

신경망 예시

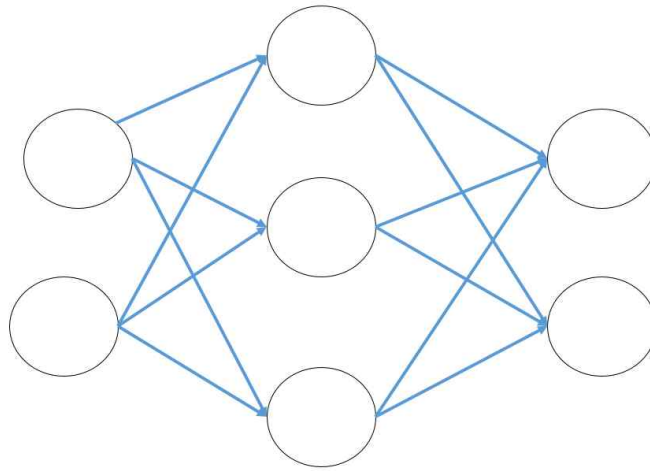


그림 1 신경망 예시

맨 왼쪽 층을 **입력층**, 가운데 층을 **은닉층**, 맨 오른쪽 층을 **출력층**

가중치를 갖는 층(화살표가 있는) 2개 이므로 '2층 신경망'

활성화 함수

활성화 함수(Activation Function) : 입력 신호의 총합을 출력 신호로 변환하는 함수
: 입력 신호의 총합이 활성화를 일으키는지 정하는 역할

$$a = b + w_1x_1 + w_2x_2$$
$$y = h(a)$$

가중치가 달린 입력 신호와 편향의 총합을 a 라 칭함

이러한 a 를 h 함수(활성화 함수)에 넣어 출력 하는 것이 y

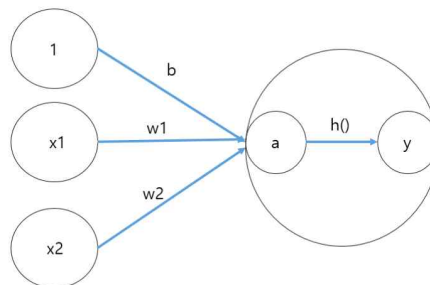


그림 2 활성화 함수의 처리 과정

가중치 신호를 조합한 결과가 **a** 라는 노드, 활성화 함수 **h()**를 통과하여 y 라는 노드로 변환

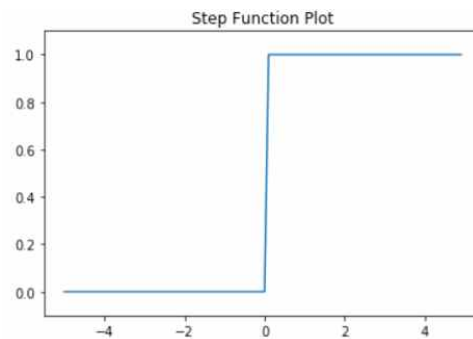
단순 퍼셉트론 : 단층 네트워크에서 계단함수(임계값을 경계로 출력이 바뀌는 함수)를 활성화 함수로 사용한 모델

다층 퍼셉트론 : 여러 층으로 구성되고 여러 활성화 함수를 사용하는 신경망 모델

활성화 함수의 종류

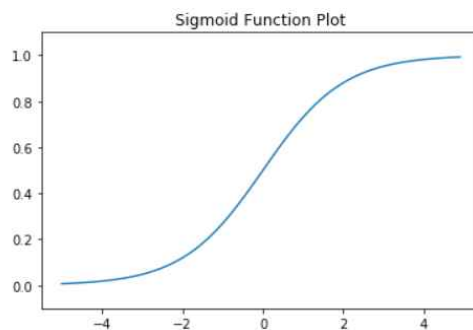
1. 계단 함수(Step Function)

$$h(x) = \begin{cases} 0 & (x \leq 0) \\ 1 & (x > 0) \end{cases} \quad \text{임계값을 경계로 출력이 바뀌는 함수}$$



2. 시그모이드 함수(Sigmoid Function)

$$h(x) = \frac{1}{1 + \exp(-x)}$$



- 시그모이드 함수와 계단 함수의 비교

차이점 :

시각적인 직관적 비교 : 매끄러움의 차이

퍼셉트론은 0 아니면 1이라는 이진적으로 값을 반환하지만 신경망은 그 사이에 있는 실수(0.731, 0.88 등)을 반환해준다.

공통점 : 입력이 작을 때는 0에 가깝고 입력이 클 때는 1에 가까워진다.

범위는 0~1 사이

비선형 함수

step function과 sigmoid function의 공통점 = 비선형 함수

신경망은 비선형 함수를 활성화 함수로 사용해야 된다.

왜? 층을 아무리 깊게 해도 의미가 없음

선형 $h(x)=cx$ 를 활성화 함수로 사용할 때

$y(x)=h(h(h(x)))$ 는 $y=c*c*c*x = y=ax(a=c**3)$ 따라서 여러층으로 구성하여도 의미가 없음

3. ReLU(Rectified Linear Unit) 함수

$$h(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases}$$

