트로피 오차

교차 엔트로피 오차 (Cross entropy error; CEE)

$$E = - \sum_i y_i \log \tilde{y}_i$$

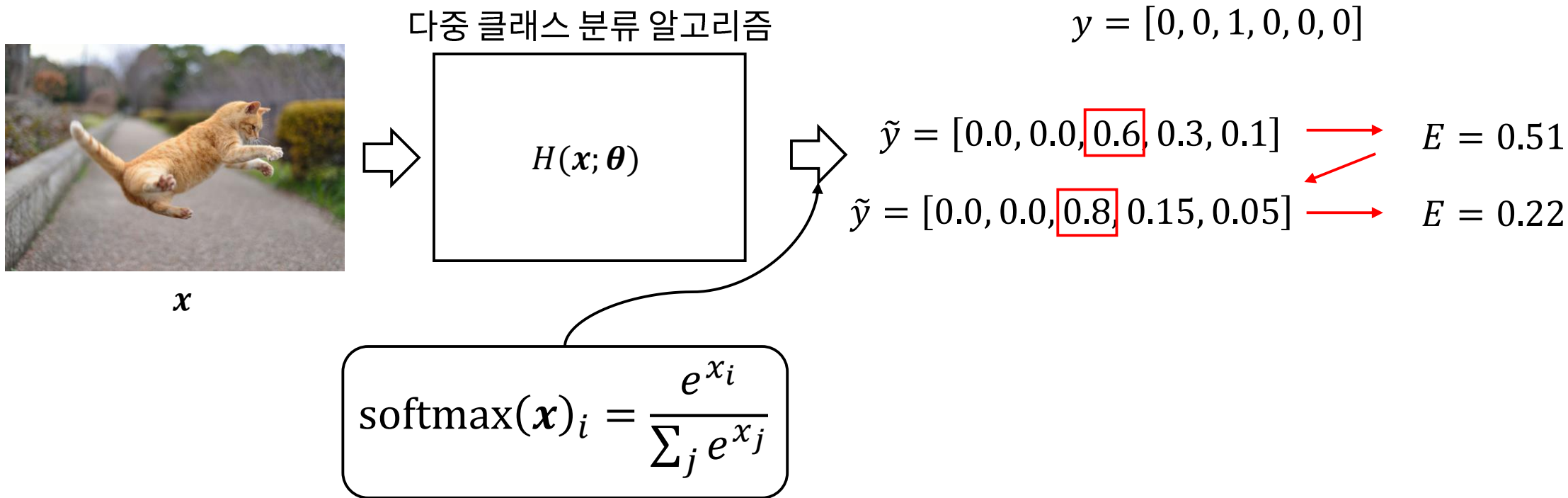
y_i : 학습 데이터 정답의 i 번째 요소 (원-핫 인코딩)

\tilde{y}_i : 학습 데이터 입력으로 추정한 출력의 i 번째 요소



원-핫 인코딩으로 인해, **정답인 클래스에 대해서만 오차를 계산**. 정확히 맞추면 오차가 0, 틀릴수록 오차가 무한히 증가하는 특징이 있다.

교차 엔트로피 오차 Cross Entropy Error



오차를 내는 과정에서는 정답 클래스만 비교하지만, 다중 클래스 분류의 활성화함수인 Softmax로 인해 다른 클래스에 대한 학습에도 영향을 준다.