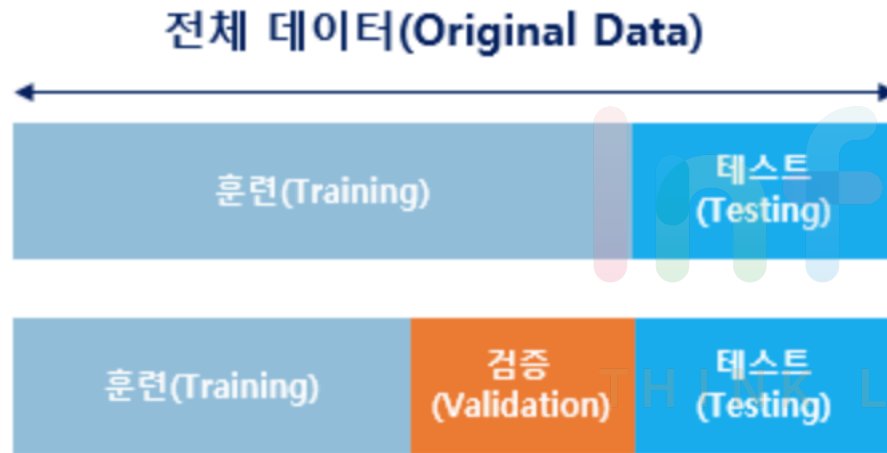


퍼셉트론

AI

머신 러닝 용어 정리



훈련 (Training) : 모델 학습을 위한 데이터

검증 (Validation) : 모델 검증용이 아닌 모델 성능을
조정하기 위한 용도

테스트(Testing) : 모델 검증용 데이터

머신 러닝 용어 정리

분류 (Classification) 와 회귀(Regression)

분류

1) 이진 분류 문제(Binary Classification)

- 주어진 입력에 대해서 둘 중 하나의 답을 정하는 문제 입니다 . 즉 시험 합격 여부 생각하시면 됩니다.

2) 다중 클래스 분류(Multi-class Classification)

- 주어진 입력에 대해서 세 개 이상의 정해진 선택지 중에서 답을 정하는 문제 입니다.

회귀

3) 회귀 문제 (Regression)

- 회귀 문제는 분류 문제처럼 0 또는 1 분리된(비연속적인) 답이 결과가 아니라 연속된 값을 결과로 가짐

머신 러닝 용어 정리

지도 학습 (Supervised Learning) 과 비지도 학습 (Unsupervised Learning)

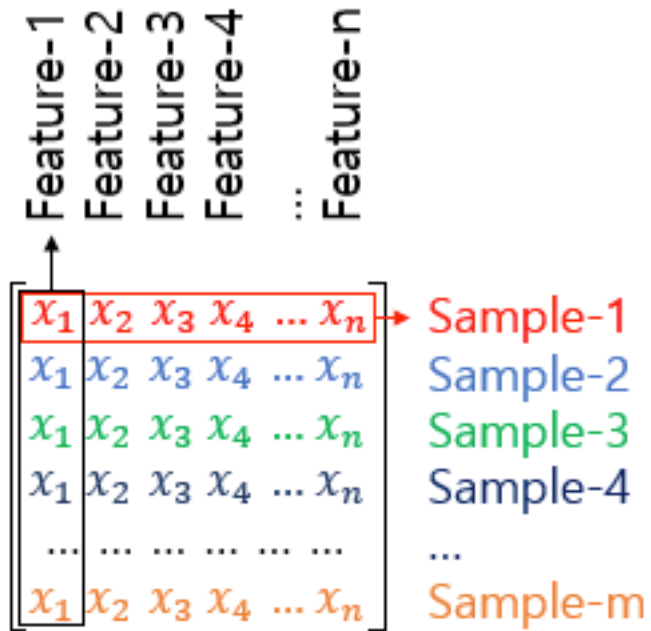
1) 지도 학습 : 레이블(Label)이라는 정답과 함께 학습하는 것을 말합니다.

2) 비지도 학습 : 기본적으로 목적 데이터(또는 레이블)이 없는 학습 방법입니다.
대표적으로 군집(clustering) 이나 차원 축소와 같은 학습 방법들을 비지도 학습이라고 합니다.

3) 강화 학습 : 어떤 환경 내에서 정의된 에이전트가 현재의 상태를 인식하여, 선택 가능한 행동들 중 보상을 최대화하는 행동 혹은 행동 순서를 선택하는 방법입니다.

머신 러닝 용어 정리

샘플(Sample)과 특성(Feature)



머신 러닝에서 하나의 데이터, 하나의 행을 샘플(Sample)이라고 부릅니다.
(데이터베이스에서는 레코드라고 부르는 단위입니다.)

종속 변수 y 를 예측하기 위한 각각의 독립 변수 x 를
특성(Feature)이라고 부릅니다.

머신 러닝 용어 정리

혼동 행렬(Confusion Matrix)

-	참	거짓
참	TP	FN
거짓	FP	TN

TP는 양성(Positive)이라고 대답하였고 실제로 양성이라서 정답을 맞춘 경우입니다.

TN은 음성(Negative)이라고 대답하였는데 실제로 음성이라서 정답을 맞춘 경우입니다.

THINK LIFE SYNC AI

$$\text{정밀도} = \frac{TP}{TP + FP}$$

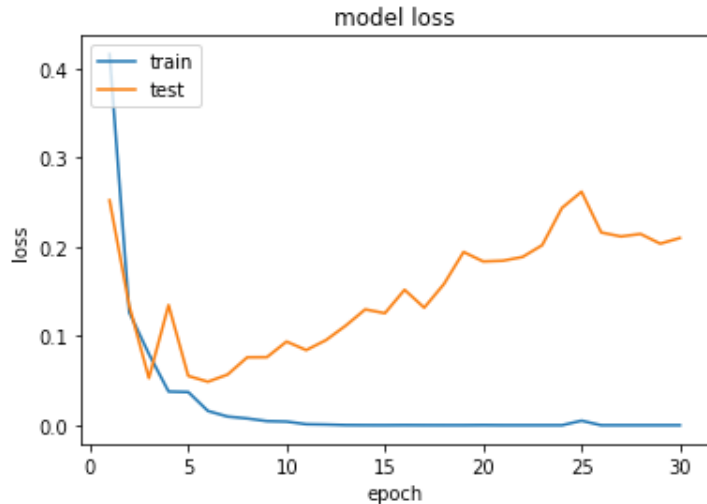
양성이라고 대답한 전체 케이스에 대한 TP 비율입니다.

$$\text{재현률} = \frac{TP}{TP + FN}$$

실제값이 양성인 데이터의 전체 개수에 대해서 TP 비율입니다.
즉 양성인 데이터 중에서 얼마나 양성인지를 예측(재현)했는지를 나타냅니다.

머신 러닝 용어 정리

과적합(Overfitting) 과 과소 적합(Underfitting)



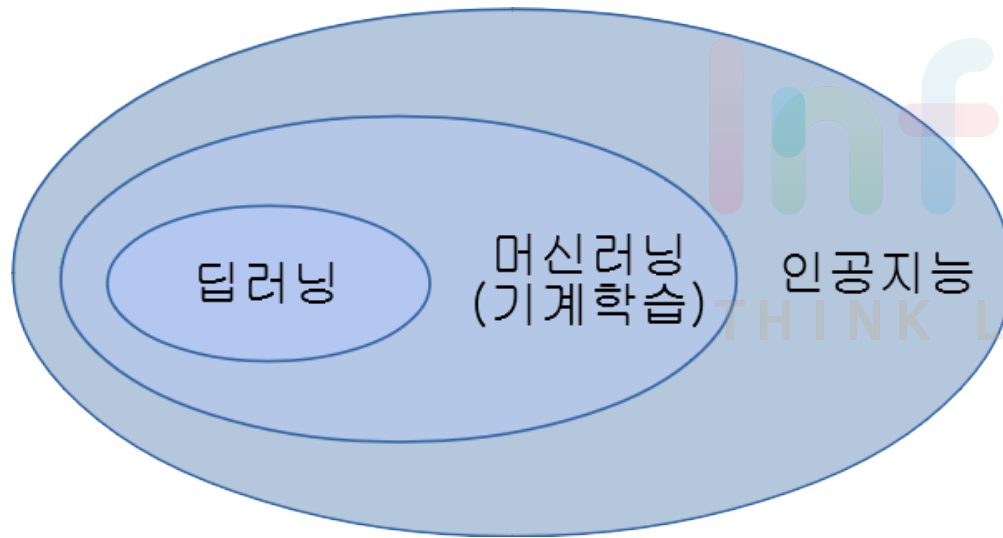
과적합(Overfitting) : 훈련 데이터를 과하게 학습한 경우를 말합니다.
즉 기계가 훈련 데이터에 대해서만 과하게 학습하면 테스트 데이터나 실제 서비스에서의 데이터에 대해서는 정확도가 좋지 않은 현상이 발생합니다.

과적합 상황에서는 훈련 데이터에 대해서는 오차가 낮지만
테스트에 대해서는 오차가 높아지는 상황이 발생합니다.

과소적합(Underfitting) : 훈련 자체가 부족한 상태이므로 과대 적합과는 달리 훈련 데이터에 대해서도 보통정확도가 낮다는 특징이 있습니다.

과소 적합을 막는 방법은 드롭아웃, 조기종료 같은 방법이 존재합니다.

인공지능과 머신러닝, 딥러닝



› 인공지능

- 지식 기반 접근
- 데이터 기반 접근

› 머신러닝

- 데이터 기반 접근

› 딥러닝

- 데이터 기반 접근

지식 기반 접근 vs 데이터 기반 접근(머신러닝)

․ 지식 기반 접근

- 초창기 열악한 컴퓨팅 환경
- 인공지능 연구 초기에 주류
- 문제 영역의 핵심 지식을 추출하여 기호로 표현해 풀어내는 방식
- 전문가가 갖는 풍부하고 섬세하며 때로 추상적이거나 막연하기까지 한 전문지식을 모순 없이 기호로 표현하는 일은 불가능

THINK LIFE SYNC AI

․ 데이터 기반 접근(머신러닝)

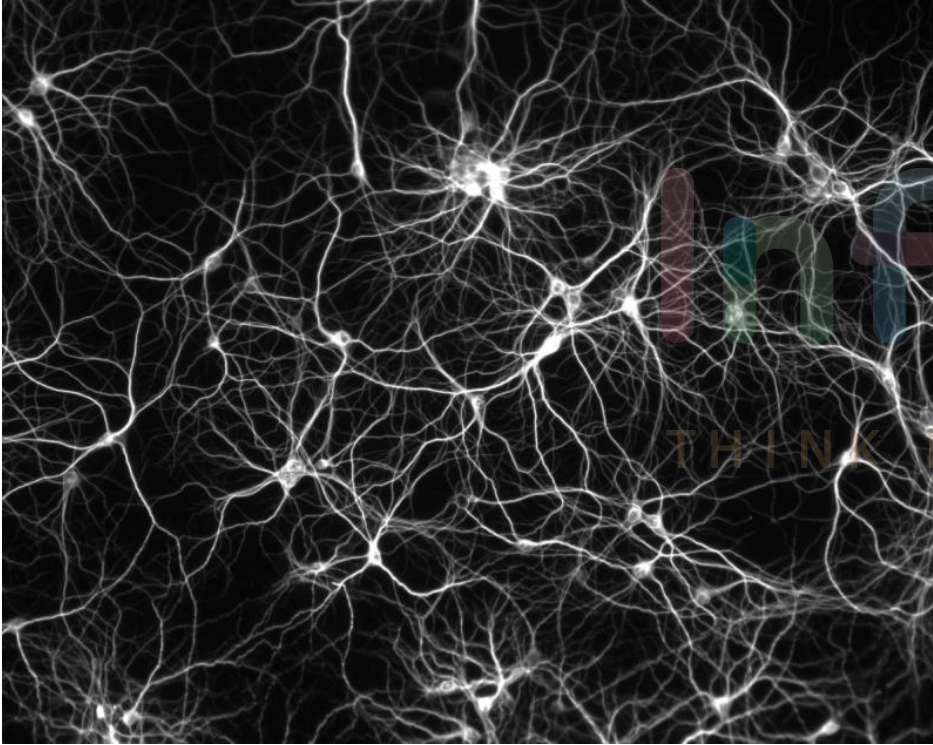
- 머신러닝 프로그램이 직접 데이터를 분석하여 숨어 있는 규칙이나 패턴을 포착해 문제를 해결
- 데이터로부터 규칙이나 패턴을 획득하는 과정을 학습이라 함

딥러닝에서의 데이터 기반 접근

› 데이터 기반 접근(딥러닝)

- 동물의 신경세포(뉴런)를 흉내 낸 퍼셉트론을 단위로 삼아 구성한 신경망 구조를 이용해 학습을 수행
- 인공지능 초창기부터 연구가 진행되었으나 여러 어려움 존재
- 2010년 무렵 부터 여러 성공적인 결과
- 알파고와 이세돌의 바둑 대결이 딥러닝의 이름을 널리 알리는 계기

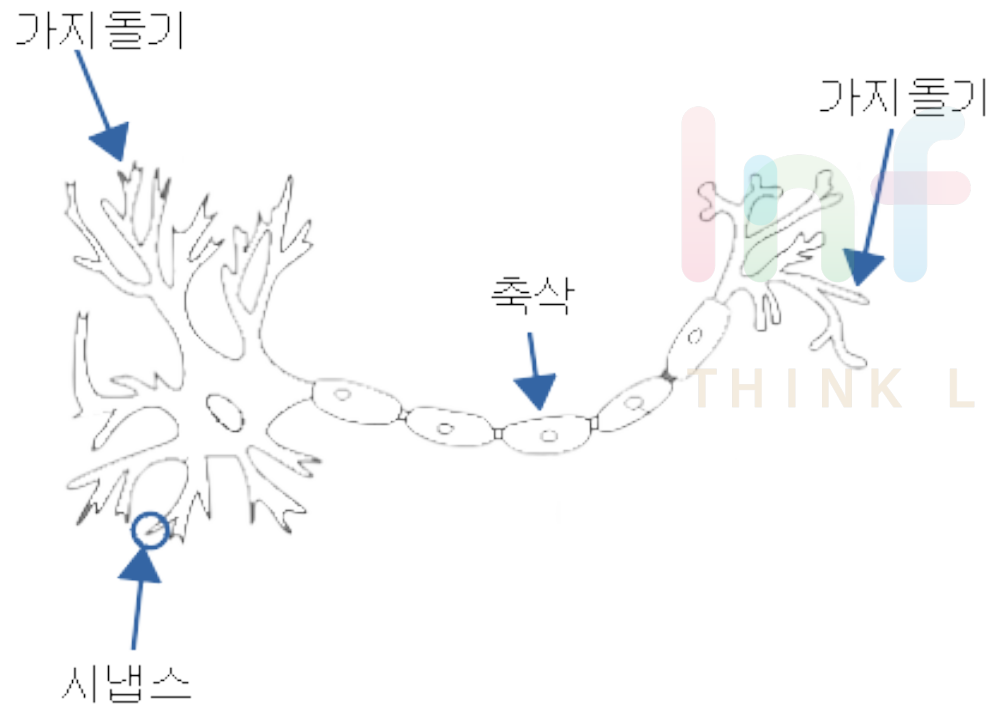
동물의 신경 세포, 뉴런(neuron)



정보 전달용 신경세포(뉴런)

- › 몸의 내부와 외부에 자극을 가하게 되면 일련의 과정을 통해 뉴런은 자극을 전달하게 되며, 최종적으로 척수와 뇌 등의 중추신경계로 도달하게 되며 중추신경계에서 처리한 정보를 다시 우리 몸으로 전달해 명령을 수행
- › 자극(impulse)의 전도와 전달

동물의 신경 세포, 뉴런(neuron)

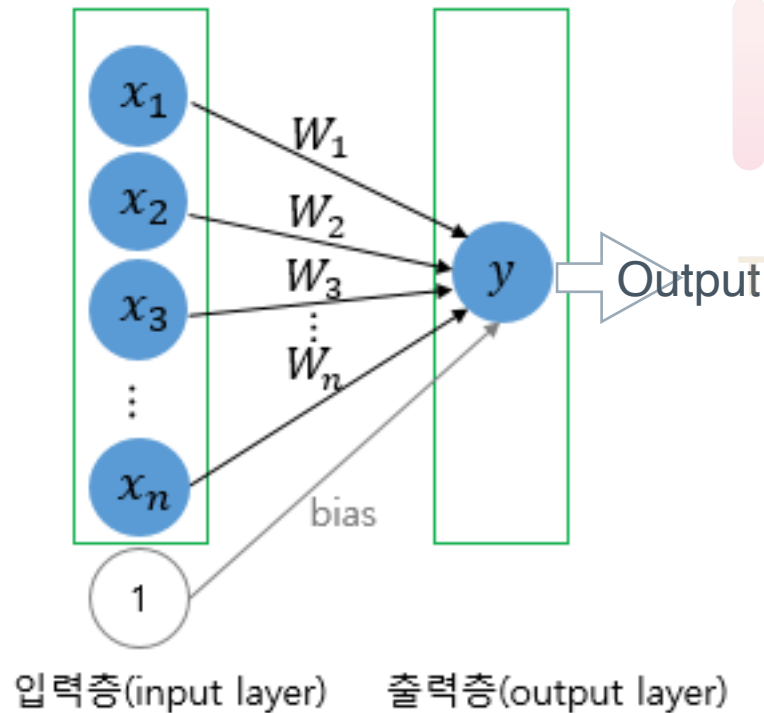


정보 전달용 신경세포(뉴런)

- 세포체 주변에 돋은 가지돌기들을 통해 다른 뉴런들로부터 전기 신호를 받음
- 전기 신호는 시냅스라는 화학적 연결 부위를 통해 전달
- 시냅스 연결 부위는 사용 빈도에 따라 발달 정도가 달라지며 같은 신호라도 더 강하게 전달
- 뉴런은 각 가지돌기로부터 전달된 전기 신호들로부터 자신의 출력 신호를 만들어내어 축삭돌기와 축삭돌기에 무수히 붙은 가지돌기를 통해 다른 뉴런에 전달
- 출력 신호는 입력 신호값의 합이 어떤 임계값 이상이면 활성화 상태가 되어 신호를 전달

단층 퍼셉트론 (Single-Layer Perceptron)

단층 퍼셉트론은 값을 보내는 단계와 값을 받아서 출력하는 두 단계로 이루어집니다.



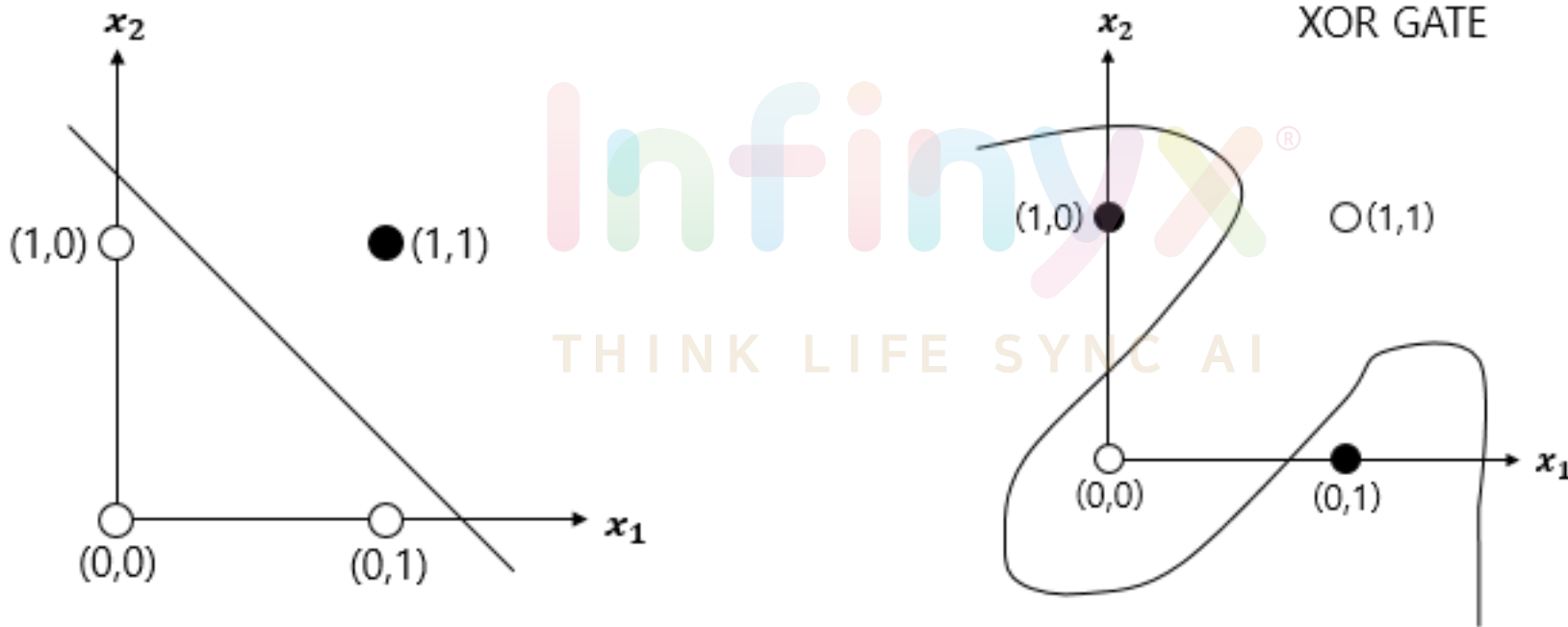
입력층은 학습 벡터 또는 입력 벡터가 입력되는 계층으로써,
입력된 데이터는 출력층 뉴런으로 전달되어 활성화함수에 따라 값이 출력

출력층 퍼셉트론 설계 시 임의의 N개의 뉴런으로 구성

단층 퍼셉트론에서는 활성화 함수가 사용되지 않음

단층 퍼셉트론 (Single-Layer Perceptron)

단층 퍼셉트론 한계점



단층 퍼셉트론은 직선 하나로 두영역을 나눌 수 있는 문제에 대해서만 구현이 가능합니다.

단층 퍼셉트론 (Single-Layer Perceptron)

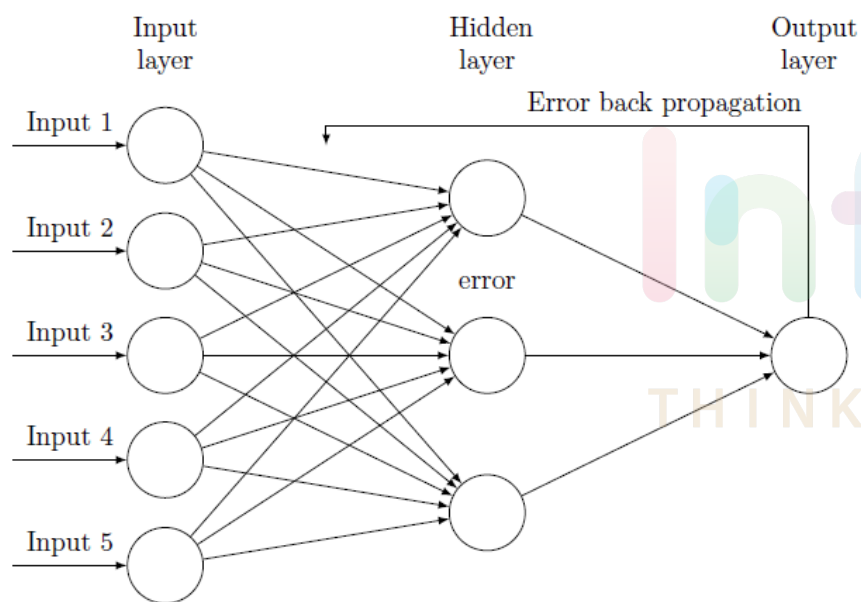
단층 퍼셉트론 실습

1. 단층 퍼셉트론 이용한 AND NAND OR 게이트 구현



THINK LIFE SYNC AI

역전파 (BackPropagation)



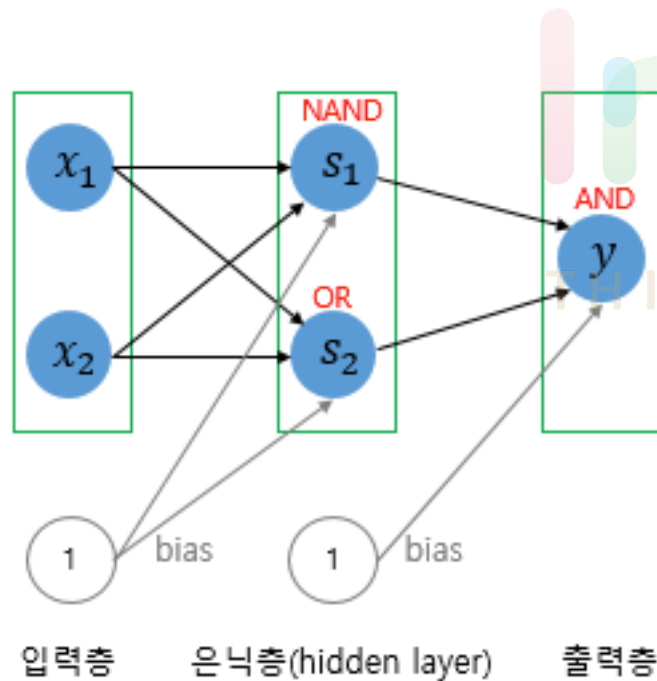
Input 이 들어오는 방향 (순전파)으로 Output layer 에서 결과 값이 나옵니다.

결과값은 오차를 가지게 되는데 역전파는 이 오차를 다시 역방향으로 hidden layer 와 input layer로 오차를 다시 보내면서 가중치를 계산 하면서 output에서 발생했던 오차를 적용시키는 것입니다.

한 번 돌리는 것을 1 epoch 주기라고 하며 epoch를 늘릴 수록 가중치가 계속 업데이트(학습)되면서 점점 오차를 줄여나가는 방법 입니다.

다층 퍼셉트론 (MultiLayer Perceptron, MLP)

XOR 게이트는 기존 AND , NAND , OR 게이트를 조합하면 만들 수 있습니다.
퍼셉트론 관점 : 층을 더 쌓으면 만들 수 있습니다.

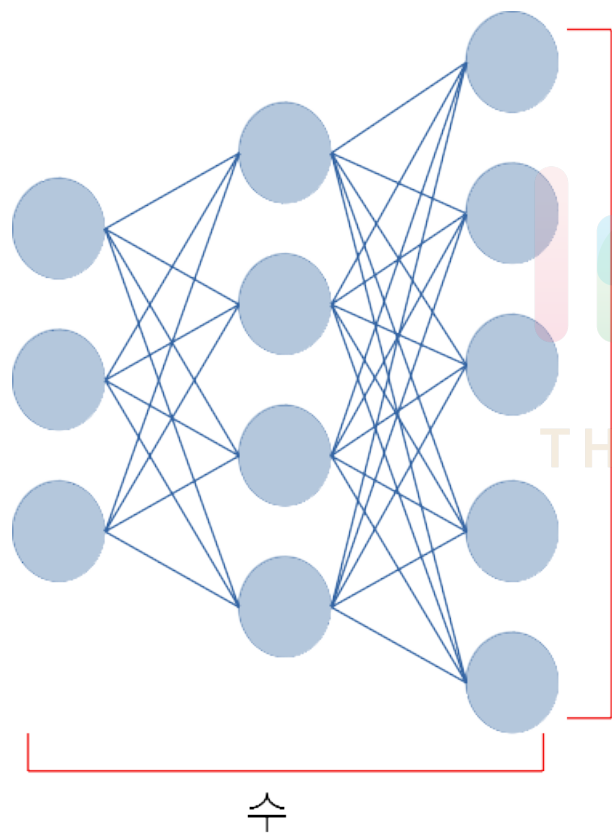


은닉층 hidden Layer : 입력층과 출력층 사이에 존재하는 층
즉 다층 퍼셉트론은 중간 은닉층이 존재한다는 점이 단층 퍼셉트론과 다릅니다.

복수의 퍼셉트론 계층을 순서를 두고 배치하여 입력 벡터로부터 중간 표현을 거쳐 출력 벡터를 얻어내는 신경망 구조

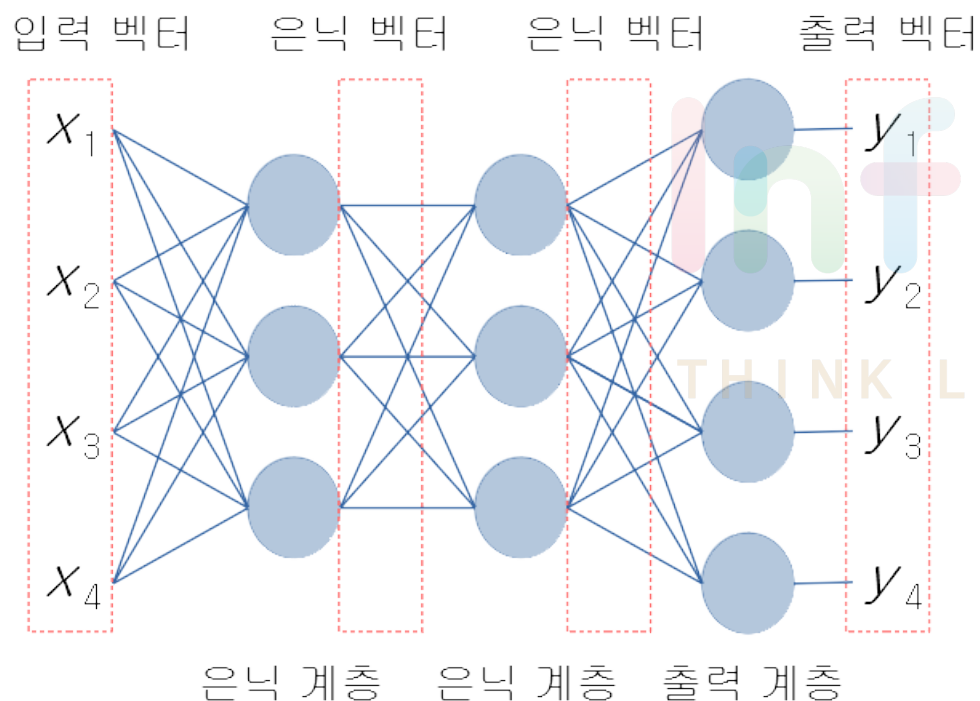
인접한 계층끼리는 앞 계층의 출력이 뒤 계층의 모든 퍼셉트론에 공통 입력으로 제공

은닉 계층의 수와 폭



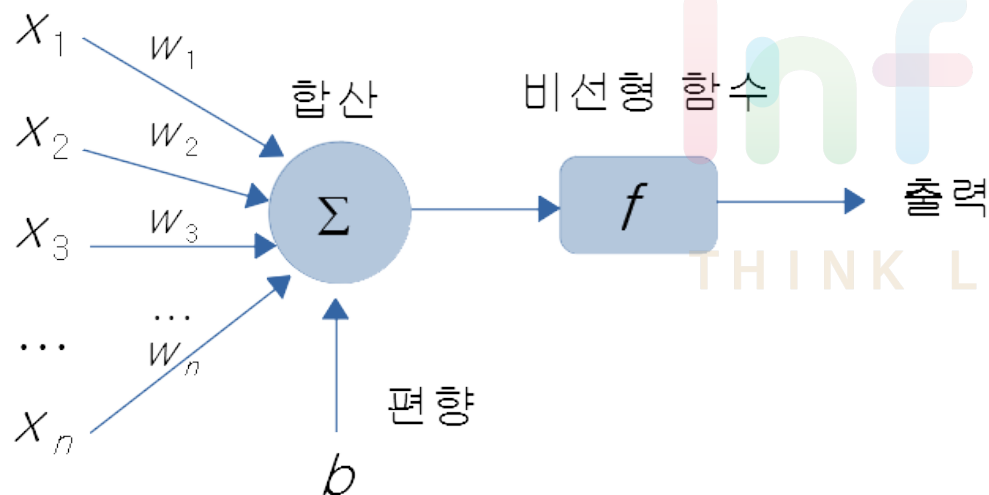
- 출력 계층
최종 단계에 배치된 출력 계층은 신경망에 주어진 원래의 임무에 따라 알맞은 형태의 출력 벡터를 생성하는 역할, 그러므로 출력 벡터의 크기(퍼셉트론의 수)는 문제의 성격에 따라 고정적으로 정해짐
- 은닉 계층
출력 계층과 같은 제약이 없고 은닉 계층의 수와 각 은닉 계층의 폭은 신경망 설계자가 자유롭게 정할 수 있음

은닉 계층의 수와 폭



- 은닉 계층의 수와 각 은닉 계층의 폭은 신경망의 품질을 결정짓는 중요한 요인이 될 수 있지만 무조건 은닉 계층의 수나 폭을 늘린다고 품질이 좋아지는것은 아님
- 은닉 계층을 추가해 학습 파라미터 수가 늘어나면 더 많은 학습 데이터가 필요해지는 경향이 존재

비선형 활성화 함수



- › 은닉 계층은 가중치와 편향을 이용해 계산된 선형 연산 결과를 바로 출력으로 내보내는 대신 한 번 더 변형시켜 내보낸다. 선형 연산 결과 뒷단에 적용되어 퍼셉트론의 출력을 변형시키려고 추가한 장치를 비선형 활성화 함수라고 한다

출력층의 활성화 함수와 오차 함수의 관계

각 문제에 따른 출력층의 활성화 함수와 비용 함수의 관계를 정리 해보면 다음과 같습니다.

문제	활성화 함수	비용 함수
이진 분류	시그모이드	<code>nn.BCELoss()</code>
다중 클래스 분류	소프트맥스	<code>nn.CrossEntropyLoss()</code>
회귀	없음	MSE

주의할 점은 `nn.CrossEntropyLoss()`는 소프트맥스 함수를 이미 포함하고 있습니다.

회귀 문제 등 `nn.CrossEntropyLoss()`를 사용하지 않을 때는 소프트맥스 함수를 이미 포함하고 있습니다.

회귀

소프트맥스

MSE

다층 퍼셉트론 실습

1. 파이토치로 다층 퍼셉트론 구현하기
2. 사이킷런 패키지에서 제공하는 분류용 데이터를 활용하여 다층 퍼셉트론으로 손글씨 분류하기

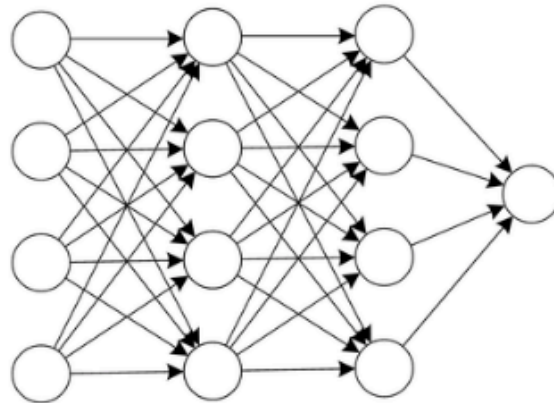


THINK LIFE SYNC AI

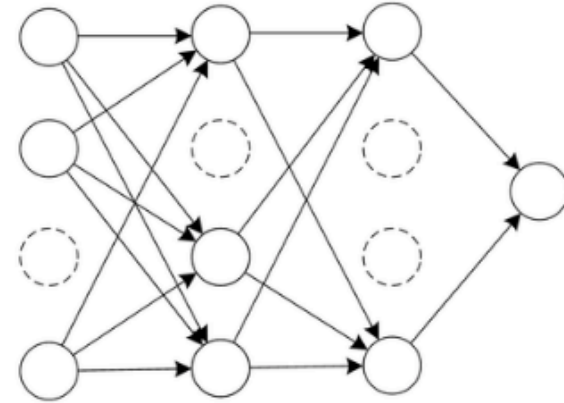
과적합을 막는 방법

1. 데이터 양을 늘리기 : 데이터 양이 적을 경우 해당 데이터 특정 패턴이나 노이즈까지 암기하기 되므로 과적합 현상이 발생할 확률이 높아집니다.
2. 모델의 복잡도 줄이기 : 인공 신경망의 복잡도는 은닉층의 수나 매개변수의 수 등으로 결정됩니다. 과적합 현상이 포착되었을 때, 신경망에서 해줄수 있는 방법은 복잡도를 줄이는 방법입니다.
3. 가중치 규제 적용하기 (Normalization)
4. 드롭아웃(Dropout)

예를 들어 드롭아웃의 비율을 0.5로 한다면 학습 과정마다 랜덤으로 절반의 뉴런을 사용하지 않고, 절반의 뉴런만을 사용합니다.



(a) Standard Neural Network



(b) Network after Dropout