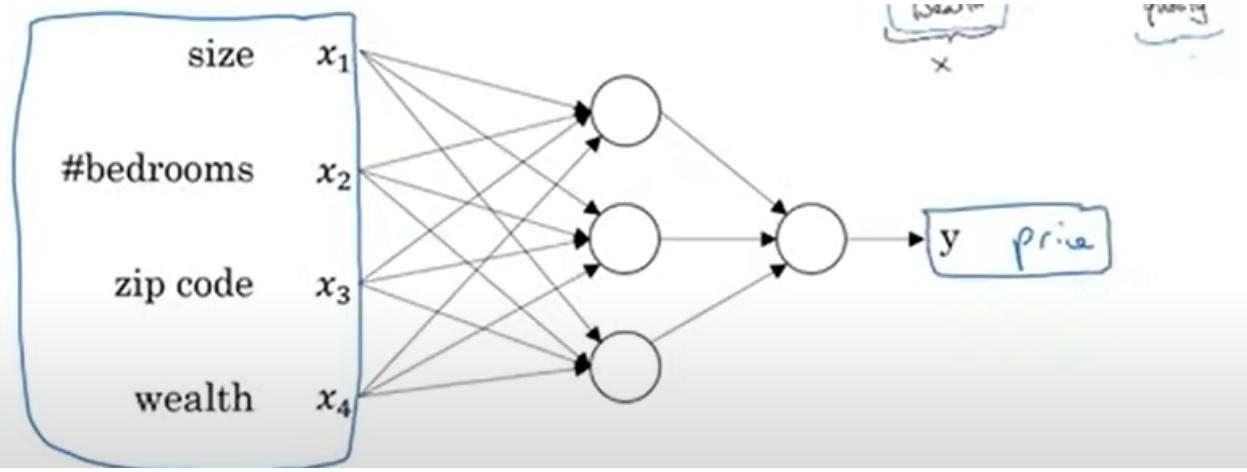
[Week1]_문가을

1. 신경망과 딥러닝

딥러닝 소개

- 딥러닝 : 신경망을 학습시키는 것
- 신경망 : 입력(x)와 출력(y)를 매칭해주는 함수를 찾는 과정



각각의 원은 노드임. 모든 입력은 중간 노드와 연결됨.

(중간에 있는 3개의 원들을 은닉 유닛이라고 부름)

지도 학습

Input(x)와 Output(y) 데이터로 컴퓨터를 학습시킴

- 이미지 → CNN (합성곱 신경망)
- 오디오, 언어 → RNN (순환 신경망 : 입력과 출력을 시퀀스 단위로 처리하는 시퀀스 모)

데이터 종류

- 구조적 데이터 : 데이터베이스로 표현된 데이터
- 비구조적 데이터 : 음성파일이나 이미지, 텍스트 데이터 (특성 : 픽셀, 단어 등)

딥러닝의 성능이 좋으려면,

1) 신경망의 크기가 커야 하고,

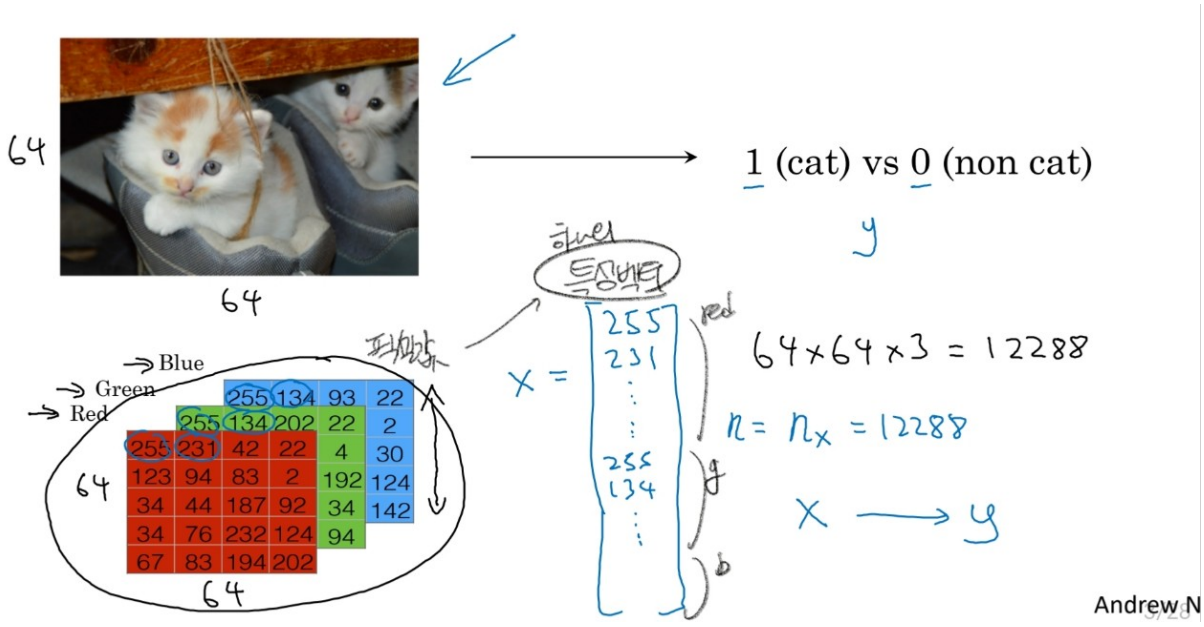
2) 데이터(레이블이 있는 데이터 / m : 데이터 개수)의 양이 많아야 함.

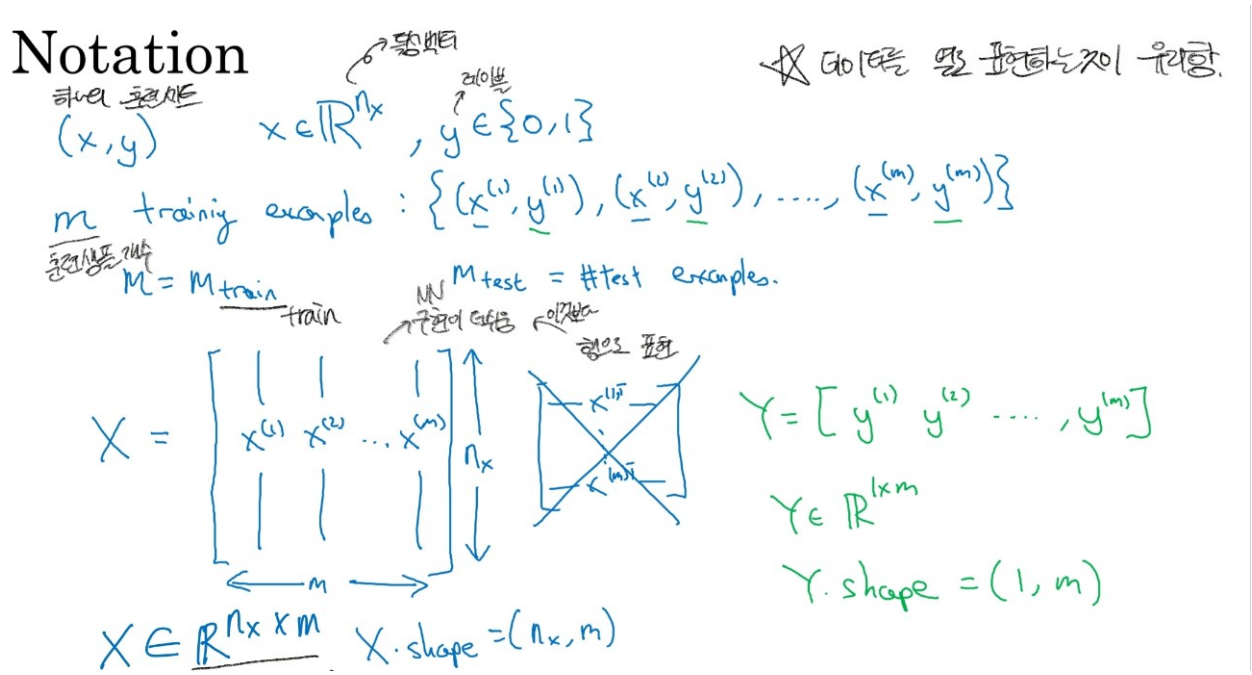
---

2. 신경망과 로지스틱 회귀

이진 분류 (Binary classification)

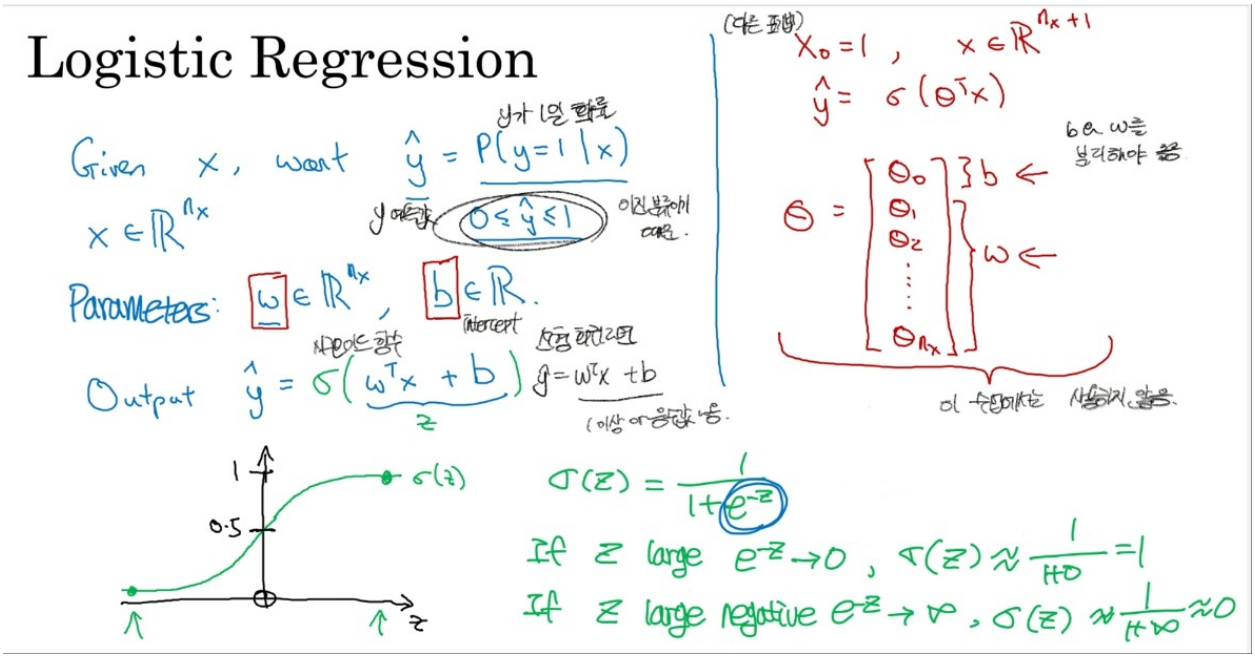
- 예 : 이미지(입력)을 통해 고양이(1)인지 아닌지(0) 분류





로지스틱 회귀

이진 분류 문제에 사용되는 알고리즘



로지스틱 회귀의 비용함수

- **손실 함수** : 하나의 입력 특성(x)에 대한 실제값(y)과 예측값(yhat)의 오차를 계산하는 함수
- 로지스틱 회귀에서 사용하는 손실함수 :

$$L(\hat{y}, y) = -(y \log \hat{y} + (1 - y) \log (1 - \hat{y}))$$

1) y가 1이라면

$$L(\hat{y}, y) = -log(\hat{y})$$

이 값이 최대한 작아지길 원하기 때문에, y 예측값이 커지길 원함.

→ y 예측값은 1보다 클 수 없기 때문에 1에 수렴하길 원함.

2) y가 0이라면

$$L(\hat{y}, y) = -log(1 - \hat{y})$$

손실 함숫값을 줄이고 싶다면 y 예측값이 작아지길 원함

→ y 예측값이 0에 수렴하도록 매개 변수를 조정할 것임.

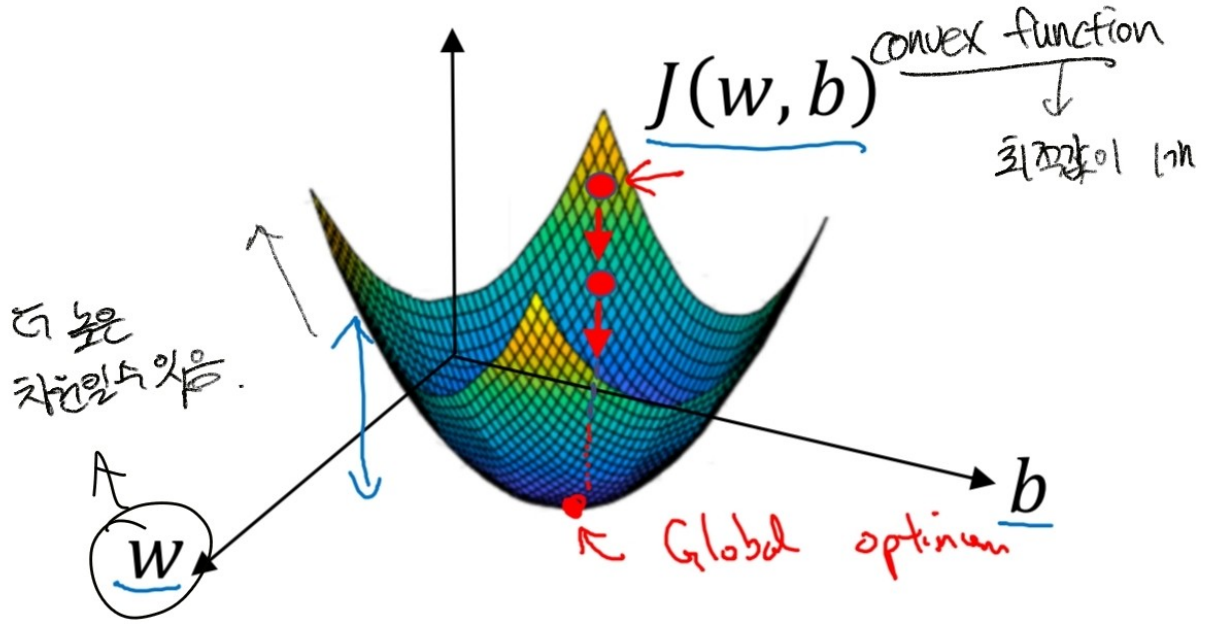
- **비용함수** : 훈련 세트 전체에 대해 얼마나 잘 추측되었는지 측정해주는 함수 모든 입력에 대한 손실 함수의 평균값

$$J(w, b) = -\frac{1}{m} \Sigma_{i=1}^{i=m} (y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log (1 - \hat{y}^{(i)}))$$

로지스틱 회귀를 학습한다는 것은 비용함수를 최소화해주는 매개 변수 w와 b를 찾는 것임.

경사하강법(Gradient Descent)

비용 함수 그래프 모양



매개 변수 값을 초기화하여 초기점에서 시작해 가장 가파른 내리막 방향으로 내려가면서 최적값에 가까워짐.

내려간다는 것은 매개 변수를 (매개 변수- 학습률 * 미분계수)를 뺀 값으로 업데이트 하는 것

$$w := w - \alpha \frac{\partial J(w,b)}{\partial w}$$

$$b := b - \alpha \frac{\partial J(w,b)}{\partial b}$$

어느 지점에서 시작해도 전역 최솟값까지 도달하게 됨.

- $dw = \frac{\partial J(w,b)}{\partial w}$

- $db = \frac{\partial J(w,b)}{\partial b}$

파이썬으로 변수를 지정할 땐 dw, db로 씀.
