

인공지능 신경망 Basic

AI - 스쿨

THINK LIFE SYNC AI

2022. 06. 02.

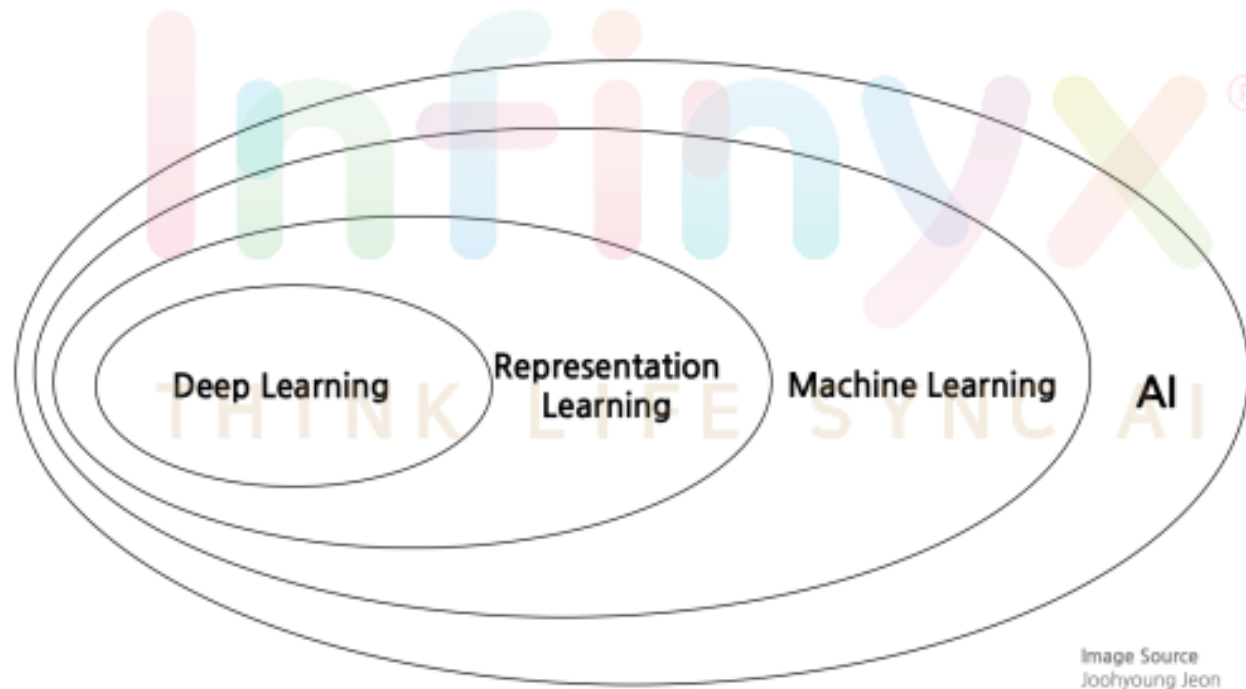


인공신경망 Basic

1. 인공지능이란 ?
2. 인공신경망의 용어
3. 실제 서비스에서 사용되는 다양한 신경망 소개
4. 퍼셉트론

인공 지능이란 ?

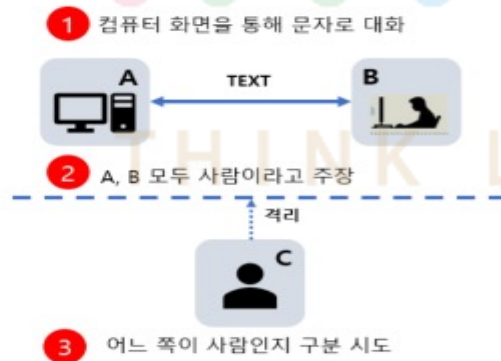
AI라는 큰 개념이 포함하고 있는 하위 3개의 세부속성으로 분류한 다이어그램입니다.



AI 이란 ?

요즘 말하는 AI(Artificial Intelligence)라는 개념은 오래전부터 존재 하였고, 과학자들의 무수한 연구 대상이었습니다.

AI의 명확한 정의는 무엇일까요? 단순히 기계가 사람처럼 움직이면 AI라고 할 수 있을까요?



튜링 테스트는 기계(A)와 사람(B)의 대화를 제3자(C)가 듣고 선, 제3자(C)가 기계(A)와 사람(B) 중 누가 기계인지 분별하기 어려운 경우, 해당 기계(A)는 튜링 테스트를 통과하게 됩니다. "컴퓨터로부터의 반응을 인간과 구별할 수 없다면, 컴퓨터는 생각 할 수 있는 것"이라고 말이죠. 이러한 튜링 테스트의 관점에서 AI란 기계에게 "지식"과 "경험"을 가르쳐주는 것이고 이를 통해 사람 처럼 행동하는 모든 유형의 기계를 의미한다고 할 수 있습니다.

Image Source

[좌] 위키백과 https://ko.wikipedia.org/wiki/%EC%95%A8%EB%9F%B0_%ED%8A%9C%EB%A7%81

[우] 온라인 검색 "튜링게임" <http://blog.skby.net/%ED%8A%9C%EB%A7%81-%ED%85%8C%EC%8A%A4%ED%8A%B8/>

머신 러닝이란 ? Machine Learning

AI 시스템으로 거듭나기 위해서는 지식과 경험을 가르쳐줘야 한다고 했습니다. 과연, 어떻게 할 수 있을까요? 왠지 그것에 대한 해답은 Machine Learning에 있을 것 같지 않나요?

"컴퓨터 프로그램이 특정 업무(T)를 수행할 때 성능(P)만큼 개선되는 경험(E)를 보이면 컴퓨터 프로그램은 해당 업무(T)와 성능(P)에 대해 경험(E)를 학습했다고 할 수 있다."

컴퓨터에게 사과와 딸기 이미지를 인식하고 분류하는 업무를 부여한다고 했을 때, 머신러닝의 관점에서는

T : 사과/딸기 이미지를 인식하고 분류하는 것

P : 사과/딸기 이미지를 정확하게 구분한 확률

E : 사과/딸기 이미지를 표현한 Matrix 와 label 값(사과 : 1/ 딸기 : 0)을 가지고 있는 데이터셋 입니다.

컴퓨터에게 E를 계속 보게 하여, T라는 업무를 수행함에 있어 과거의 P가 50%였던 것을 80%까지 끌어 올렸다면, 미첼이 말했던 머신러닝 관점에서는 30%의 성능향상이 있었기에 해당 프로그램은 T업무와 이를 잘 분류한 80%(P)에 대해 경험(E)를 학습했다는 것 입니다. 이것이 머신러닝 입니다.

Representation learning

Representation은 그 뜻 자체로 '표현', '대표' 등의 뜻이 있습니다. 무엇인가를 대표한다는 것이죠. 잠시 머신러닝의 예제였던 사과와 딸기를 인식하고 분류하는 문제로 넘어가겠습니다. 우리는 T라는 과업에서 P의 향상을 위해 E를 학습한다고 했습니다. 여기서 E는 데이터셋 인데요, 머신러닝에서 말하는 representation learning이란 훈련하는 알고리즘이 학습 과정에서 데이터(E)의 중요한 특징을 스스로 파악한다는 것 입니다.

우리는 어렸을 적 낱말카드를 보며 사물을 인식하는 T를 해 왔습니다. 사과, 딸기가 그려진 그림을 보며 맞추는 행동이었죠.

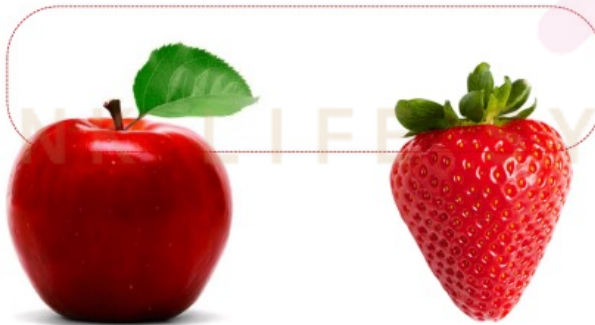


Image Source
[좌] 사과 사진, 구글 검색, <https://kr.123rf.com/>
[우] 딸기 사진, 네이버 블로그, <https://m.blog.naver.com/PostList.nhn?blogid=yogurtman00>

이처럼, Representation learning은 데이터 에서 중요한 특징(Feature)을 알고리즘이 감지하는 것을 의미합니다.

Deep learning

이제야 딥러닝의 영역까지 왔습니다. 앞에서 나온 (1),(2),(3)의 정의를 쭉 정리해보자면 다음과 같습니다.

"무엇이 AI시스템이냐?"

라는 질문에 대답하기 위해 여러 관점 중 앨런 튜링의 '튜링 테스트'를 인용하면, AI란 *"제 3자가, 보이지 않는 두 화자의 대화에서 어떤 화자가 컴퓨터인지 구분할 수 없다면, 그 시스템은 AI시스템이다."*라는 것이 있었습니다. 그러면,

"어떻게 AI 시스템을 만들 것이냐?"

라는 질문에는 기계에게 "지식"과 "경험"을 가르쳐야 한다고 했습니다. 이를 위해 머신러닝 계열의 알고리즘은 데이터를 통해 경험을 학습시켜 줍니다. 또한, 수 많은 경험(데이터)에서 알아서 핵심적인 경험을 추출하고 학습하는 알고리즘을 Representation learning) 이라고 합니다.

딥러닝이라는 기술은 결국 AI라는 광범위한 분야 중, 머신러닝 이면서 동시에 데이터에서 핵심을 잘 선별해내는 기술이라고 할 수 있습니다. 그리고 이러한 분야에 매우 유능한 알고리즘이 하나 등장합니다. 바로 **인공신경망(Artificial Neural Networks)** 이죠.

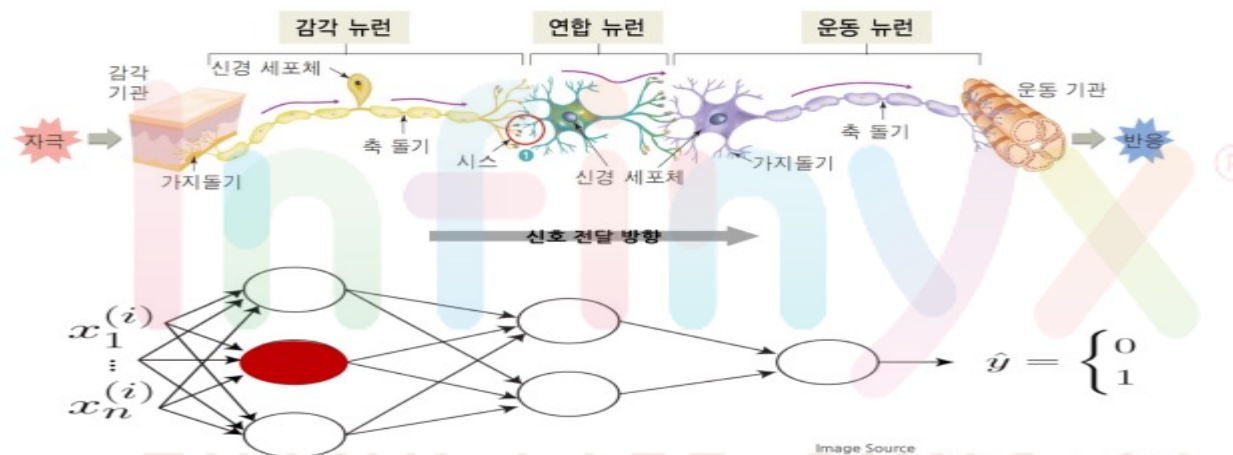
인공신경망 == AI ?

인공신경망은 인간의 신경체계를 모사한 알고리즘 이기도 하기 때문에, 이름 부터가 인공신경망 (Artificial Neural Networks) 입니다.

이러한 이유로, 혹 누군가는 "인공신경망 = AI" 라고 생각하시는 분들도 있는데, "인공신경망 = AI"는 결코 아닙니다. AI라는 거대한 집합(Set)이 있으면 (그것을 구현하기 위한 하나의 기술) 집합의 한 원소(Element)로 인공신경망이 존재하는 것 입니다.

THINK LIFE SYNC AI

인간의 신경망과 인공신경망



인간의 신경망

- 사람의 뇌 속에는 다수의 뉴런들이 네트워크 형태로 구성되어 있습니다. 서로 연결되어 있어 뉴런 하나가 다른 뉴런에게 신호를 전달하는 구조입니다

인간신경망

- 그림 맨 좌측 는 입력신호를 말합니다. 인간의 신경계로 따지면, 감각기관을 통해 받아들여진 일종의 감각입니다. 데이터들은 화살표를 통과하여 다음 원으로 전달됩니다. 데이터를 받아들이는 부분은 일종의 가지돌기 역할을 수행합니다. 그런 다음 신경세포체에게 전달됩니다. 그림에서 신경세포체는 원을 말합니다. 그래서 많은 인공신경망 교과서에는 저 부분을 '뉴런'과 비슷하다고 말합니다.

인공신경망의 용어

인공신경망에서는 어떻게 구현하고 있으며 무엇이라 부르는지 한번 살펴보겠습니다.

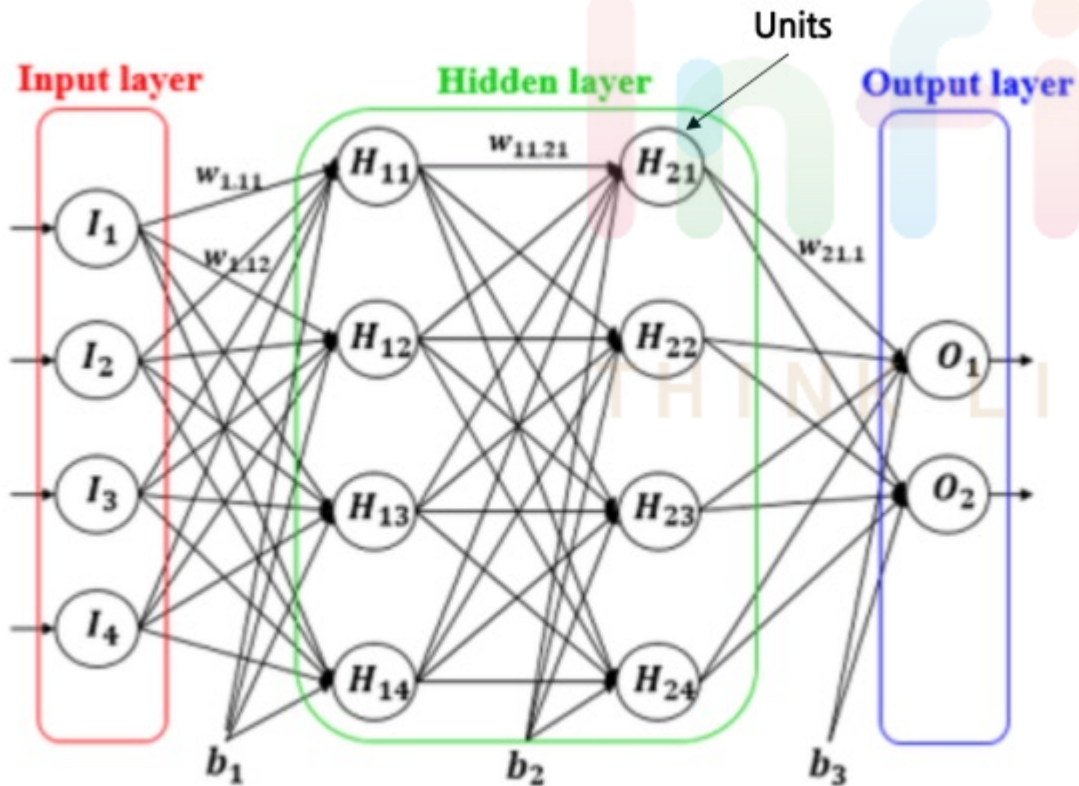


Image Source
<https://prideasanengineer.tistory.com/20>

- **Input Layer (입력층)** : 데이터가 입력되는 계층 (가지돌기의 역할)
- **Hidden Layer (은닉층)** : 데이터가 전달되는 계층 (뉴런-시냅스-축삭돌기-뉴런 등 신호의 전달 계층)
- **Output Layer (출력층)** : 데이터가 출력되는 계층 (운동기관을 통해 반응여부를 결정하는 기능)
- **Units** : 데이터를 받아들여 다음 계층으로 전달할지 판단 (신경망 뉴런의 역할)
- **Weights** : 전/후 Units를 잇는 화살표, 각 Units의 연결강도를 결정하는 가중치 (시냅스 틈을 타고 전달되는 신경 전달 물질의 강도, 특정 감각에 대하여 민감한지 둔감한지 여부)
- **Activation (Function)** : Units에서 다음 신호로 보낼지 판별하는 함수 (뉴런에서 계산된 값이 임계치를 넘는지 여부와 이를 토대로 다음 신호 전달여부를 활성화 할 것인지 판단)

입력층은 인공신경망 전체 구성을 말할 때 빼줍니다.

실제 서비스에 사용되는 다양한 신경망 소개

인공신경망에서는 어떻게 구현하고 있으며 무엇이라 부르는지 한번 살펴보겠습니다.

서비스 유형	딥러닝 모형 (주로 활용되는 알고리즘)
이미지 판별, 객체 인식, 얼굴 인식	CNN (VGG, ResNet, GoogleNet, Inception, R-CNN, YOLO)
챗봇, 음성스피커, 기계번역	RNN, LSTM, Seq2Seq, Attention, BERT
이미지 합성, 이미지 생성, 화자판별	Neural Style Transfer, GAN, DCGAN
자율주행	CNN 계열을 통한 동적/정적 객체 인식, 강화학습 통한 움직임 정의

[표 1] 딥러닝 기반 서비스 유형과 이를 구현하기 위한 기술

실제 서비스에 사용되는 다양한 신경망 소개

1. 이미지 인식

- 이미지 처리와 관련된 대부분의 알고리즘은 CNN(Convolution Neural Networks)을 기반으로 하고 있습니다.

DNN의 경우 망이 깊어질 수록 weights가 많아지며 이는 Computation Power에 큰 영향을 미칩니다.

만약, 여러분 휴대전을 통해 찍은 사진에서 물체를 인식하는 문제를 DNN을 통해 해결해야 한다고 가정하겠습니다. 입력되는 사진의 화소가 $2,048 \text{ px} \times 2,048 \text{ px}$ 이라면, 입력층의 유닛 수는 $2,048 \times 2,048 = 4,194,304$ 개 입니다.

그리고, 바로 다음에 연결된 첫 번째 히든레이어 유닛의 수가 정확하게 1,000개라면 입력층과 바로 다음 히든레이어가 연결될 때 필요한 weights의 개수는 $4,194,304,000 + 1,000(\text{bias})$ 로 대략 41억개 입니다. 입력층 하나와, 히든레이어 딱 하나로 구성된 DNN에서 필요한 weights의 수가 41억개라는 것은 말도 안되는 것 입니다.

CNN은 Convolution 및 Pooling 연산을 반복적으로 수행하여 이미지 좌우 크기의 차원을 축소하고, 3차원 volume을 넓혀가면서 이미지에서 중요한 특징을 추출합니다. 이를 통해 DNN 보다 획기적으로 weights를 줄입니다.

실제 서비스에 사용되는 다양한 신경망 소개

2. 이미지 합성, 이미지 생성

- 이미지 처리와 관련된 대부분의 알고리즘은 CNN(Convolution Neural Networks)을 기반으로 하고 있습니다.

딥러닝을 설명하는 여러가지 글 중에서, 고흐 "별이 빛나는 밤" 등 명화와 일반 사진을 합성한 사진을 보셨을 겁니다. 이처럼 Texture가 되는 이미지를 학습하고(명화) 여기에 새로운 사진을 입력하여, 두 개의 이미지 패턴이 서로 합성된 새로운 이미지를 만들어냅니다. 이러한 딥러닝 모형을 Neural Styler Transfer 이라고 합니다.

연예인 이미지를 학습하고 학습한 연예인 사진과 최대한 유사해 보이는 얼굴을 만들어 냅니다.

이걸 가능하게 해주는 모형이 GAN(Generative Adversarial Networks)입니다. 간략하게 설명드리면, 생성자(Generator)와 판별자(Discriminator)가 서로 경쟁하여 생성자는 자신이 입력받은 데이터와 최대한 유사한 분포를 가진 데이터를 생성하려고 하고, 판별자는 생성자가 만든 데이터가 진짜인지 가짜인지 판별합니다. 위의 두 가지 네트워크가 서로 경쟁하여 학습합니다. GAN이 주로 활용되는 사례는 AI 스피커에서 사용자의 음성이 실제로 사용자의 음성인지 아닌지를 판별하는 서비스가 있습니다.

실제 서비스에 사용되는 다양한 신경망 소개

3. 자율 주행

카메라를 통해 입력된 이미지에서 자율주행에 필요한 정적인 환경정보를 수집하고, 구별합니다. 이때 CNN 기반의 딥러닝 모형이 활용됩니다. 정적인 정보들은 차선, 표지판, 교통신호, 운전가능도로 등이 해당됩니다.

또한, 동적인 정보도 수집합니다. 주변 보행자, 차량 등등 이죠. 이러한 것들을 검출하기 위해 매우 빠르고 정확한 CNN 모형이 필요합니다. 카메라 정보 뿐 아니라 레이더, LiDAR, 각종 센서들이 감각 기관의 역할을 수행합니다.

카메라를 포함한 각종 감각정보를 통해 무인 차량은 자신의 움직임을 결정해야 합니다.

이때는 주로 강화학습(Reinforcement Learning)을 활용 할 수 있습니다.

강화학습은 특정 상태(State)에서 정의된 차량(Agent)이 현재 환경을 인식하여 선택가능한 보상(Reward)을 최대화 하는 행동을 결정하는 것을 말합니다.

이는, 카메라 또는 센서로 인식된 환경 하에서 최적의 행동을 찾도록 합니다.

예를 들어, 막히는 도로를 피해야 하는 상황, 교차로로 뛰어는 보행자를 피하는 것, 차선 변경하는 차량이 있으면 속도를 줄이거나 멈추는 행동등.. 특정 상황에서 받은 보상을 학습합니다.

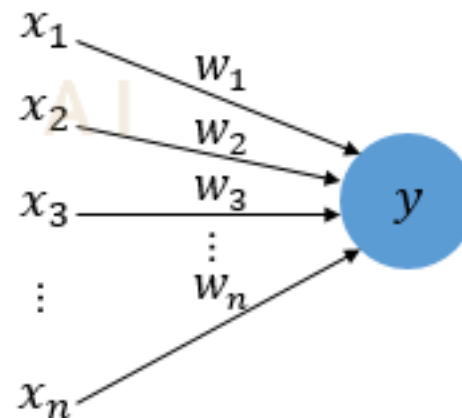
계속 운전을 시도하면서 보상을 받고, 이를 통해 스스로 운전하는 방법을 배웁니다.

인공 지능 신경망 개념

- 퍼셉트론(Perceptron)
- 1957년 프랭크 로젠블라트(Frank Rosenblatt)가 제안한 알고리즘
- 다수의 입력값을 연산 처리 하여 하나의 값을 반환함
- 인공신경망의 초기 형태로 평가됨



신경 세포 뉴런



퍼셉트론(Perceptron)

인공 지능 신경망 개념

- 퍼셉트론(Perceptron) 동작 방식

각 노드의 입력치 x_n , 가중치 w_n 를 곱한 값을 모두 합함
합한 값을 활성화함수로 판단하여 임계치 이상인 경우 1, 그렇지 않으면 0 출력

$$output = \begin{cases} 0 & (w_1x_1 + w_2x_2 + \dots + w_nx_n \leq \theta) \\ 1 & (w_1x_1 + w_2x_2 + \dots + w_nx_n > \theta) \end{cases}$$

- 가중치의 크기는 일반적으로 입력값의 중요도를 나타냄

즉 해당하는 특징값과 가중치를 곱한 결과를 모두 더하여 output를 구하고 activation function 을 적용하여 최종 출력 $y = +1$ or -1 이 됩니다.

인공 지능 신경망 개념

- 단층 퍼셉트론 이용하여 AND NAND OR 게이트 구현 실습

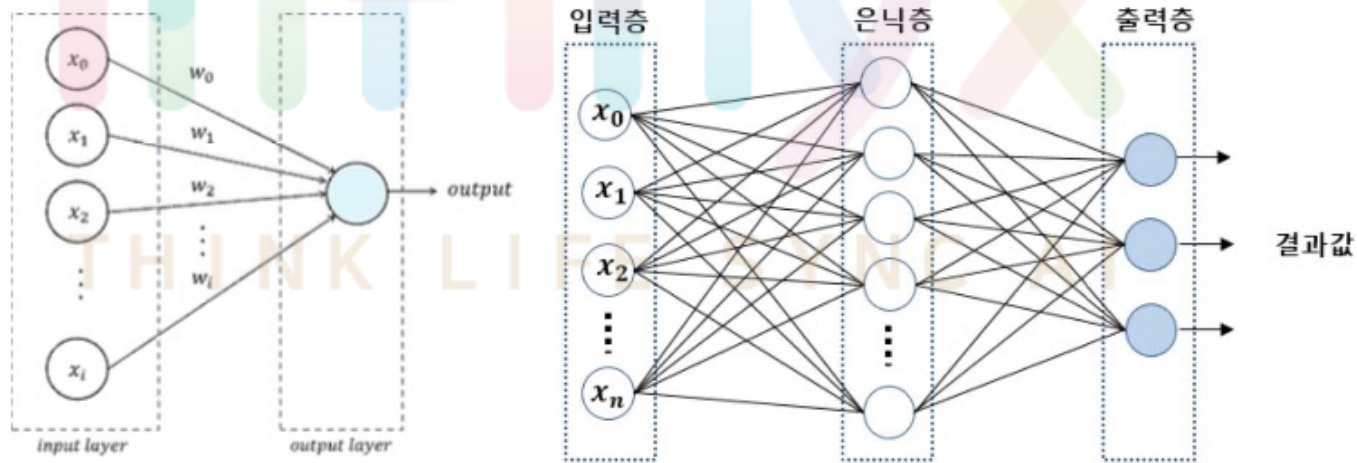


인공 지능 신경망 개념

- 다층 퍼셉트론(Multi-Layer Perceptron : MLP)

단층 퍼셉트론 : 데이터 입력 층(input layer) 와 출력층(output layer) 만 존재

다층 퍼셉트론 : 입력층과 출력층 사이에 하나 이상의 은닉층(hidden layer) 존재



[단층 퍼셉트론과 다층 퍼셉트론]

복잡한 문제를 해결하기 위해 다층 퍼셉트론 내 은닉층의 개수가 많아짐(깊어짐)에 따라, 이를 가리키기 위해 심층 신경망(Deep Neural Network: DNN) 용어 사용

- 심층 신경망 : 입력층과 출력층 사이 여러개의 은닉층으로 구성된 인공신경망(ANN)

인공 지능 신경망 개념

- 다층 퍼셉트론을 이용한 XOR_gate 구현 실습
- 파이토치로 다층 퍼셉트론 구현하기 실습
Background : Linear, Sigmoid, optimizer(의미, 종류 대표적인 SGD, adam) 모델 정의, Cuda, Loss function
- 사이킷런 패키지에서 제공하는 분류용 예제 데이터를 이용한 다층 퍼셉트론으로 손 글씨 분류 하기 실습
Background : sklearn 라이브러리, zip(python), 모델 정의, Loss function

THINK LIFE SYNC AI

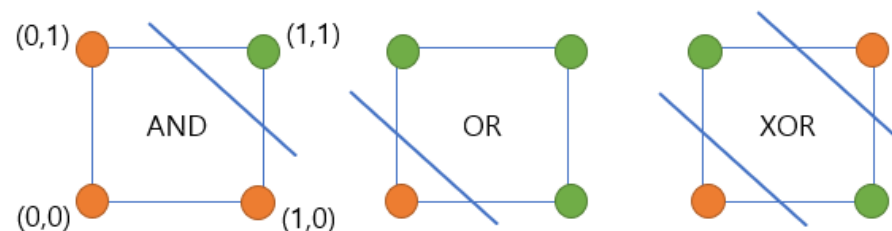
인공 지능 신경망 개념

- 선형 분류자의 한계점

직선형 영역만 표현할 수 있다. 데이터들의 차이가 확연히 두드러지지 않는다면 두 가지의 종류조차도 분류하지 못할 수도 있다

단층 퍼셉트론에서 논리함수를 예를 들어보면 AND, NAND, OR 논리회로를 계산할 수 있지만 XOR(배타적 논리합)의 경우 단일 선형 프로세스로 데이터를 분류해낼 수 없다. 여러 개의 선형 분류자를 이용해서 데이터를 나누거나 한 선으로 곡선으로 나눌 수 있기 때문에 비선형 분류이다.

입력 A	입력 B	논리곱 (AND)	논리합 (OR)	배타적 논리합 (XOR)
0	0	0	0	0
0	1	0	1	1
1	0	0	1	1
1	1	1	1	0





감사합니다.

THINK LIFE SYNC AI

Infinyx®