

인공지능 신경망 Basic

AI - 스쿨

THINK LIFE SYNC AI

2022. 06. 02.



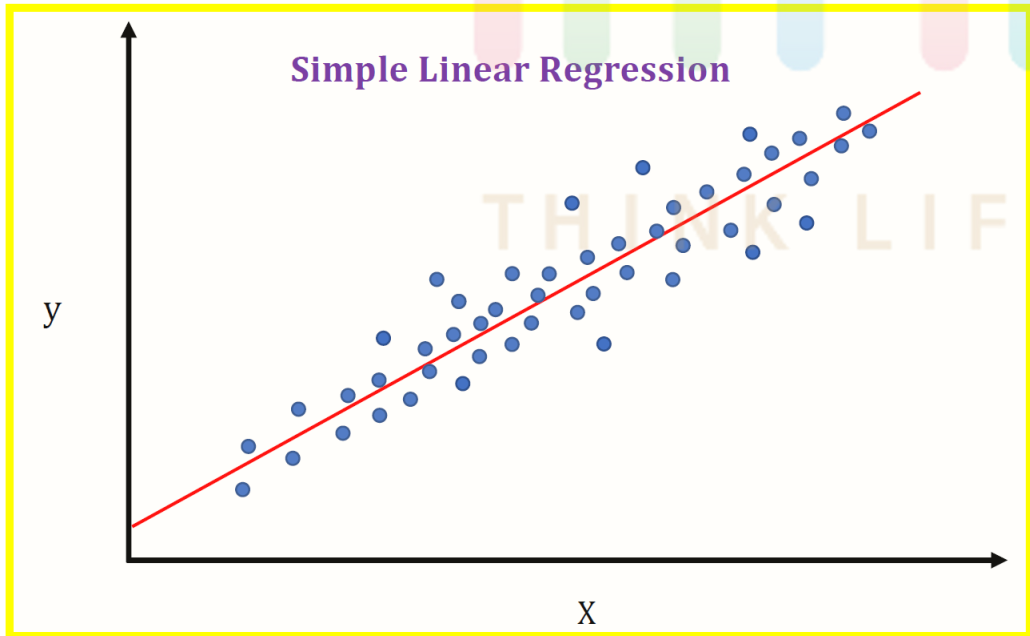
인공신경망 Basic

1. 선형분류 개념 및 실습
2. 이진분류 개념 및 실습

인공 지능 신경망 개념

- 선형 회귀 (Linear Regression)

실제 데이터를 바탕으로 모델을 생성해서 만약 다른 입력 값을 넣었을 때 발생할 아웃풋을 예측하는 데 있다. 이때 우리가 찾아낼 수 있는 가장 직관적인 간단한 모델은 선(line)이다. 그래서 데이터를 놓고 그걸 가장 잘 설명할 수 있는 선을 찾는 분석하는 방법을 선형 회귀(Linear Regression) 분석이라 부른다.

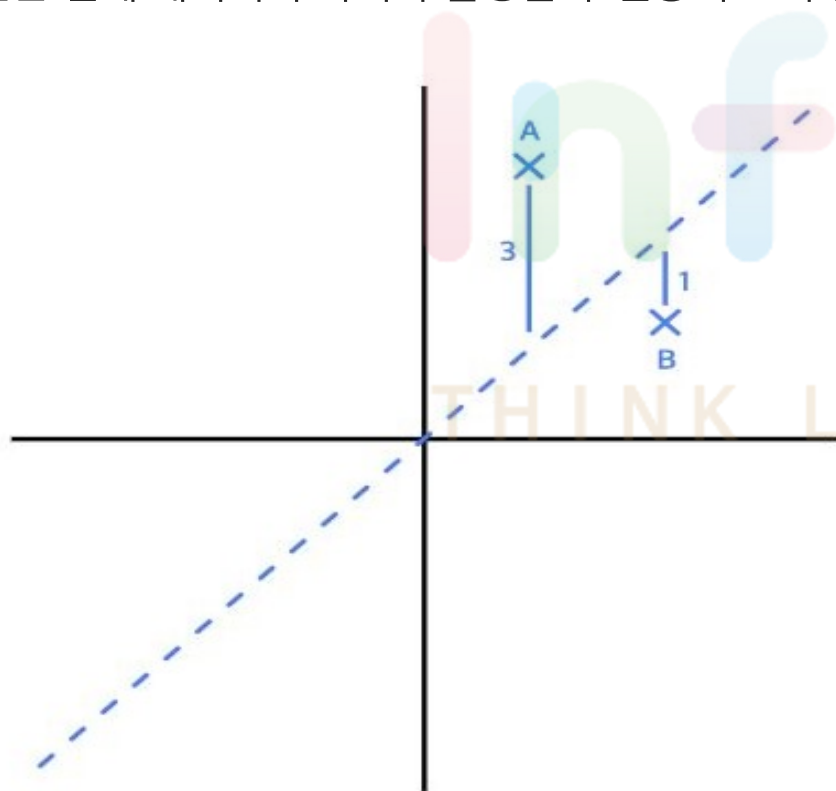


공식 다음과 같습니다. $Y = mx + b$

- 기울기: m , 절편: b 에 따라 그 선의 모양이 정해지기 때문에 x 를 넣었을 때 y 를 구할 수 있다. 선형 회귀 분석의 목적도 결국 우리가 가진 데이터를 가장 잘 설명할 수 있는 m 과 b 를 얻는다.

인공 지능 신경망 개념

- 선형 회귀 (Linear Regression) 발생하는 오차, 손실
선은 실제 데이터와 차이가 발생한다. 일종의 오차 손실 이라고 합니다.



그림을 보면 A는 3 B는 1만큼 손실이 발생했습니다.

선과 실제 데이터 사이에 얼마나 오차가 있는지 구하려면 양수, 음수 관계 없이 동일하게 반영되도록 모든 손실에 제곱 해야 합니다.

선형 회귀 손실을 구하는 걸 평균 제곱 오차(mean squared error, 이하 MSE)라고 부릅니다. 손실을 구할 때 가장 널리 쓰이는 방법입니다.

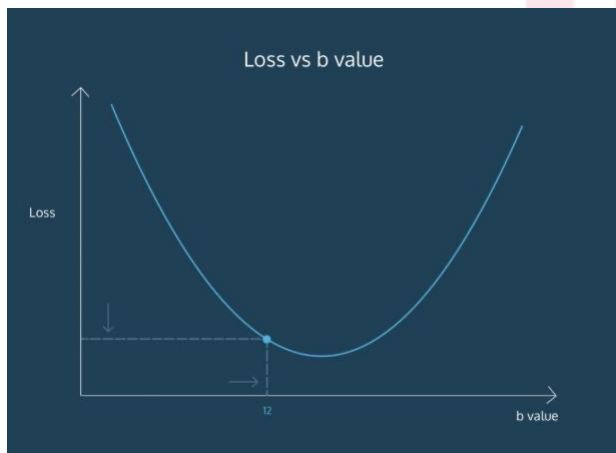
손실 구하는 이 외의 방법으로 MSE처럼 제곱하지 않고 그냥 절대값으로만 바로 평균 절대(mean absolute error, MAE), MSE와 MAE를 절충한 후버 손실 (Huber loss), $1 - \text{MES} / \text{VAR}$ 으로 구하는 결정 계수(coefficient of determination)등이 있습니다.

인공 지능 신경망 개념

- 선형 회귀 모델의 목표는 모든 데이터로부터 나타나는 오차 평균을 최소화할 수 있는 최적의 기울기 와 절편을 찾는것.

손실을 최소화 하기 위한 방법, 경사하강법(Gradient Descent)

손실(Loss)을 함수로 나타내면 이렇게 아래로 볼록한 모양이라는 거다.



파라미터를 임의로 정한 다음에 조금씩 변화시켜가며 손실을 점점 줄여가는 방법으로 최적의 파라미터를 찾아간다.

이러한 방법을 경사 하강법이라고 부릅니다.

절편을 구할 때 사용하는 공식은 아래와 같다.

$$\frac{2}{N} \sum_{i=1}^N -(y_i - (mx_i + b)) \text{ 이 값이 최소가 되도록 계속 } b \text{ 값을 변화시키는 거다.}$$

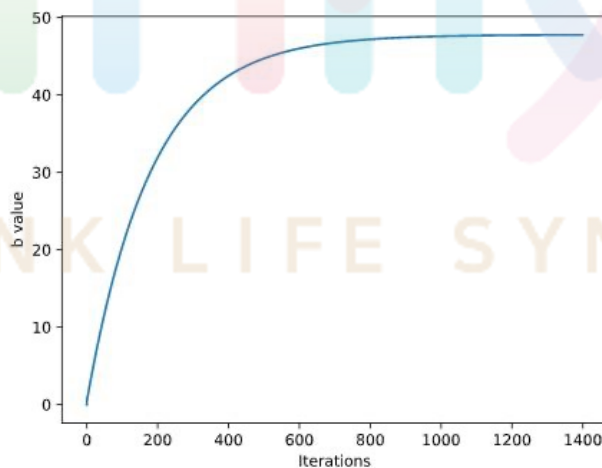
기울기를 구하는 공식은 아래와 같다.

$$\frac{2}{N} \sum_{i=1}^N -x_i(y_i - (mx_i + b)) \text{ 마찬가지로 이 값이 최소가 되도록 } m \text{ 값을 변화시키는 거다.}$$

인공 지능 신경망 개념

수렴(Convergence)

선형 회귀 분석을 수행하면 기울기와 절편을 계속 변경해가면서 최적의 값을 찾게 될 텐데, 이것 언제까지 할지 정해줘야 한다. 무작정 계속 시킬 수는 없으니까. 어차피 파라미터를 계속 조정 하다보면 어느정도 최적의 값으로 수렴(converge)한다.



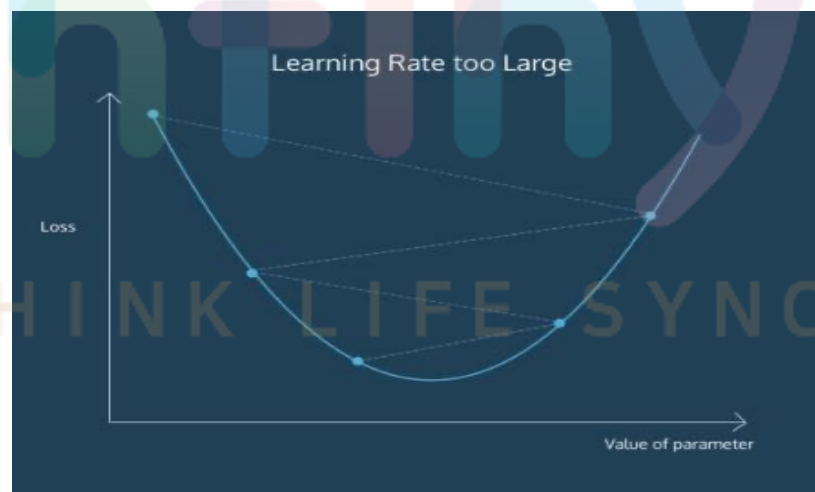
위 그림을 보면 1000번 반복하니깐 b값이 결국 약 47에 수렴하는 걸 알 수 있다. 그 이상 시도하는 건 별로 의미가 없어진다.

이걸 우리가 어떻게 결정할까? 어차피 머신러닝 알고리즘이 알아서 잘 수렴할 거니 걱정 말자.

인공 지능 신경망 개념

학습률(Learning Rate)

우리는 학습률(Learning Rate)이라는 걸 정해줄 필요가 있다.
아래 그림을 보면 학습률이 너무 커서 파라미터를 듬성듬성 조정한다.



이렇게 학습률을 크게 설정하면 최적의 값을 제대로 찾지 못한다. 일을 대충하는 거다. 대신 일을 빨리 하긴 하겠지.

그렇다고 학습률을 작게 설정하면 최적의 값으로 수렴할 때까지 시간이 오래 걸린다.
그래서 모델을 학습시킬 때는 최적의 학습률을 찾는 게 중요하다. 효율적으로 파라미터를 조정하면서도 결국 최적의 값을 찾아한다.

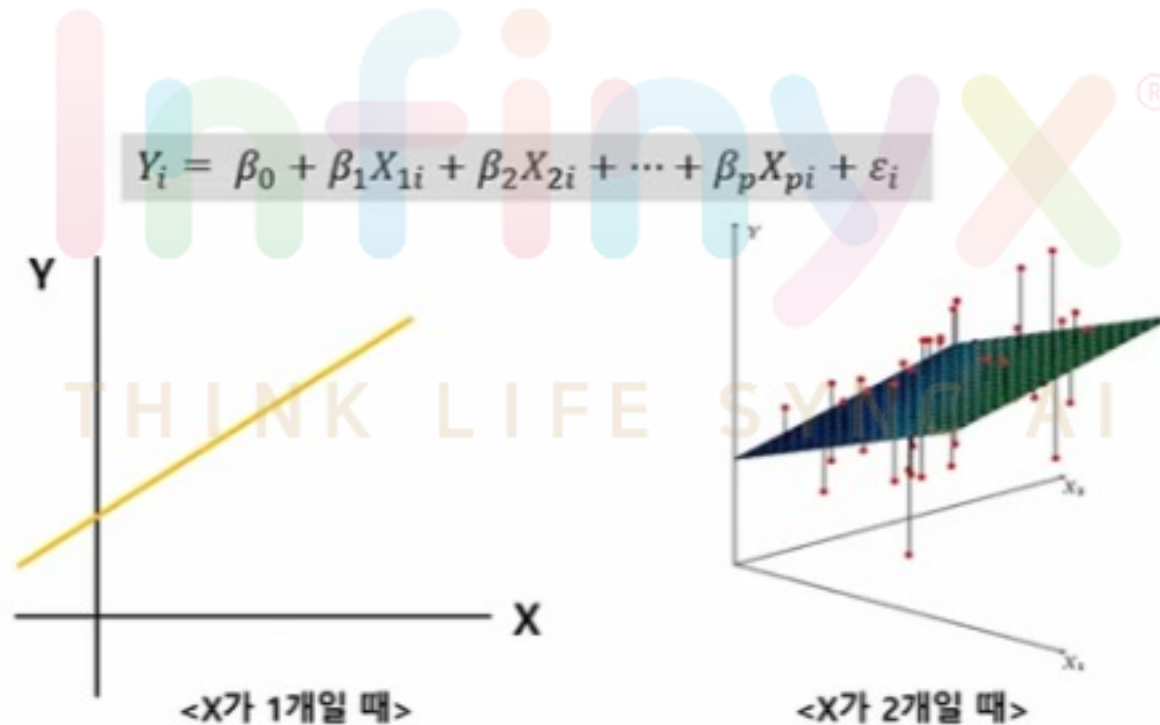
인공 지능 신경망 개념

선형회귀 (Linear Regression) Pytorch 구현 실습



인공 지능 신경망 개념

- 다중선형회귀 (Multiple Linear Regression) 개념
여러 개의 특성을 이용해 종속변수를 예측하기 때문에 일반 선형회귀보다 더 좋은 성능을 기대할 수 있습니다.



인공 지능 신경망 개념

- 다중선형회귀 (Multiple Linear Regression) 구성

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4$$

다중 선형회귀분석의 일반식

- b_1, b_2, \dots 를 편 회귀 계수(Partial regression coefficient)라고 합니다.
- 계수를 통해 설명변수 x_1 이 종속변수에 대한 영향력을 나타냅니다.
- 즉 독립변수들이 설명변수에 얼마나 영향을 주는지 알 수 있게 해 주며, 이를 통해 예측을 할 수 있게 합니다.
- 단순회귀와 마찬가지로 계수는 최소제곱법을 통해 편미분하여 계수를 추정합니다.

인공 지능 신경망 개념

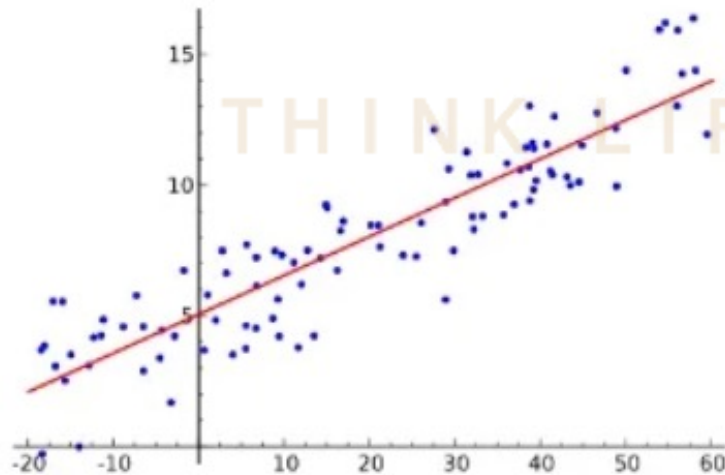
다중 선형회귀 (Linear Regression) Pytorch 구현 실습



인공 지능 신경망 개념

- 이진 분류(Logistic Regression)

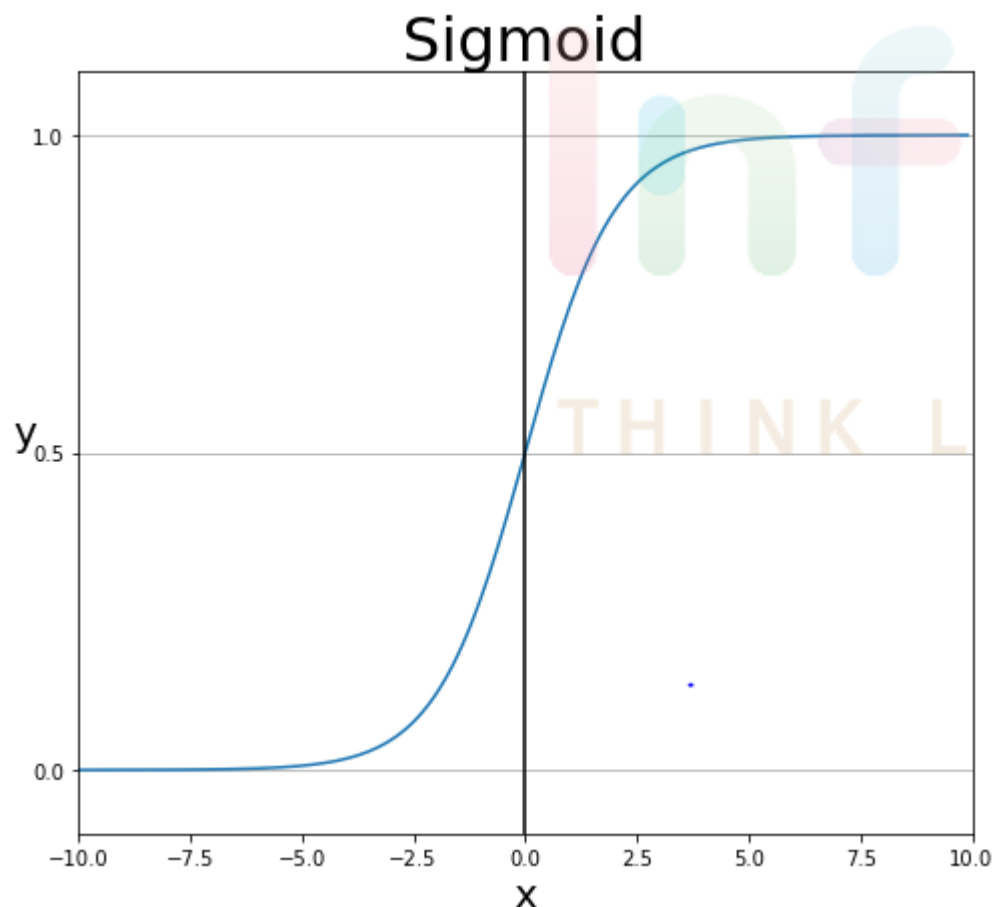
입력값에 따라 모델이 분류한 카테고리가 두 가지인 분류 알고리즘입니다. 주로 어떤 대상에 대한 규칙이 참(True) 인지 거짓(False)인지를 분류하는데 쓰입니다. 예를 들면 암 종양을 분류하는 모델은 어떤 종양을 입력으로 받았을 때 이 종양이 암 종양이 암 종양인지(True) 암 종양이 아닌지(False) 분류합니다.



$$\hat{y} = w \times x + b$$

인공 지능 신경망 개념

- 시그모이드 함수(Sigmoid function)



시그모이드 함수는 0에서 1 사이의 함수이며, 값이 들어왔을 때, 0~1 사이의 값을 반환한다.

연속형 데이터이기 때문에 계단 함수가 끊기지 않는 매끄러운 모양으로 바뀐 것을 알 수 있다.

동시에 이상치가 들어온다 할지라도, 시그모이드 함수는 0과 1에 수렴하므로, 이상치 문제도 해결하면서, 연속된 값을 전달할 수 있다.

시그모이드 함수를 활성화 함수로 사용하면, 0과 1에 가까운 값을 통해 이진분류를 할 수 있습니다.

인공 지능 신경망 개념

- 로지스틱 회귀의 손실 함수

로지스틱 회귀의 경우 손실 함수로 이진교차 엔트로피(Binary Cross-Entropy)를 사용합니다. 이 함수는 다음과 같은 형태 입니다.

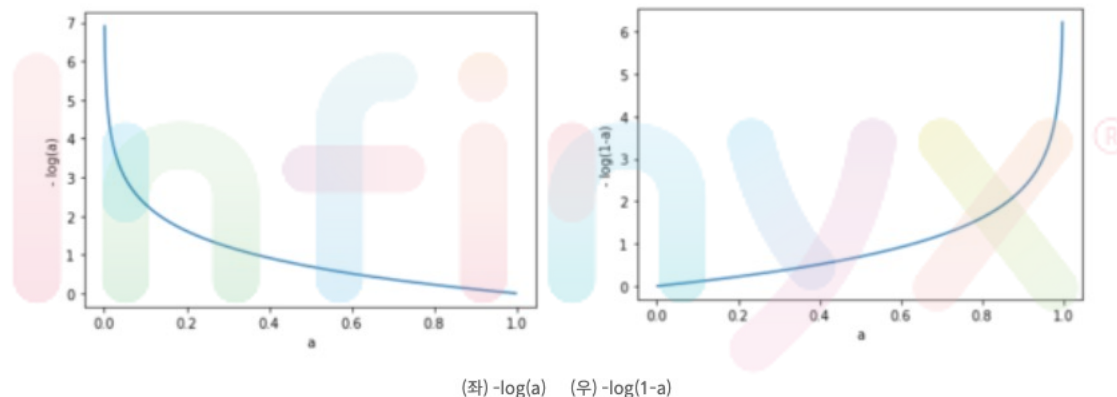
$$L = -\frac{1}{n} \sum_{i=1}^n (y \log(a) + (1 - y) \log(1 - a))$$

a 는 활성화 함수를 통과한 값입니다. 타겟, 즉 y 값은 0 또는 1이기 때문에 i 번째 데이터에 대한 손실은 다음과 같은 값을 가집니다.

$$L_i = \begin{cases} -\log(a) & (y_i = 1) \\ -\log(1 - a) & (y_i = 0) \end{cases}$$

인공 지능 신경망 개념

- 로지스틱 회귀의 손실 함수



시그모이드 함수를 통과하고 나온 a 는 0부터 1까지의 값을 가졌었습니다. 위의 그래프는 각각 $-\log(a)$ 의 그래프와 $-\log(1-a)$ 의 그래프를 그린 것입니다. 그래프에서 알 수 있듯이 $-\log(a)$ 는 a 값이 1에 가까워질수록 그 값이 작아집니다. 반대로 $-\log(1-a)$ 의 그래프는 a 값이 0에 가까워질수록 값이 작아집니다.

즉 손실 함수의 값을 최소화하는 과정에서 타깃이 참(True)이라면 $-\log(a)$ 의 값이 작아져 a 값이 1에 가까워지고, 타깃이 거짓(False)이라면 $-\log(1-a)$ 의 값이 작아져 a 값이 0에 가까워지게 됩니다. 이제 마지막으로 손실 함수의 값을 최소로 만들기 위해 가중치와 바이어스를 찾아야 합니다.

인공 지능 신경망 개념

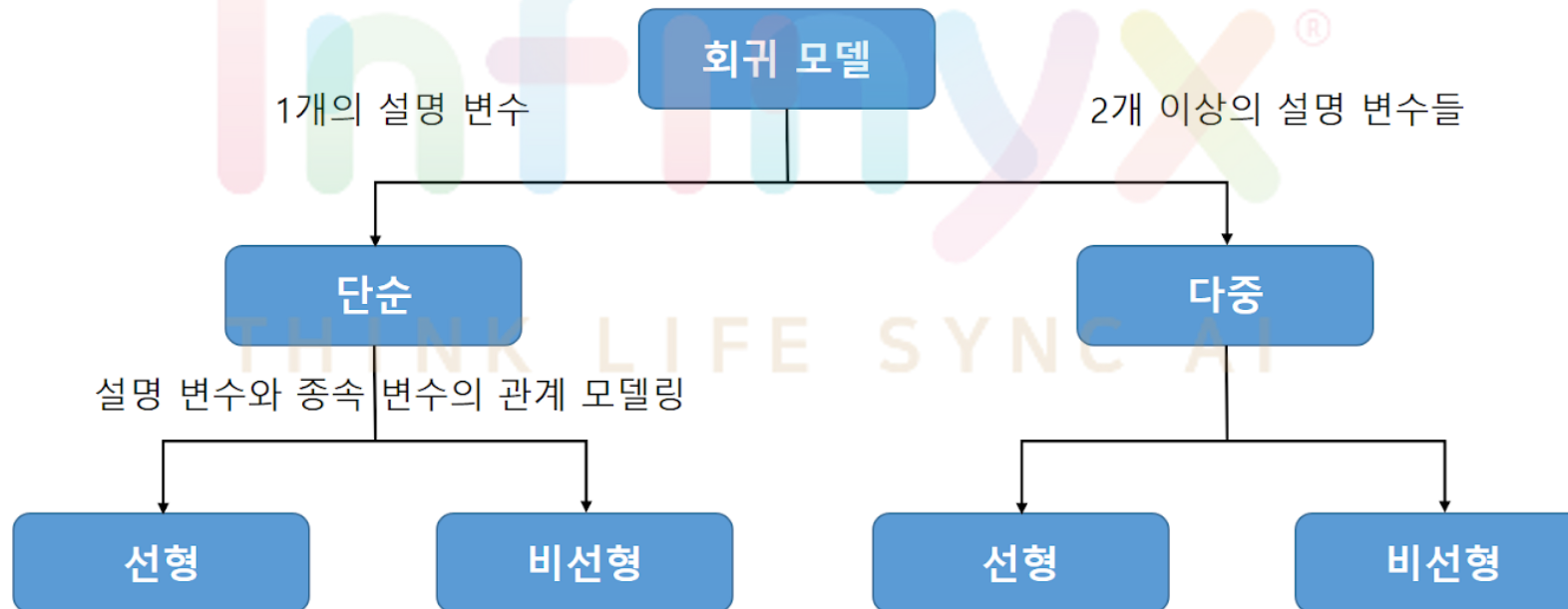
이진분류 실습



인공 지능 신경망 개념

- 회귀의 종류 정리

회귀의 모델의 종류는 다음과 같습니다.



인공 지능 신경망 개념

- 참고사항 Pytorch 에서 다음과 같은 오차 함수를 제공

오차 함수	설명
L1 Loss	데이터 정규화에 사용된다.
MSE Loss	회귀에서 주로 사용되는 함수로 오차의 평균을 제공해서 최종 오차를 계산한다. Vanishing Gradient 가 발생할 수 있다.
Cross-Entropy Loss	이진 분류와 다중 분류 문제에서 사용되는 오차 함수이다.
NLL Loss	분류 문제에서 특정 가중치를 사용해 데이터셋 불균형을 처리할 때 사용된다.
NLL Loss2d	앞서 언급한 NLL Loss 와 유사하지만, 주로 이미지 분할과 관련된 문제에서 픽셀 단위 분류에 사용된다.

인공 지능 신경망 개념

- 참고사항 기술 동향



<https://ai.googleblog.com/>



<https://ai.facebook.com/>



<https://www.amazon.science/>



<https://ai.sony/publications/>

최신 AI 기술 분야별 논문 확인 가능 <https://paperswithcode.com/>

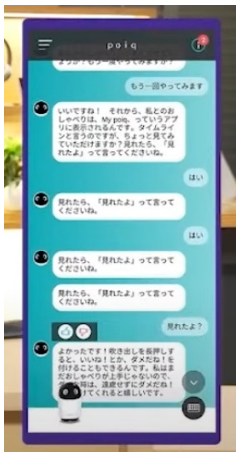
학회에 나오기전 논문 확인 가능 한 사이트 <https://arxiv-sanity-lite.com/>

인공 지능 신경망 개념

- 참고사항 기술 동향 : 반려 로봇 관련 이야기



- 소니 : 반려로봇 포이크(poiq)
- 사용자의 대화를 통해서 학습을 진행
 - 그 사람에 맞춰서 로봇마다 성격을 맞춰 가는 것이 특징
 - 움직임 감지 및 자율 이동 가능



학습 방법 : 앱을 통한 대화 리스트 확인체크후 나쁜 대화 제외하고 학습하고자 하는 문구를 학습을 진행가능

구하는 방법 : 육성자(신청자 5만원 제품을 획득) 학습을 시키면서 육성자 재미난 방식 -> 학습을 시키는 방식

테슬라도 같은 방식 아마존 아스트로 로봇도(300만원) 시제품 주고 반영하는 방식의 마케팅

<https://www.youtube.com/watch?v=sj1t3msy8dc> -> 아스트로 로봇 영상



감사합니다.

THINK LIFE SYNC AI

Infinyx®