

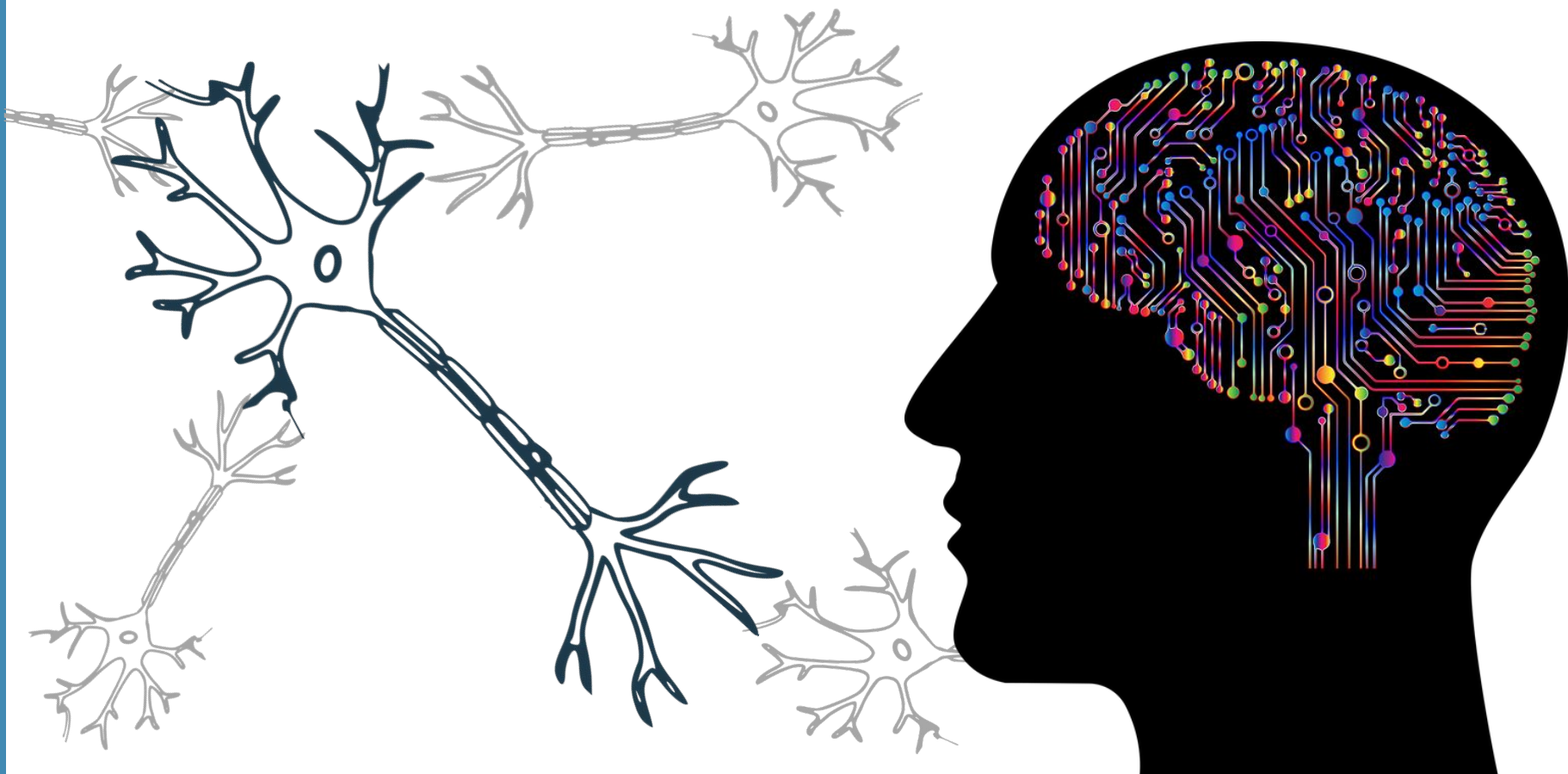
인공지능 개론

L08 Artificial Neural Network

국민대학교
소프트웨어융합대학원
박하명

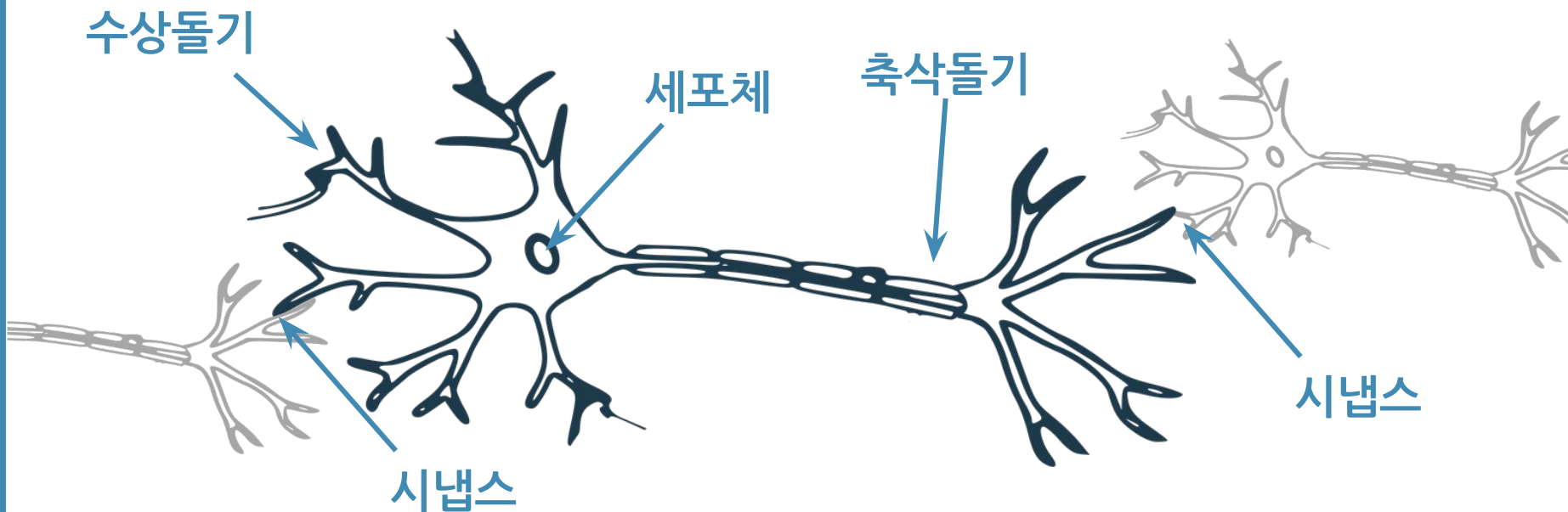
뇌는 어떻게 동작하는가?

인간의 뇌는 조밀하게 서로 연결된 신경 세포 '뉴런'으로 이루어져 있다.



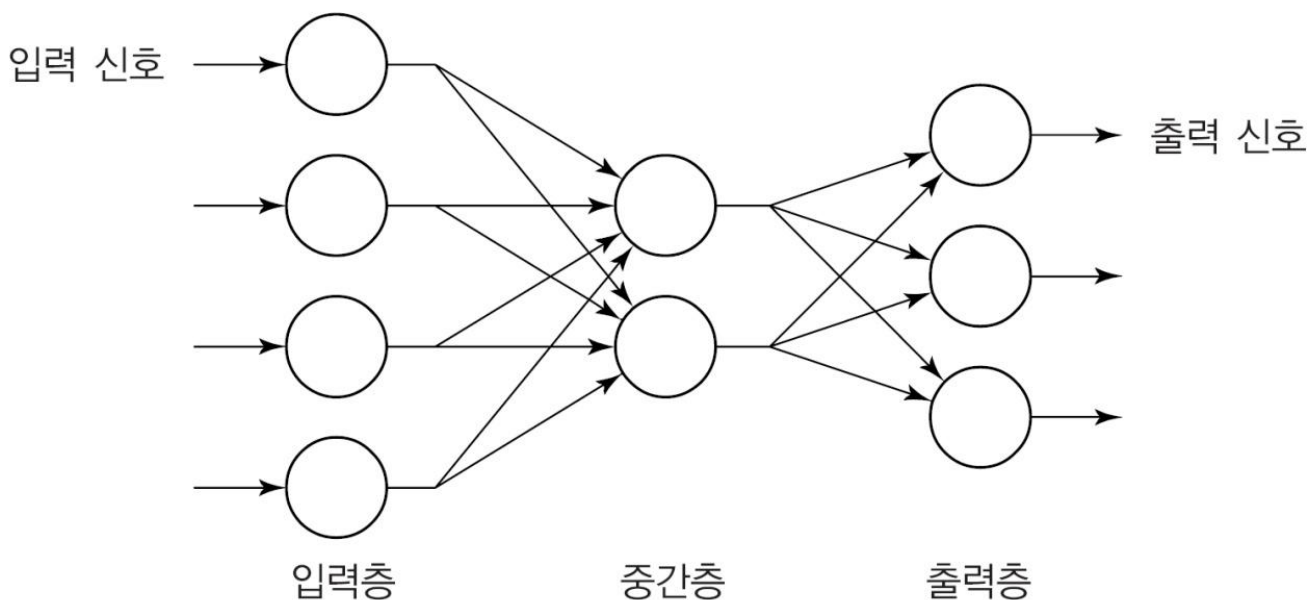
뇌는 어떻게 동작하는가?

- 뉴런 **Neuron**: 기본 정보 처리 단위
- 시냅스 **Synapse**: 뉴런간의 연결 부위, 화학 혹은 전기적인 신호로 정보 전달
- 인간의 뇌는 100억 개의 뉴런과 6조 개의 시냅스의 결합체
- 신경망의 주요 특징 - 적응성 → 학습 가능함!
 - 적응성에 따라 '잘못된 답'으로 이끄는 뉴런들 사이의 연결은 약화되고, '올바른 답'으로 이끄는 연결은 강화된다.



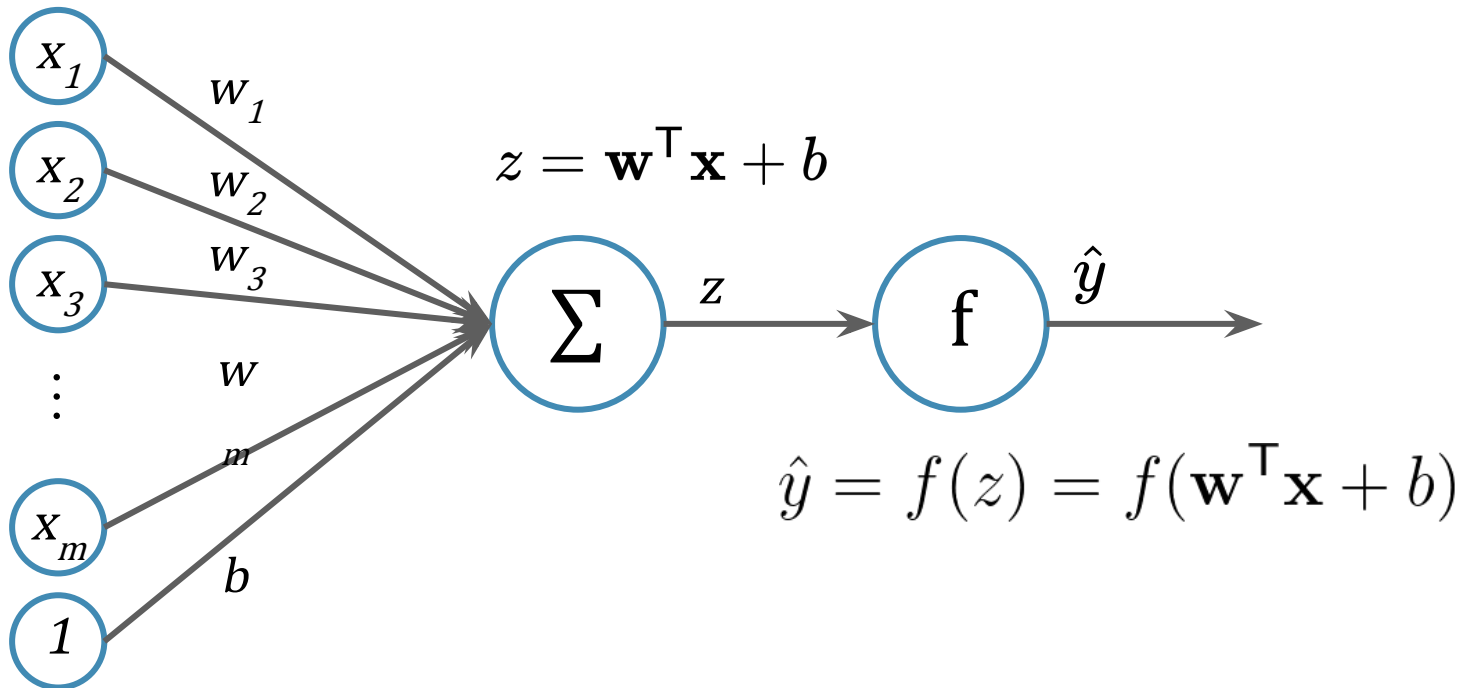
인공신경망 Artificial Neural Network

- 인간의 뇌를 흉내내어 무언가를 **학습**할 수 있지 않을까?



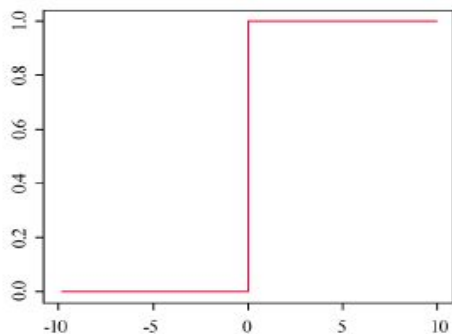
인공 뉴런

- **인공 뉴런**: 인공신경망의 기본 정보 처리 단위
- 여러개의 입력 신호에 가중치를 곱하여 합하고 활성화함수에 넣는다.
- 활성화함수의 결과는 다른 인공뉴런의 입력이 되거나 출력된다.

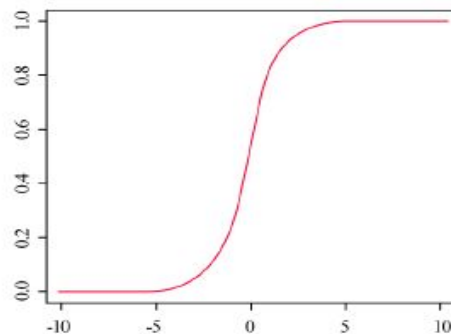


활성함수 Activation function

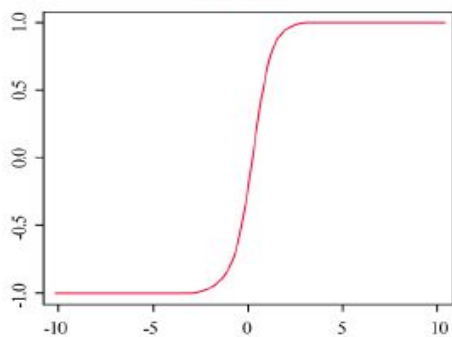
- 활성화함수로 어떤 함수든 사용 가능
- 예) 계단함수, 시그모이드 함수, 쌍곡탄젠트 함수, ReLU (Rectified Linear Unit) 함수



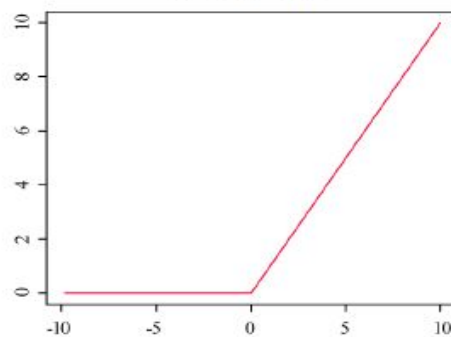
계단함수



시그모이드 함수



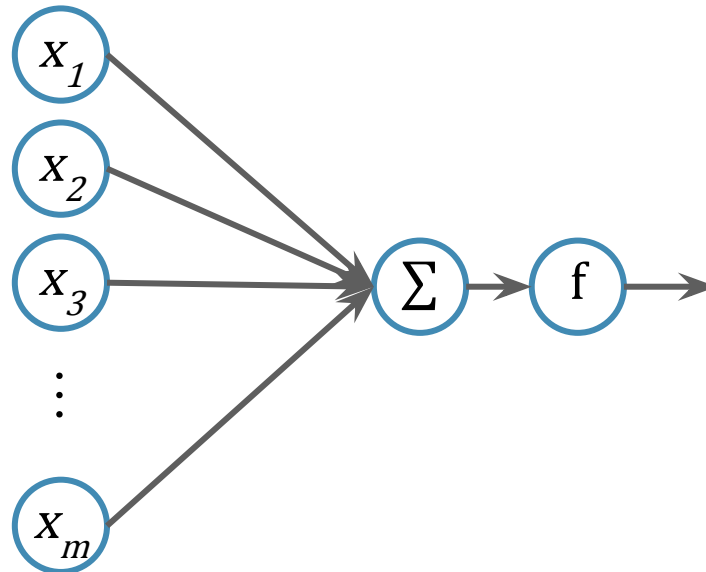
쌍곡탄젠트 함수



ReLU 함수

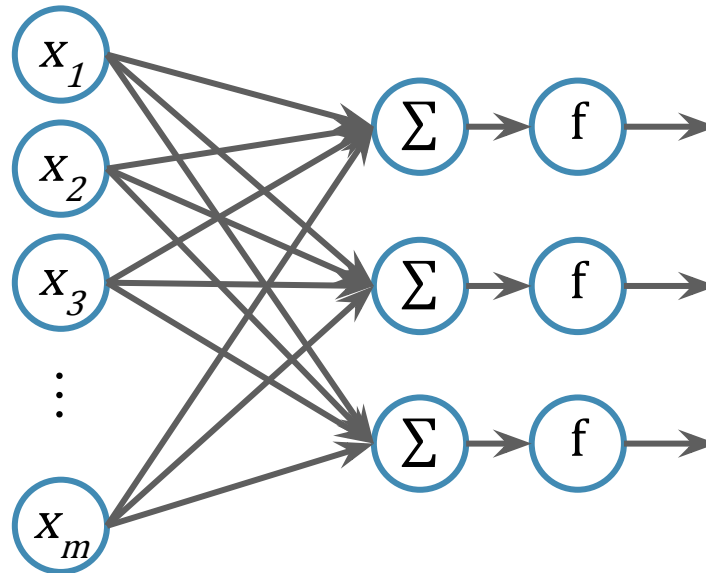
인공신경망: 여러개의 인공 뉴런

- 인공 뉴런 하나가 단독으로 사용되어 학습할 수 있다.
 - 예1) Linear Regression (활성함수가 시그모이드 함수인 인공 뉴런)
 - 예2) 로젠블랫의 퍼셉트론 (활성함수가 계단함수)



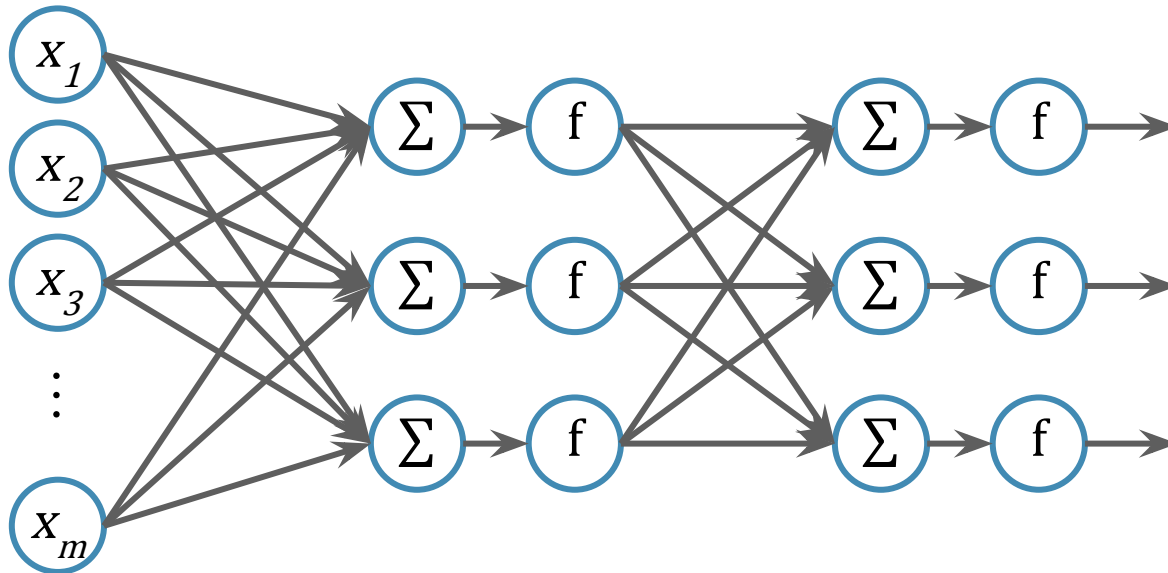
인공신경망: 여러개의 인공 뉴런

- 인공 뉴런이 **병렬적으로 연결**되어 여러 개의 출력을 가질수도 있다.
 - 예) Softmax Regression



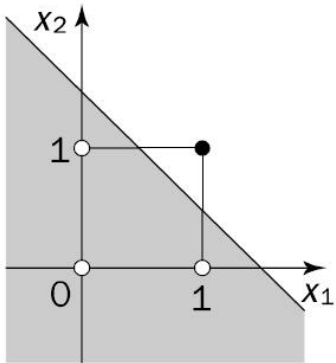
인공신경망: 여러개의 인공 뉴런

- 인공 뉴런의 출력은 다른 인공 뉴런의 입력으로 활용될 수 있다.
 - 보다 복잡한 문제를 해결할 수 있다.
 - 예) 1개의 층으로 된 인공신경망은 xor 문제를 학습할 수 없다..!

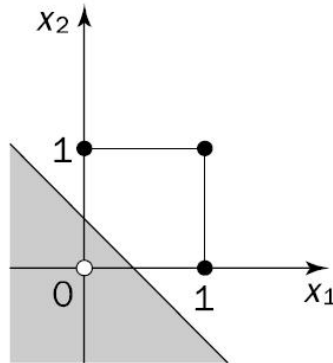


다층 인공신경망의 필요성: XOR 문제

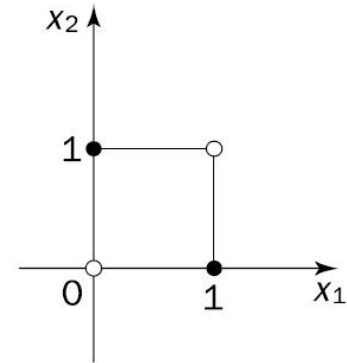
- AND, OR 연산은 단층의 인공신경망으로 학습 가능.
- 하지만 XOR은 학습 불가능.. Why?
 - 단층 신경망은 선형 분리가 가능한 함수만 학습 가능하다.



(a) AND 연산 ($x_1 \cap x_2$)



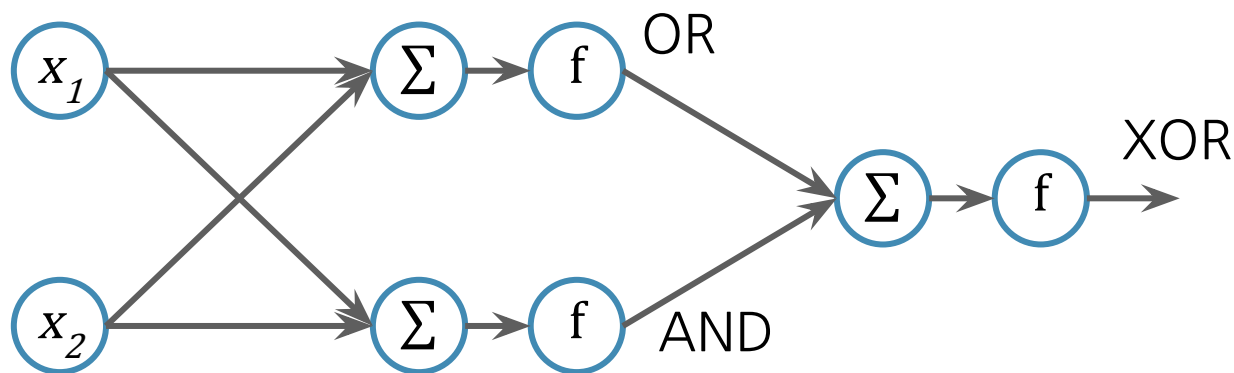
(b) OR 연산 ($x_1 \cup x_2$)



(c) Exclusive-OR 연산 ($x_1 \oplus x_2$)

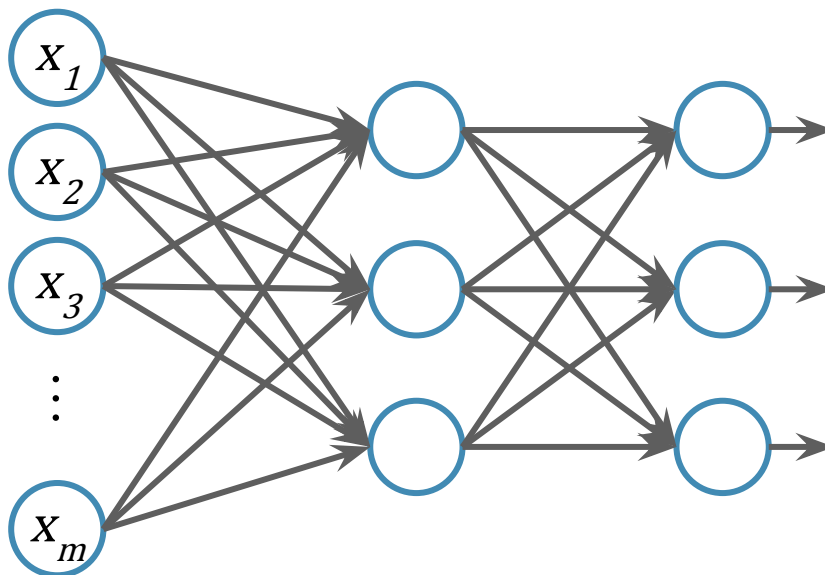
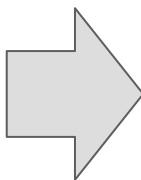
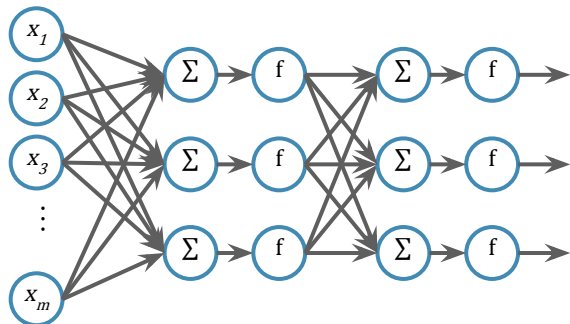
다층 인공신경망의 필요성: XOR 문제

- 선형 분리가 불가능한 입력이 여러 층의 인공신경망을 거쳐서 선형 분리가 가능한 입력으로 변신할 수 있다..!



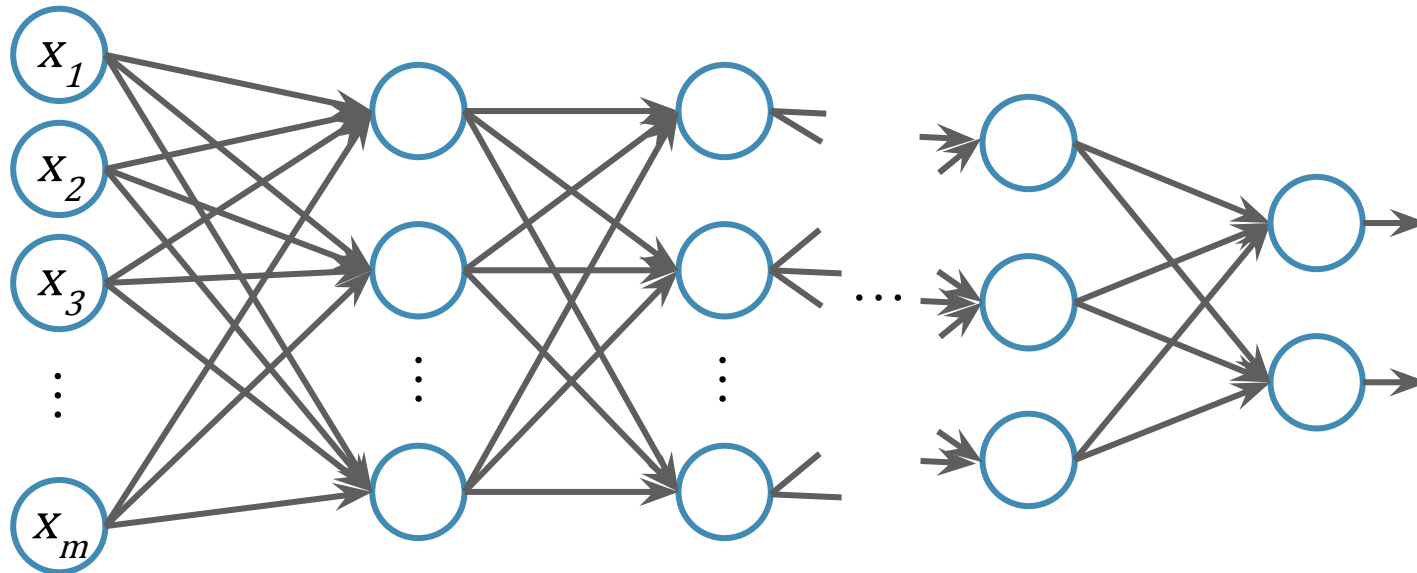
인공신경망: 여러개의 인공 뉴런

- 다층 인공 신경망의 표현을 단순화



Deep Neural Networks

- 층이 많아질 수록 더욱 더 복잡한 문제를 해결 가능하다...!
 - 하지만... 계산 복잡도가 크게 증가하여 학습이 오~래걸린다..!



인공신경망 학습

- 비용 함수를 최소화 하도록 학습.
- 가중치(+bias)에 대한 비용함수의 기울기를 구하고, 기울기에 따라 가중치 값을 조절한다.
 - 이 과정을 여러번 반복. 수렴할 때까지.

역전파 알고리즘

- 네트워크가 복잡해졌는데, 어떻게 기울기를 구할 수 있을까..?
- 우리는 걱정이 없습니다. pytorch에서 **backward()** 메소드만 있으면 기울기가 자동으로 구해지니깐요 o(^__^)
- 그렇다면, 원리는? 연쇄법칙! (Chain Rule)

$$(f(g(x)))' = f'(g(x))g'(x)$$

$$\frac{dt}{dx} = \frac{dt}{du} \frac{du}{dx}$$

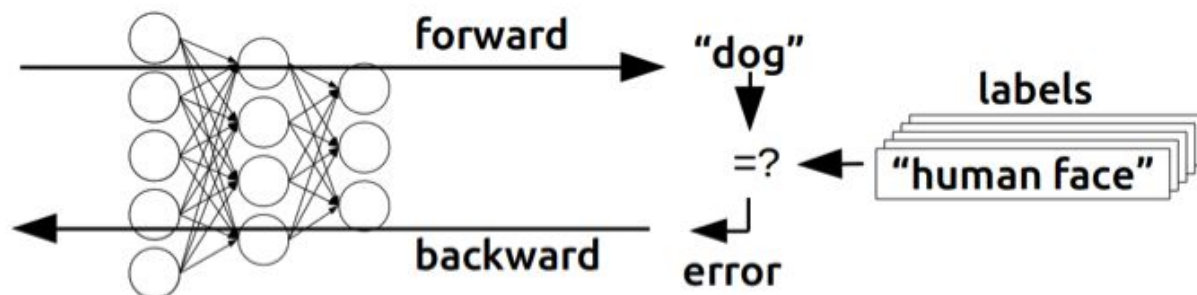
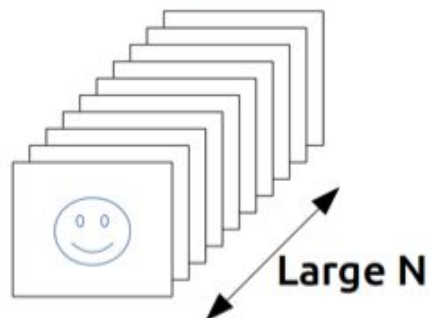
미분, 편미분 다시보기

- $f(x) = 3$
- $f(x) = 2x$
- $f(x) = x^2$
- $f(x) = 1/x$
- $f(x) = e^x$
- $f(x,y) = 2xy$

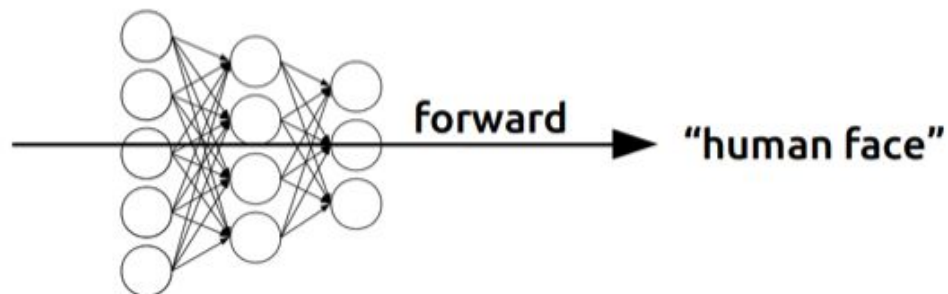
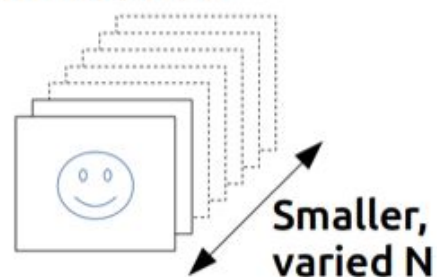
$$f'(x) = \frac{d}{dx} f(x) = \lim_{\Delta x \rightarrow 0} \frac{f(a + \Delta x) - f(a)}{\Delta x}$$

역전파 알고리즘

Training

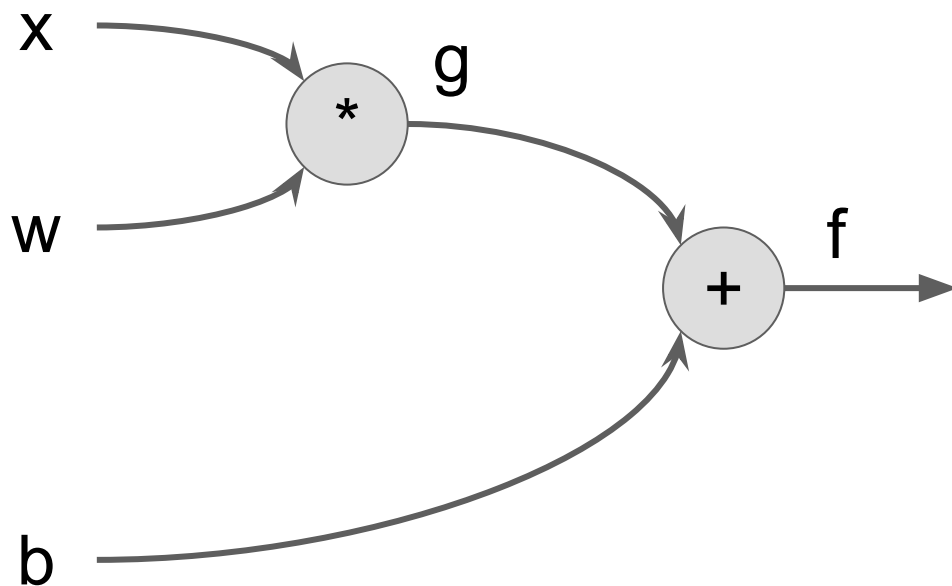


Inference



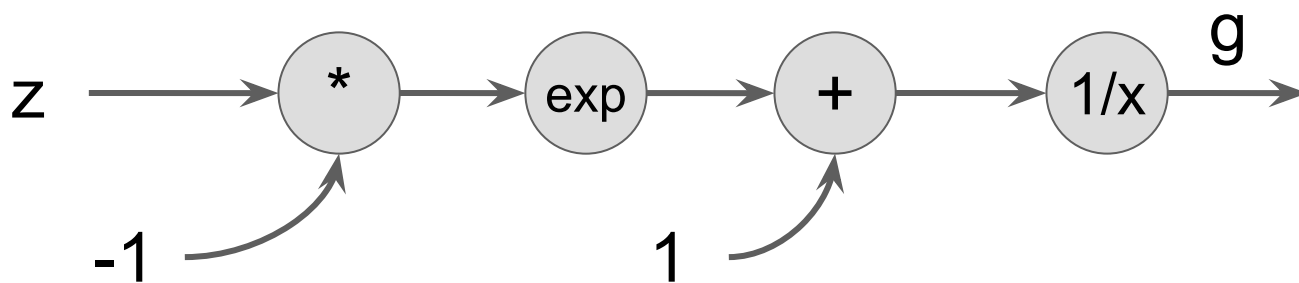
역전파 알고리즘

- $f = wx + b, g = wx \rightarrow f = g + b$



역전파 알고리즘

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



Question?