



셋째마당

딥러닝의 시작, 신경망

8장 다층 퍼셉트론

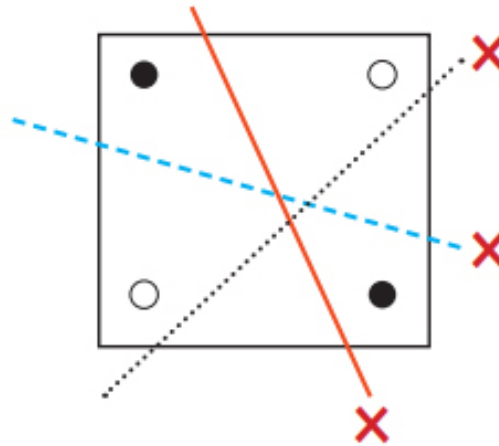
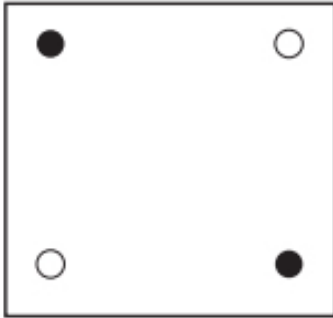
- 1 다층 퍼셉트론의 등장
- 2 다층 퍼셉트론의 설계
- 3 XOR 문제의 해결
- 4 코딩으로 XOR 문제 해결하기



1 다층 퍼셉트론의 등장

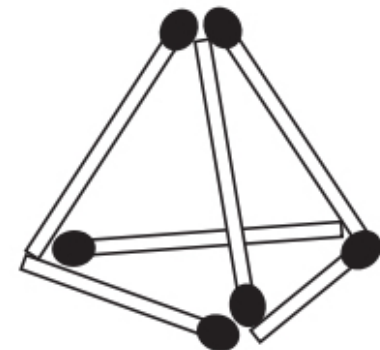
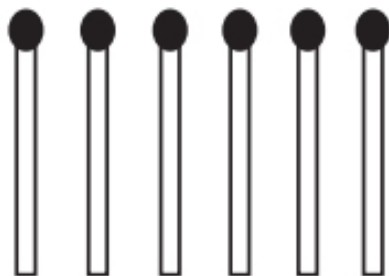


1 다층 퍼셉트론의 등장



● 다층 퍼셉트론의 등장

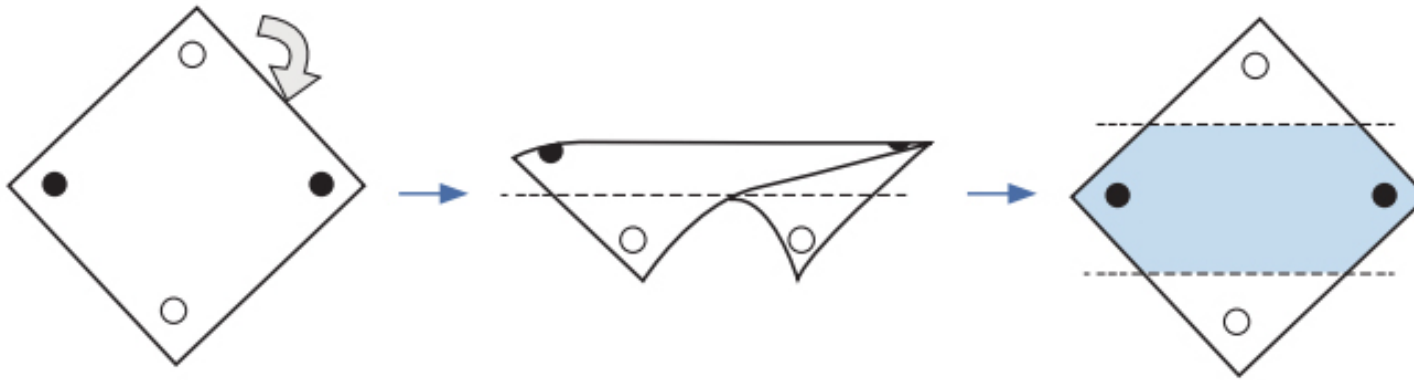
- 앞서 종이 위에 각각 엇갈려 놓인 검은색 점 두 개와 흰색 점 두 개를 하나의 선으로는 구별할 수 없다는 것을 살펴보았습니다.
- 언뜻 보기에 해답이 없어 보이는 이 문제를 해결하려면 새로운 접근이 필요하다는 것입니다.
- 성냥개비 여섯 개로 정삼각형 네 개를 만들라는 문제가 있을 때
- 2차원 평면에서만 해결하려는 고정 관념을 깨고 피라미드 모양으로 성냥개비를 쌓아 올리니 해결할 수 있듯이





1 다층 퍼셉트론의 등장

▼ 그림 8-3 | XOR 문제의 해결은 평면을 휘어 주는 것!

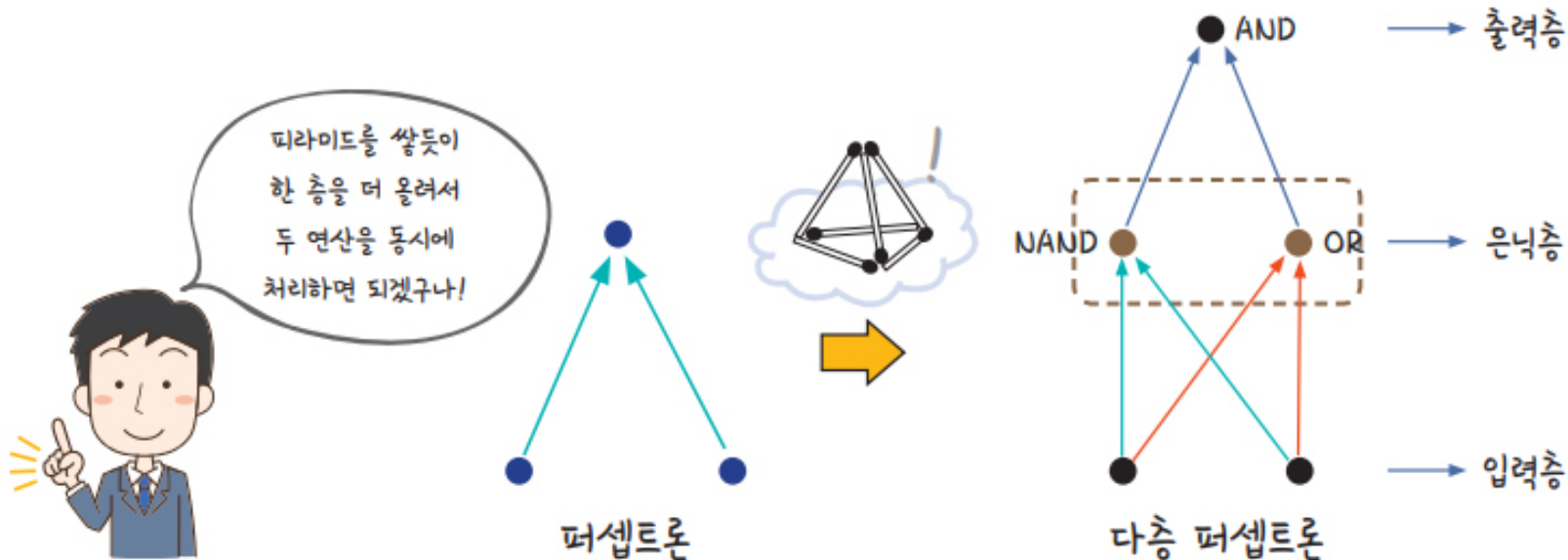


● 다층 퍼셉트론의 등장

- 앞에 예시처럼 주어진 구조를 변경
- 즉 아래 그림 처럼 평면을 휘게 되면 두개의 선이 만들어짐
- XOR 문제에 적용하면 '퍼셉트론 두 개를 한 번에 계산'하면 된다는 결론에 이름
- 그래서 퍼셉트론 두 개를 각각 처리하는 은닉층(hidden layer)이라는 것을 만들게 되었다.



1 다층 퍼셉트론의 등장



● 다층 퍼셉트론의 등장

- 파란색 라인의 퍼셉트론이 있고 빨간색 라인의 퍼셉트론이 있다.
- 이 두개의 퍼셉트론 결과에 의해 또 하나의 퍼셉트론이 가동을 하게 된다.
- 은닉층을 두므로 두개의 퍼셉트론이 실행되게 되었고 XOR문제를 해결할 수 있게 된 것이다.



1 다층 퍼셉트론의 등장

▼ 그림 8-4 | 은닉층이 XOR 문제를 해결하는 원리

XOR (배타적 논리합) 하나만 1이어야 1			2 NAND (부정 논리곱) 하나라도 0이면 1			3 OR (논리합) 두 개 중 한 개라도 1이면 1			4 AND (논리곱) 두 개 모두 1일 때 1		
x_1	x_2	결괏값	x_1	x_2	결괏값 1	x_1	x_2	결괏값 2	결괏값 1	결괏값 2	결괏값
0	0	0	0	0	1	0	0	0	1	0	0
0	1	1	0	1	1	0	1	1	1	1	1
1	0	1	1	0	1	1	0	1	1	1	1
1	1	0	1	1	0	1	1	1	0	1	0

● 다층 퍼셉트론의 등장

- 먼저 ① x_1 과 x_2 를 두 연산으로 각각 보냄
- ② 첫 번째 연산에서는 NAND 처리를 함
- ③ 이와 동시에 두 번째 연산에서 OR 처리를 함
- ②와 ③을 통해 구한 결괏값 1과 결괏값 2를 가지고 ④ AND 처리를 하면 우리가 구하고자 하는 출력 값을 만들 수 있음



2 다층 퍼셉트론의 설계

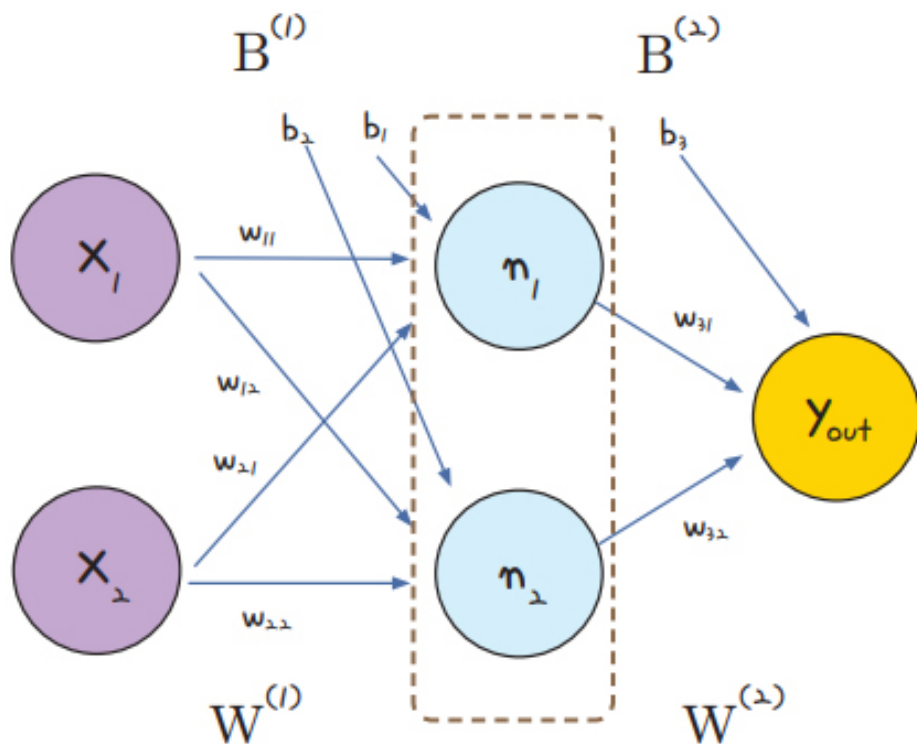


2 다층 퍼셉트론의 설계

● 다층 퍼셉트론의 설계

- 다층 퍼셉트론이 입력층과 출력층 사이에 숨어 있는 은닉층을 만드는 것을 그림으로 나타내면 그림 8-5와 같음

▼ 그림 8-5 | 다층 퍼셉트론의 내부



$$W^{(1)} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \quad B^{(1)} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$W^{(2)} = \begin{bmatrix} w_{31} \\ w_{32} \end{bmatrix} \quad B^{(2)} = [b_3]$$

- 은닉층을 포함해 가중치 여섯 개와 바이어스 세 개가 필요함

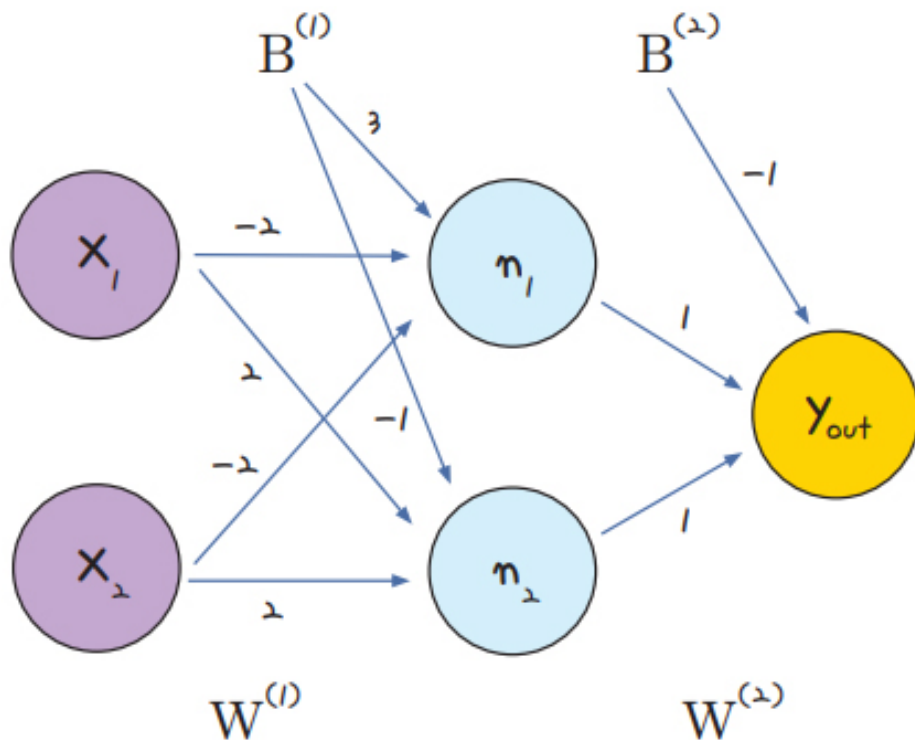


3 XOR 문제의 해결

● XOR 문제의 해결

- 최소제곱법에 의해서 만들어진 갯봉치를 대입하면 다음 그림과 같을 것이다.

▼ 그림 8-6 | 다층 퍼셉트론의 내부에 변수 채우기



$$W^{(1)} = \begin{bmatrix} -2 & 2 \\ -2 & 2 \end{bmatrix} \quad B^{(1)} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

$$W^{(2)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad B^{(2)} = [-1]$$

● XOR 문제의 해결

- 먼저 알고 있는 가중치 값과 비이어스를 이용해보면



3 XOR 문제의 해결

● XOR 문제의 해결

- 시그모이드를 이용하게 되면 0에 가까우면 0, 1에 가까우면 1이 됨
- 그림과 같을 때 계산을 해보면

▼ 표 8-1 | XOR 다층 문제 해결

x_1	x_2	n_1	n_2	y_{out}	우리가 원하는 값
0	0	$\sigma(0 * (-2) + 0 * (-2) + 3) \approx 1$	$\sigma(0 * 2 + 0 * 2 - 1) \approx 0$	$\sigma(1 * 1 + 0 * 1 - 1) \approx 0$	0
0	1	$\sigma(0 * (-2) + 1 * (-2) + 3) \approx 1$	$\sigma(0 * 2 + 1 * 2 - 1) \approx 1$	$\sigma(1 * 1 + 1 * 1 - 1) \approx 1$	1
1	0	$\sigma(1 * (-2) + 0 * (-2) + 3) \approx 1$	$\sigma(1 * 2 + 0 * 2 - 1) \approx 1$	$\sigma(1 * 1 + 1 * 1 - 1) \approx 1$	1
1	1	$\sigma(1 * (-2) + 1 * (-2) + 3) \approx 0$	$\sigma(1 * 2 + 1 * 2 - 1) \approx 1$	$\sigma(0 * 1 + 1 * 1 - 1) \approx 0$	0

- \approx 기호는 '거의 같다'를 의미