

# [딥러닝 1단계] 강의 필기

## 개념

인공지능 vs 머신러닝 vs 딥러닝

머신러닝

지도 학습

비지도 학습

강화 학습

딥러닝

신경망 / 로지스틱 회귀

선형 회귀 (Logistic Regression)

## 개념

### 인공지능 vs 머신러닝 vs 딥러닝

- 인공지능 ⊃ 머신러닝 ⊃ 딥러닝



### [요약]

- 정의
  - **인공지능** : 인간의 지적 능력을 컴퓨터에서 구현하는 기술
  - **머신러닝** : 컴퓨터를 인간처럼 학습시킴으로써 인간의 도움 없이 컴퓨터가 스스로 새로운 규칙을 생성할 수 있지 않을까 하는 발상으로부터 시작 / 데이터에서 패턴을 스스로 찾아내고 이를 기반으로 다른 데이터가 주어졌을 때 예측하는 기술
  - **딥러닝** : 신경망을 이용한 머신러닝 방법 / 신경망을 학습시키는 것

- [참고자료](#)

## 머신러닝



### [요약]

- 종류
  - **지도 학습** (Supervised Learning)
  - **비지도 학습** (Unsupervised Learning)
  - **강화 학습** (Reinforcement Learning)

### 지도 학습

- label(정답)이 붙은 데이터로 학습하는 것

### 비지도 학습

- label이 없는 데이터로 학습하는 것
- 종류 : 군집화

### 강화 학습

- 차이점 :
  - 지도 학습/비지도 학습 → 데이터가 주어진 상태 + 변화가 없는 정적인 환경에서 학습
  - 강화 학습 → 어떤 환경 안에서 정의된 주체(agent)가 현재의 상태(state)를 관찰하여 선택할 수 있는 행동(action)들 중에서 가장 최대의 보상(reward)을 가져다주는 행동이 무엇인지를 학습

## 딥러닝



### [요약]

- 정의
  - **딥러닝** : 신경망을 학습시키는 것
  - **신경망** : 뉴런이 복잡하게 연결된 신경망을 컴퓨터에서 인공적으로 구현한 것 / 입력(x)와 출력(y)를 매칭해주는 함수를 찾는 과정

- 신경망
  - = 인공신경망  
= **ANN** (Artificial Neural Network)
  - 종류 : (아래 모두 지도학습으로 학습시킬 수 있음.)

Standard NN	~ <u>Neural Network</u>	
CNN	합성곱 신경망 Convolutional <u>Neural Network</u>	이미지
RNN	순환 신경 Recurrent <u>Neural Network</u>	오디오 (1차원 시퀀스 데이터)

# 신경망 / 로지스틱 회귀

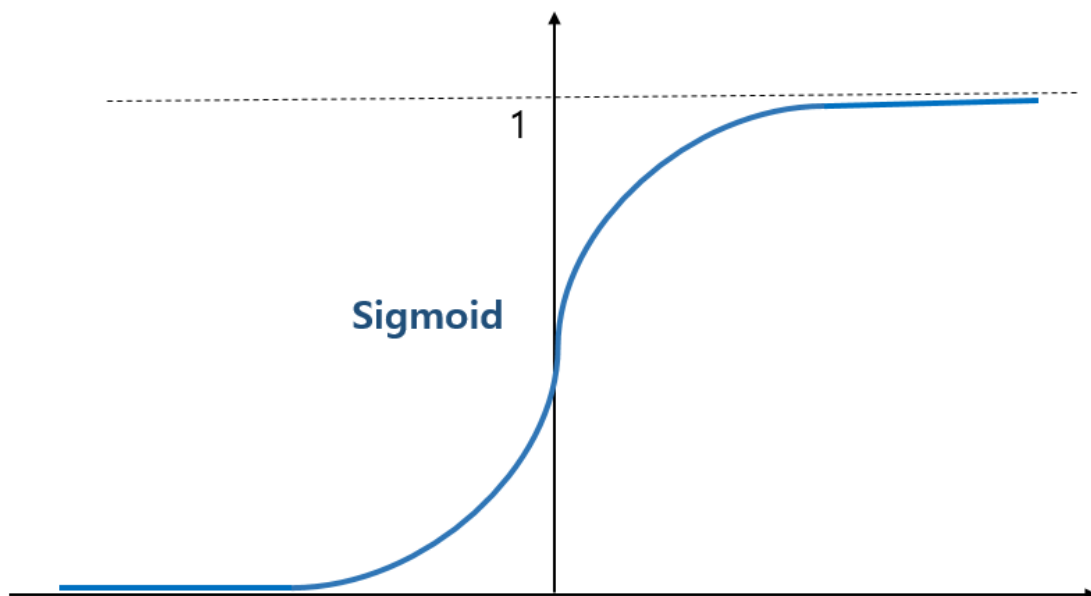
지도학습

## 선형 회귀 (Logistic Regression)

$$\hat{y} = W^T X + b$$

- **Binary Classification** : 0 or 1로 분류
  - 이거 하려고 선형 회귀 씬.
- **sigmoid 함수**
  - Binary Classification에서  $\hat{y}$ 는 **y=1일 확률** 을 의미
  - 따라서 결과값을 0~1로 바꿔주기 위해 sigmoid를 씀

$$\hat{y} = \sigma(W^T X + b)$$



(1) Sigmoid 함수

- 오차( $y - \hat{y}$ ) 계산하기

- **Loss 함수** : 한 입력 data에서의 오차

- ▼ 구체적인 수식

- 보통 손실함수는  $L(\hat{y}, y) = \frac{1}{2}(\hat{y} - y)^2$  식을 사용하지만 로지스틱 회귀에서 이러한 손실 함수를 사용하면 지역 최소값에 빠질 수 있기 때문에 사용하지 않습니다. (해당 내용은 향후 다시 나올것이니 걱정하지 않으셔도 됩니다.)

- 로지스틱 회귀에서 사용하는 **손실 함수**는 다음과 같습니다.

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

- 이 함수를 직관적으로 이해하기 위해 두 가지 경우로 나누어 생각해 볼 수 있습니다.

- 1)  $y = 0$  인 경우  $L(\hat{y}, y) = -\log(1 - \hat{y})$  가 0에 가까워지도록  $\hat{y}$  는 0에 수렴하게 됩니다.

- 2)  $y = 1$  인 경우  $L(\hat{y}, y) = -\log \hat{y}$  가 0에 가까워지도록  $\hat{y}$  는 1에 수렴하게 됩니다.

$$L(\hat{y}, y)$$

- **Cost 함수** : 전체 입력 data에서의 오차

- Loss 함수의 평균

$$J(w, b)$$

- 학습 목표 : **cost 함수**의 값이 최소가 되도록 하는  $w$ 와  $b$ 를 찾는 것