

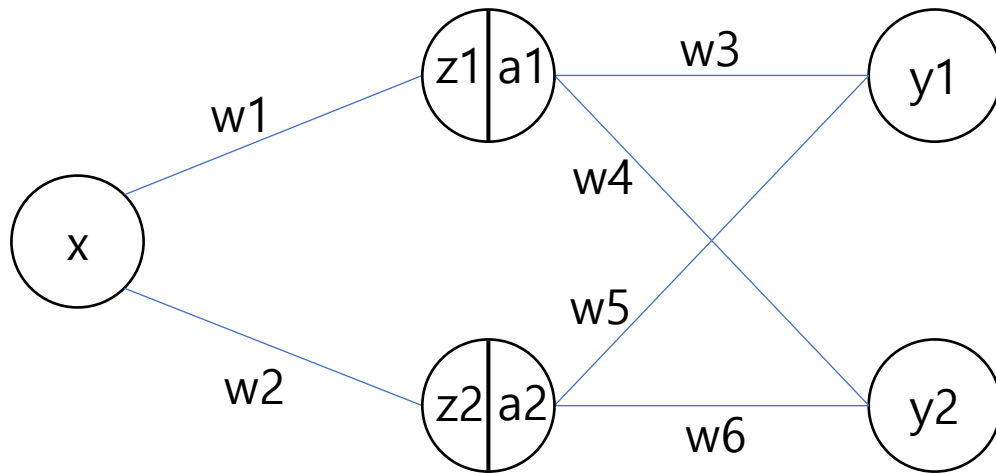
딥알못 탈출하기 #7

고려대학교 지능시스템 연구실

유민형 제작

Nonlinearity

- 1. Gradient Descent Algorithm을 할 때부터 $f(x)=x(\text{if } x>0)$ or $0(\text{else})$ 함수가 있었다. 이 함수는 입력 값을 받아서 기준치 (threshold)를 넘으면 값을 통과시키는 역할을 한다.



$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{else} \end{cases}$$

$$z1 = w1*x + b1$$

$$z2 = w2*x + b2$$

$$a1 = f(z1)$$

$$a2 = f(z2)$$

$$y1 = w3*a1 + w5*a2$$

$$y2 = w4*a1 + w6*a2$$

Nonlinearity

- 이 함수 $f(x)=x(\text{if } x>0) \text{ or } 0(\text{else})$ 는 활성화함수(Activation function)이라고 하며 생물학(신경학/뇌공학)에서 뉴런의 전기적 신호가 역치(threshold)를 넘었을 때만 신경 물질을 다음 뉴런에 전달하는 원리에서 영감을 받아 만들어졌다.
- 처음에는 Sigmoid라는 함수를 사용하다가 ReLU라는 함수를 사용하여 신경망을 이용한 머신러닝이 크게 발전하였다. (잠깐 상식 퀴즈: 누가 Sigmoid 사용했고, 누가 ReLU 썼을까?)

Nonlinearity

- (문제1) 활성화 함수는 신경망에 어떤 성질을 부여하기 위해 사용된다. 어떤 성질일까? 또한 그 성질을 부여해야 하는 이유는 무엇일까?
- (문제2) Sigmoid 함수를 사용했을 때 어떤 문제가 발생하는가? ReLU함수는 이 문제를 어떻게 해결했는가?
- (문제3) MLP에서 활성화 함수를 사용하지 않았을 때 그 신경망은 hidden layer가 없는 신경망으로 대체할 수 있는가? 그 신경망은 Linear regression이라고 볼 수 있나?

Loss function

- Loss를 정의할 때 항상 MSE와 같은 함수를 사용하지 않는다. 다양한 loss 함수에 대해서 알아보자.
- (문제4) (1) Softmax 활성화함수와 cross-entropy loss함수의 관계에 대해 설명하시오. (2) Softmax를 통과해 나온 output은 어떤 방식으로 예측하는가? (3) Softmax를 ensemble 관점으로 설명하시오.
- (문제5) Gradient Descent를 이용해 Gradient Ascent를 해야한다면 이를 어떻게 활용해야 하는가?
- (문제6) 여러 종류의 Loss 함수에 대해 조사하고(BCEloss 외 4개 이상), Loss 함수를 정의하는 것이 주어진 task에 어떻게 영향을 미치는지 설명하시오.