

# 심층학습

## 02 인공 신경망(1)

SW융합학부 양희경

주교재: 오렐리앙 제롱, 핸드온 머신러닝(사이킷런과 텐서플로를 활용한 머신러닝, 딥러닝 실무), 한빛미디어, 2018.04

# 학기 내용

1. 심층학습 소개 Deep learning
- 2. 인공 신경망 Neural network (1/2)**
3. 역전파 Backpropagation
4. 정규화 Regularization
5. 합성곱 신경망 Convolutional neural network(CNN)
6. 오토인코더 Auto encoder(AE)
7. 적대적 생성 네트워크 Generative adversarial network(GAN)
8. 순환 신경망 Recurrent neural network(RNN)

# 내용

## 2.1 회귀와 분류 (간략히) → 기계학습 2, 4장

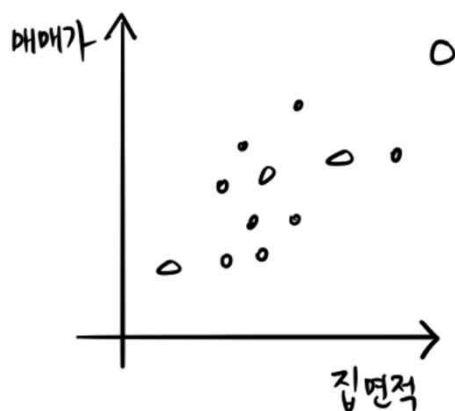
- 회귀와 분류의 정의
- 비용함수

## 2.2 인공 신경망: 퍼셉트론

## 2.3 다중 분류와 인공 신경망

## 2.1 회귀와 분류 (간략히)

### 회귀 Regression



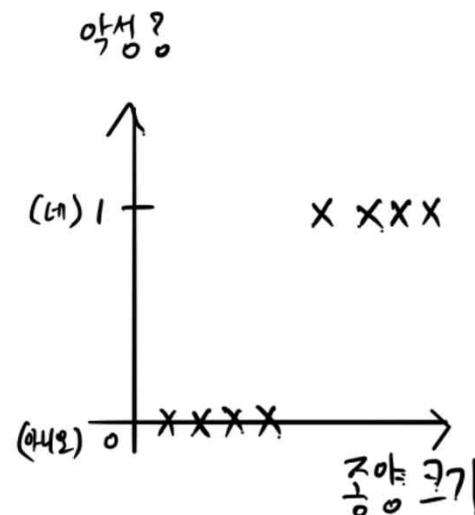
$$y = ax + b$$

$$y = wx + b$$

$$y = \theta_1 x + \theta_0$$

데이터를 이용하여 특성의 가중치(weight(w)) 합과 편향(bias(b))이라는 상수를 더해 **연속적인 값**의 결과를 예측하는 과정

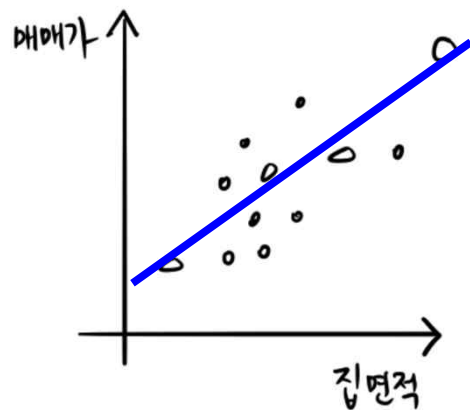
### (이진)분류 Binary classification



**특정 카테고리**로 구분하는 작업.  
카테고리가 두 개인 경우 이진 분류,  
세 개 이상인 경우 다중 분류라 함.

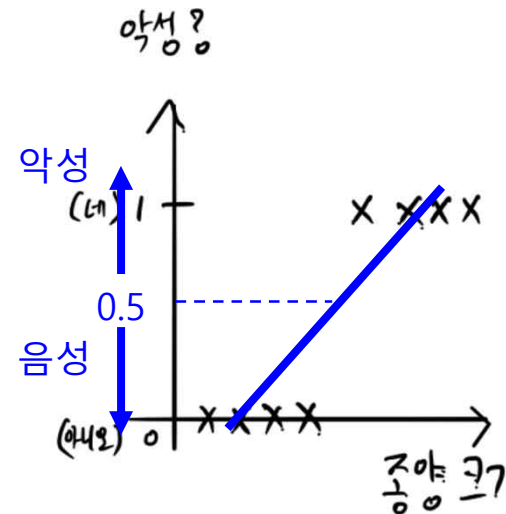
## 2.1 회귀와 분류 (간략히)

### 회귀 Regression



$$\hat{y} = h_w(x) = wx + b$$

### (이진)분류 Binary classification

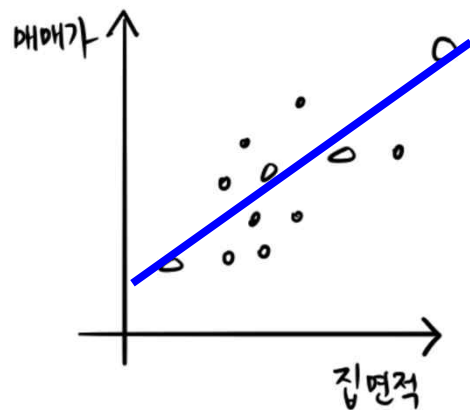


이런 데이터가 있다면?

$$\begin{aligned} \hat{y} \geq 0.5 \text{ 이면 } \hat{y} &= 1 \\ \hat{y} < 0.5 \text{ 이면 } \hat{y} &= 0 \end{aligned}$$

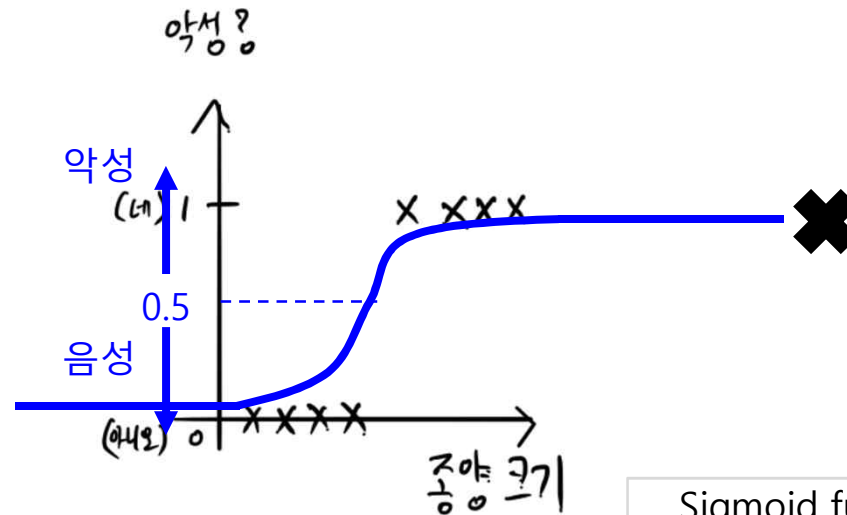
## 2.1 회귀와 분류 (간략히)

회귀 Regression



$$h_w(x) = wx + b$$

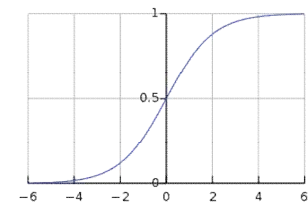
(이진)분류 Binary classification



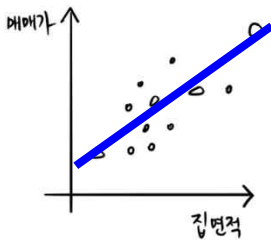
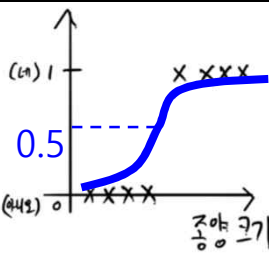
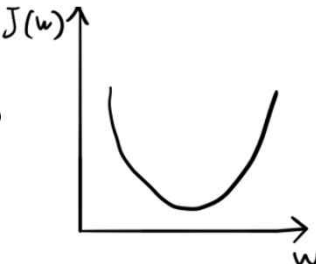
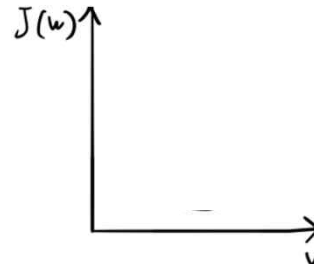
$$y = g(wx + b) \\ = \frac{1}{1 + e^{-(wx + b)}}$$

Sigmoid function  
Logistic function

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

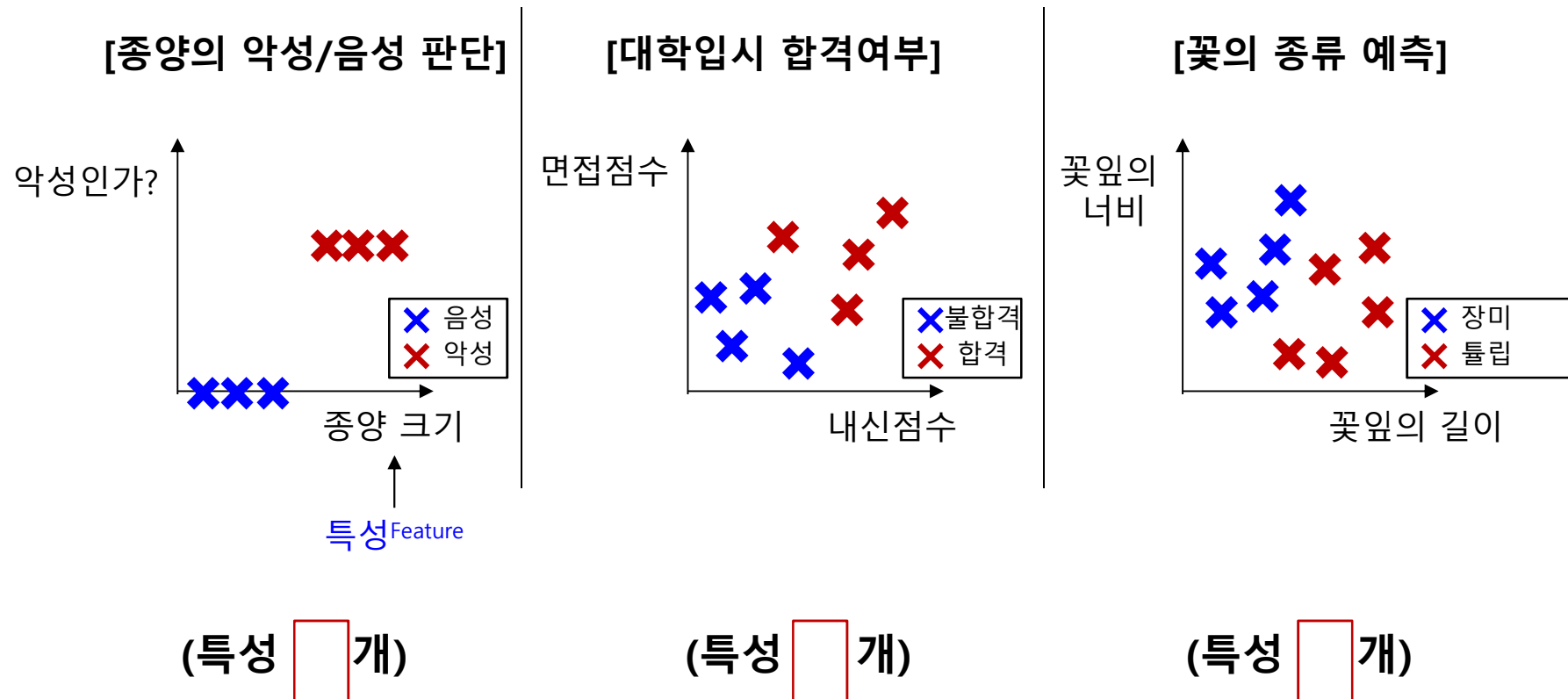


## 2.1 회귀와 분류 (간략히)

	회귀Regression	(이진)분류Binary classification
예제		
모델Model	$h_w(x) = wx + b$	$h_w(x) = g(wx + b)$
비용함수 Cost function	$J(w) = \frac{1}{2m} \sum_{i=1}^m ((wx^{(i)} + b) - y^{(i)})^2$ <p> <math>m</math>: 데이터 개수  <math>\hat{y}</math>: 모델을 통한 예측값  <math>y</math>: 실제 데이터의 값(정답)         </p> 	$J(w) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)})]$ 
최적화 Optimization	Gradient descent, Stochastic gradient descent, Momentum, NAG, Adagrad, AdaDelta, Adam, ...	

## 2.1 회귀와 분류 (간략히)

- 분류 문제의 예

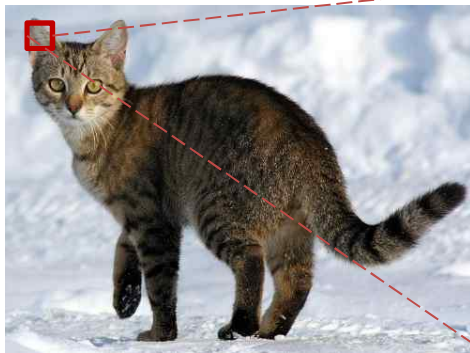




## 2.1 회귀와 분류 (간략히)

- 분류 문제의 예
  - 데이터가 이미지라면?

[고양이 사진인지 판단하기]



©Wikipedia

[ 189 203 216 ]	[ 187 204 214 ]	[ 187 204 214 ]
[ 180 193 210 ]	[ 180 193 209 ]	[ 181 194 211 ]
[ 173 195 208 ]	[ 174 196 207 ]	[ 176 195 209 ]
[ 185 193 212 ]	[ 186 194 215 ]	[ 188 196 215 ]
[ 186 200 211 ]	[ 185 201 214 ]	[ 186 202 215 ]
[ 189 203 216 ]	[ 191 204 220 ]	[ 195 208 224 ]
[ 193 207 218 ]	[ 193 207 218 ]	[ 194 208 219 ]
[ 198 207 222 ]	[ 195 207 223 ]	[ 197 210 226 ]
[ 182 206 210 ]	[ 187 206 213 ]	[ 192 206 215 ]
[ 201 208 218 ]	[ 204 209 215 ]	[ 209 214 220 ]
[ 204 222 226 ]	[ 207 222 227 ]	[ 207 222 225 ]

...

⋮

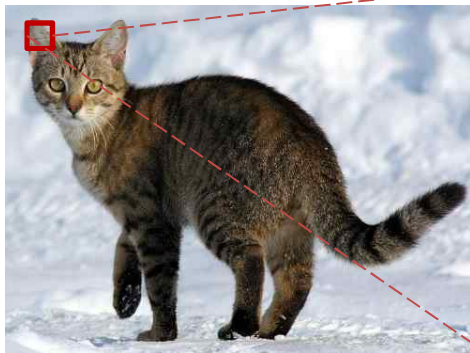
(특성  개)

(width, height, # of channel) = (300,200,3)

## 2.1 회귀와 분류 (간략히)

- 분류 문제의 예
  - 데이터가 이미지라면?

[고양이 사진인지 판단하기]



©Wikipedia

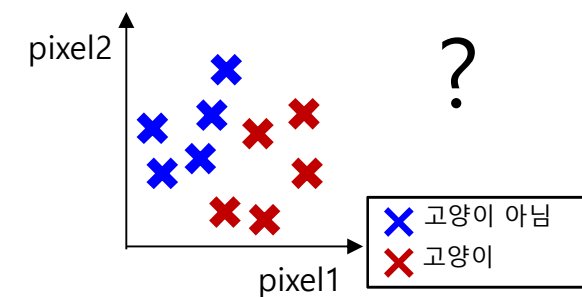
[ 189 203 216 ]	[ 187 204 214 ]	[ 187 204 214 ]
[ 180 193 210 ]	[ 180 193 209 ]	[ 181 194 211 ]
[ 173 195 208 ]	[ 174 196 207 ]	[ 176 195 209 ]
[ 185 193 212 ]	[ 186 194 215 ]	[ 188 196 215 ]
[ 186 200 211 ]	[ 185 201 214 ]	[ 186 202 215 ]
[ 189 203 216 ]	[ 191 204 220 ]	[ 195 208 224 ]
[ 193 207 218 ]	[ 193 207 218 ]	[ 194 208 219 ]
[ 198 207 222 ]	[ 195 207 223 ]	[ 197 210 226 ]
[ 182 206 210 ]	[ 187 206 213 ]	[ 192 206 215 ]
[ 201 208 218 ]	[ 204 209 215 ]	[ 209 214 220 ]
[ 204 222 226 ]	[ 207 222 227 ]	[ 207 222 225 ]

...

⋮

(특성  개)

(width, height, # of channel) = (300,200,3)



# 내용

## 2.1 회귀와 분류 (간략히)

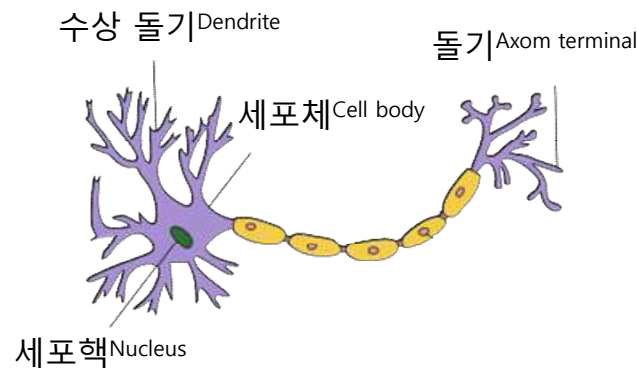
## 2.2 인공 신경망

- 생물학적 신경망과 인공 신경망
- 퍼셉트론의 정의와 표현

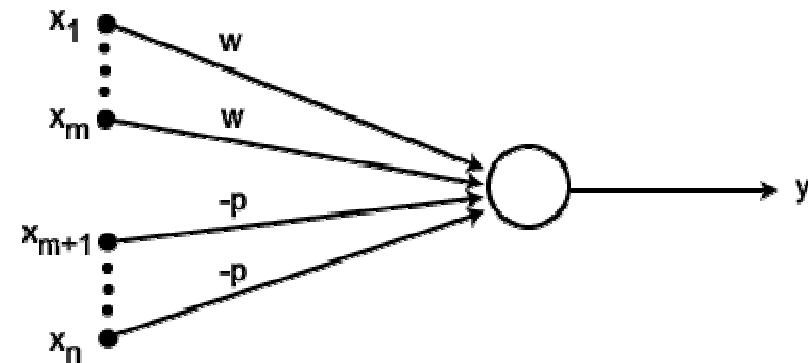
## 2.3 다중 분류와 인공 신경망

## 2.2 인공 신경망

- 생물학적 신경(뉴런)Neuron과 인공 뉴런Artificial neuron



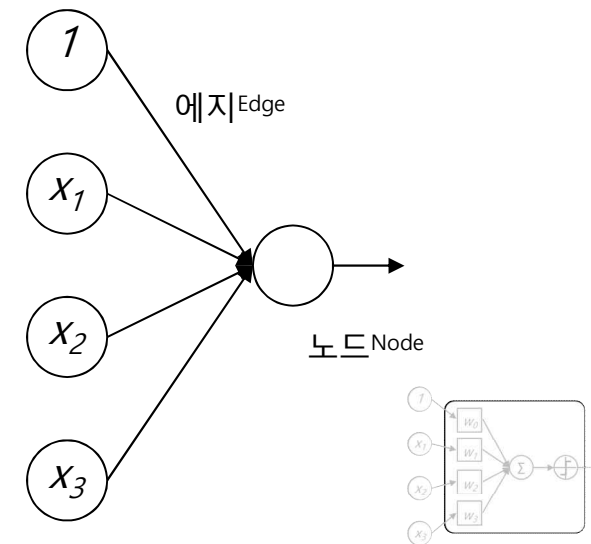
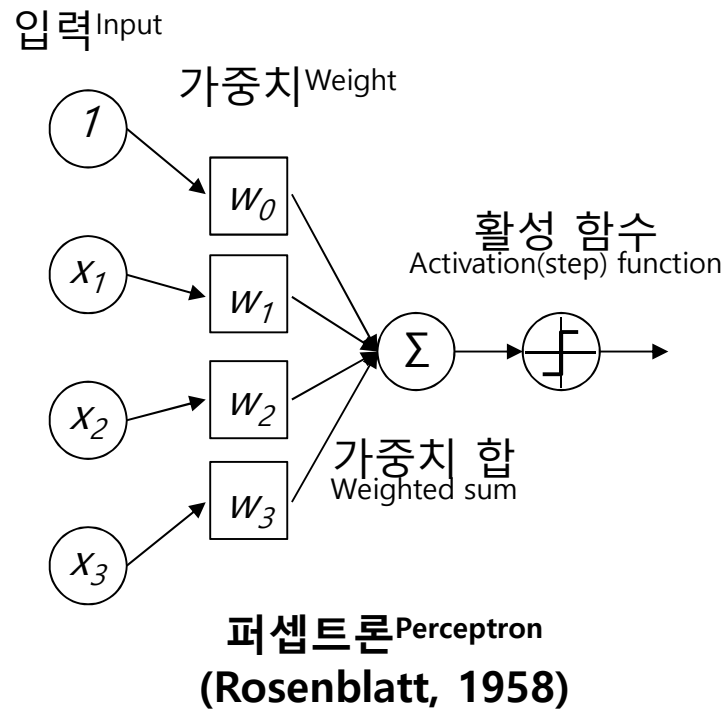
생물학적 신경(뉴런)Neuron



최초의 인공 뉴런Artificial neuron  
(McCulloch & Pitt, 1943)

## 2.2 인공 신경망

- 퍼셉트론 Perceptron (뉴런)
  - 현대 인공 신경망의 Artificial neural network 기본 단위

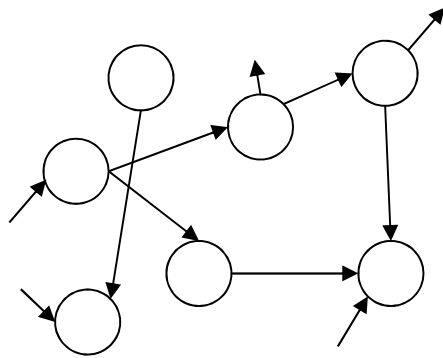


### 퍼셉트론의 그래프 표현

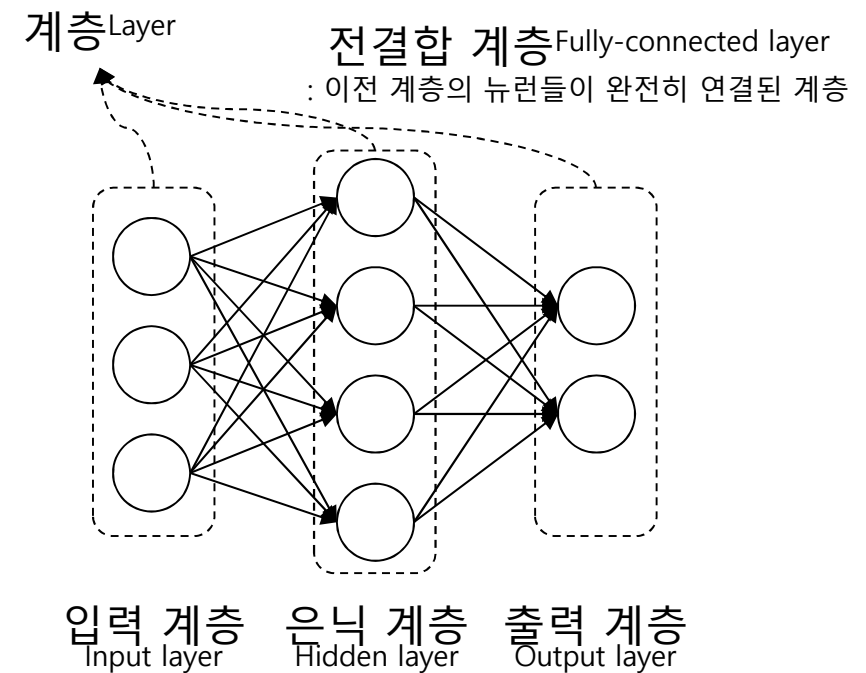
- 가중치, 가중치 합, 활성 함수를 한 개의 노드로 표현.
- 에지는 뉴런의 연결성의 의미

## 2.2 인공 신경망

- 현대적인 인공 신경망: 퍼셉트론들이 계층구조를 이룸

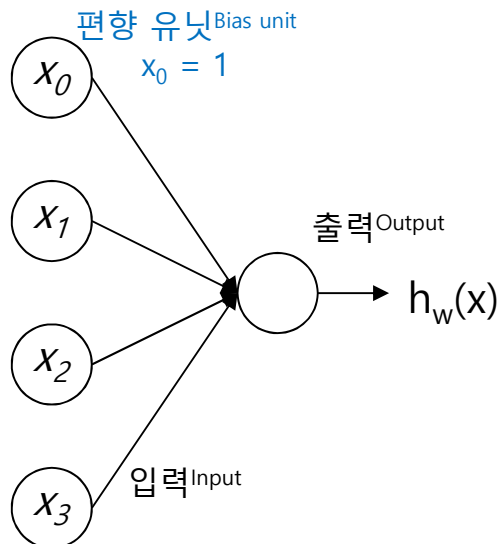


인공 신경망



## 2.2 인공 신경망

### • 퍼셉트론의 표현



$$X = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad W = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

가중치(파라미터)  
Weights(parameters)

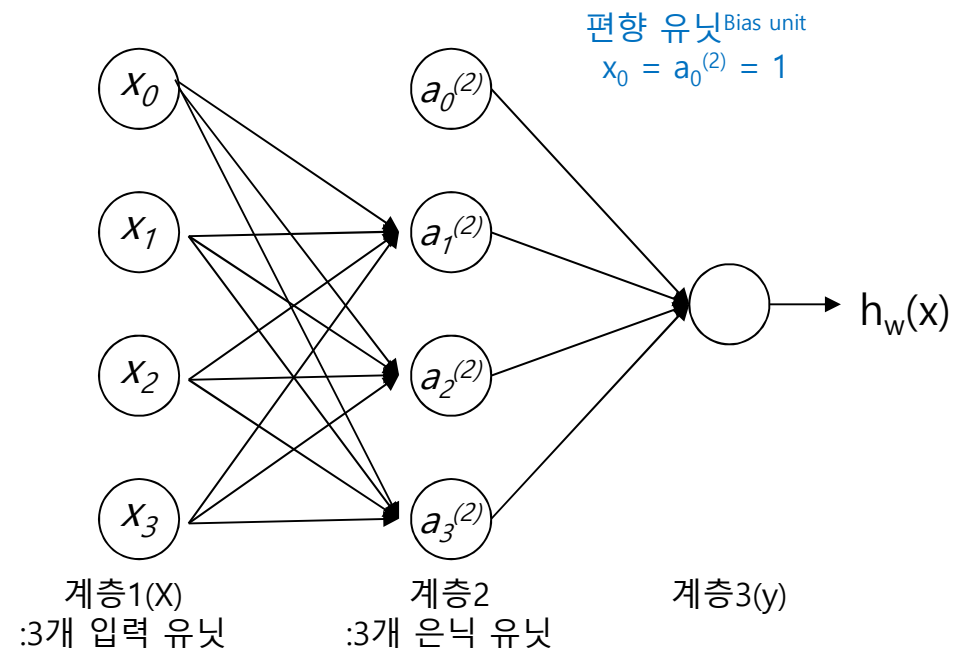
- $h_w(x) = g(W^T X)$   

$$= \frac{1}{1 + e^{-W^T X}}$$
- 시그모이드(로지스틱) 활성화 함수 Sigmoid(logistic) activation function

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

## 2.2 인공 신경망

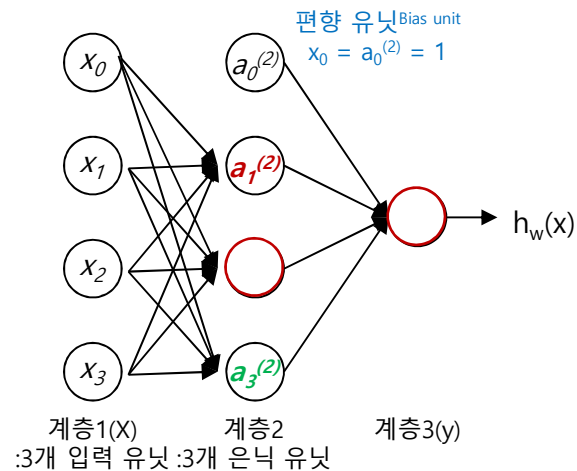
- 퍼셉트론의 표현의 예





## 2.2 인공 신경망

### • 퍼셉트론의 표현의 예



- $a_i^{(j)}$ : j번째 레이어의 i번째 유닛의 활성화값
- $W^{(j)}$ : j번째에서 j+1번째 레이어로 매핑하는 함수의 파라미터의 행렬

$$a_1^{(2)} = g(W_{01}^{(1)}x_0 + W_{11}^{(1)}x_1 + W_{21}^{(1)}x_2 + W_{31}^{(1)}x_3)$$

$$a_2^{(2)} = g(W_{02}^{(1)}x_0 + W_{12}^{(1)}x_1 + W_{22}^{(1)}x_2 + W_{32}^{(1)}x_3)$$

$$a_3^{(2)} = g(W_{03}^{(1)}x_0 + W_{13}^{(1)}x_1 + W_{23}^{(1)}x_2 + W_{33}^{(1)}x_3)$$

$$h_W(x) = a_1^{(3)} = g(W_{01}^{(2)}a_0^{(2)} + W_{11}^{(2)}a_1^{(2)} + W_{21}^{(2)}a_2^{(2)} + W_{31}^{(2)}a_3^{(2)})$$

#### ※ 행렬의 형태 추측

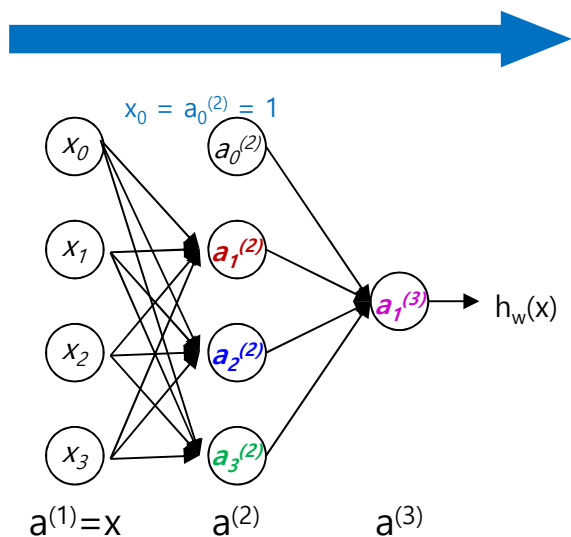
네트워크의 j번째 레이어에  $S_j$ 개 유닛이 있고,

j+1번째 레이어에  $S_{j+1}$ 개 유닛이 있다면,

행렬  $W^{(j)}$ 는  행렬이 됨

## 2.2 인공 신경망

- 정방향 전파 Forward propagation: 벡터화된 구현법 Vectorized implementation



$$a^{(2)} = g(\underline{z^{(2)}})$$

$$\underline{z^{(2)}} = W^{(1)} \cdot a^{(1)}$$

$$a^{(3)} = g(\underline{z^{(3)}})$$

$$\underline{z^{(3)}} = W^{(2)} \cdot a^{(2)}$$

$$\therefore h_w(x) = a^{(3)} = g(\underline{z^{(3)}})$$

$$a_1^{(2)} = g(W_{01}^{(1)}x_0 + W_{11}^{(1)}x_1 + W_{21}^{(1)}x_2 + W_{31}^{(1)}x_3)$$

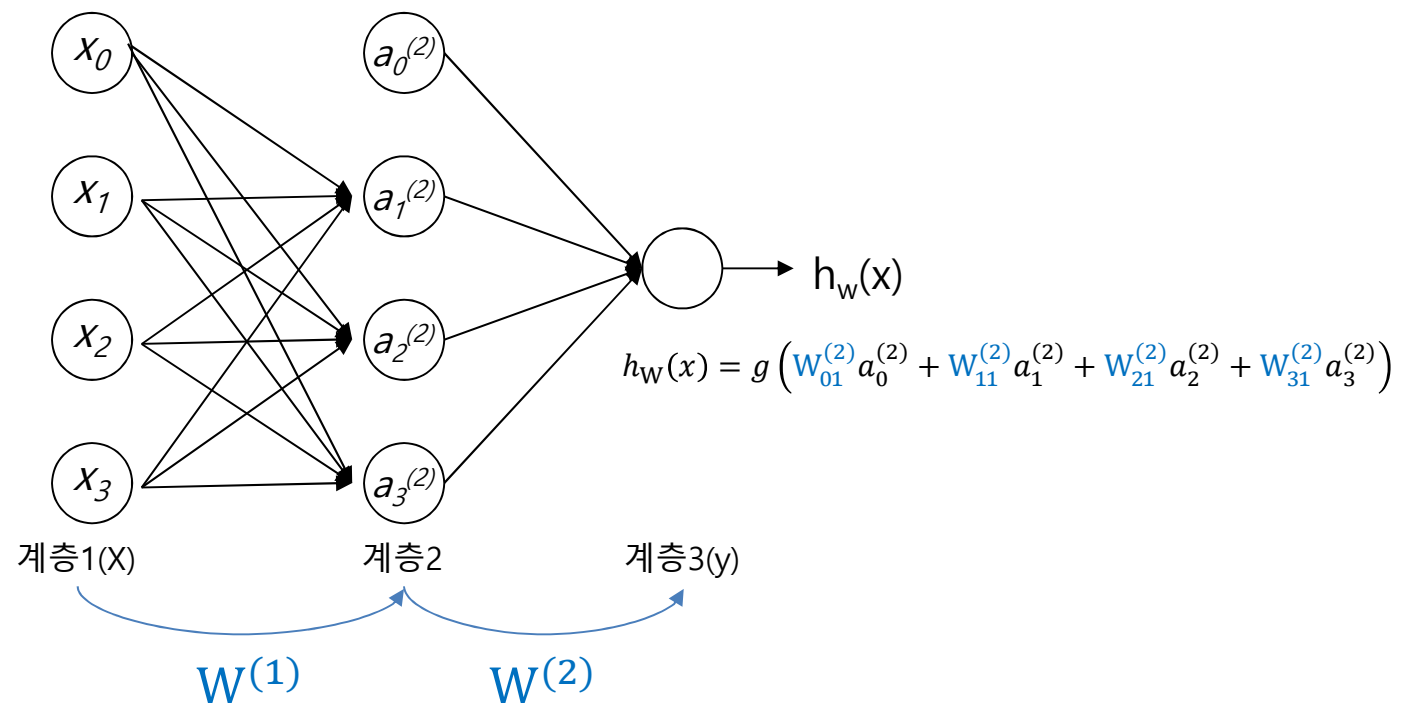
$$a_2^{(2)} = g(W_{02}^{(1)}x_0 + W_{12}^{(1)}x_1 + W_{22}^{(1)}x_2 + W_{32}^{(1)}x_3)$$

$$a_3^{(2)} = g(W_{03}^{(1)}x_0 + W_{13}^{(1)}x_1 + W_{23}^{(1)}x_2 + W_{33}^{(1)}x_3)$$

$$h_w(x) = a_1^{(3)} = g(W_{01}^{(2)}a_0^{(2)} + W_{11}^{(2)}a_1^{(2)} + W_{21}^{(2)}a_2^{(2)} + W_{31}^{(2)}a_3^{(2)})$$

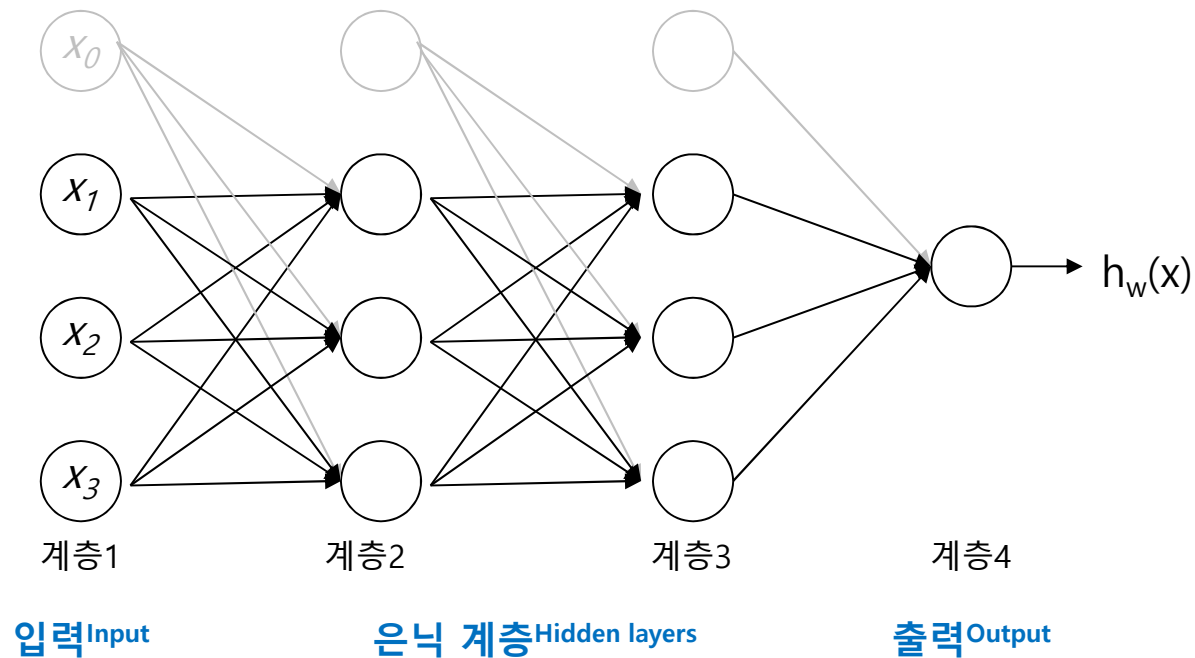
## 2.2 인공 신경망

- 인공 신경망의 학습
  - '인공 신경망을 학습한다': 가중치의 값을 결정함
  - 가중치<sup>Weight</sup>(파라미터<sup>Parameter</sup>)



## 2.2 인공 신경망

- 인공 신경망의 예: 다층 퍼셉트론 Multi-layered perceptron
  - 일반적으로 편향을 생략하여 표현함



# 내용

2.1 회귀와 분류 (간략히)

2.2 인공 신경망: 퍼셉트론

**2.3 다중 분류와 인공 신경망**

## 2.3 다중 분류와 인공 신경망

- 다중 분류 Multiclass classification
  - 다중 출력 유닛 Multi output units: 일대다 문제 One-vs-all.



고양이



개

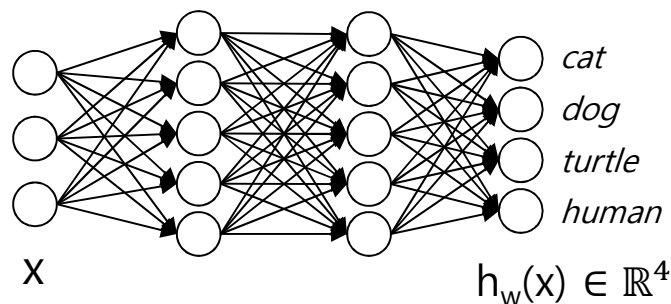


거북이



사람

©Wikipedia



<원하는 출력 유닛의 형태>

(고양이 사진일 때)  $h_w(x) \approx \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$

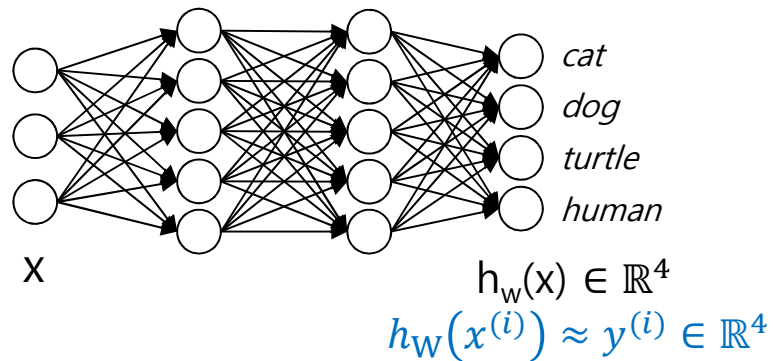
(개일 때)  $h_w(x) \approx \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},$

(거북이일 때)  $h_w(x) \approx \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \dots$



## 2.3 다중 분류와 인공 신경망

### • 다중 분류 Multiclass classification



<원하는 출력 유닛의 형태>

(고양이 사진일 때)  $h_W(x) \approx \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$

(개일 때)  $h_W(x) \approx \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},$

(거북이일 때)  $h_W(x) \approx \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \dots$

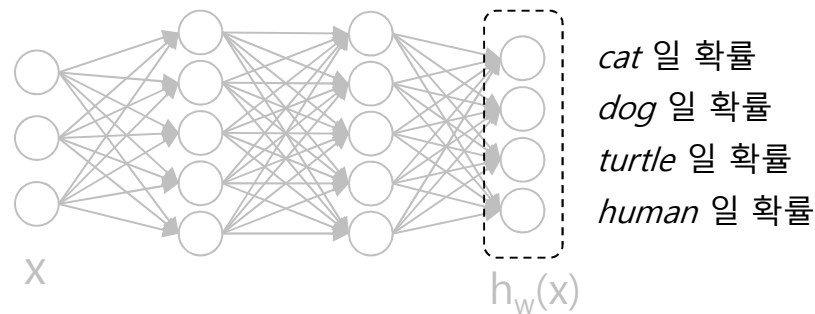
- 학습 데이터셋  
 $(x^1, y^1), (x^2, y^2), \dots, (x^m, y^m)$

- $y^{(i)}$  는 다음 중 하나

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

## 2.3 다중 분류와 인공 신경망

- 소프트맥스 함수 Softmax function
  - 출력층에 추가되는 함수
  - 각 카테고리일 확률을 의미함
  - 0~1 사이로 출력 유닛의 값을 매핑함
  - 모두 더하면 1





## 2.3 다중 분류와 인공 신경망

- 소프트맥스 함수 Softmax function
  - 각 입력의 지수함수를 정규화함
  - $\text{softmax}(x)_i = \frac{e^{x_i}}{\sum_j e^{x_j}}$
  - N가지 카테고리로 분류하는 다중 분류 문제에 사용됨



# 내용

2.1 회귀와 분류 (간략히) → 기계학습 2, 4장

2.2 인공 신경망: 퍼셉트론

2.3 다중 분류와 인공 신경망

# 학기 내용

1. 심층학습 소개 Deep learning
- 2. 인공 신경망 Neural network (2/2)**
3. 역전파 Backpropagation
4. 정규화 Regularization
5. 합성곱 신경망 Convolutional neural network(CNN)
6. 오토인코더 Auto encoder(AE)
7. 적대적 생성 네트워크 Generative adversarial network(GAN)
8. 순환 신경망 Recurrent neural network(RNN)