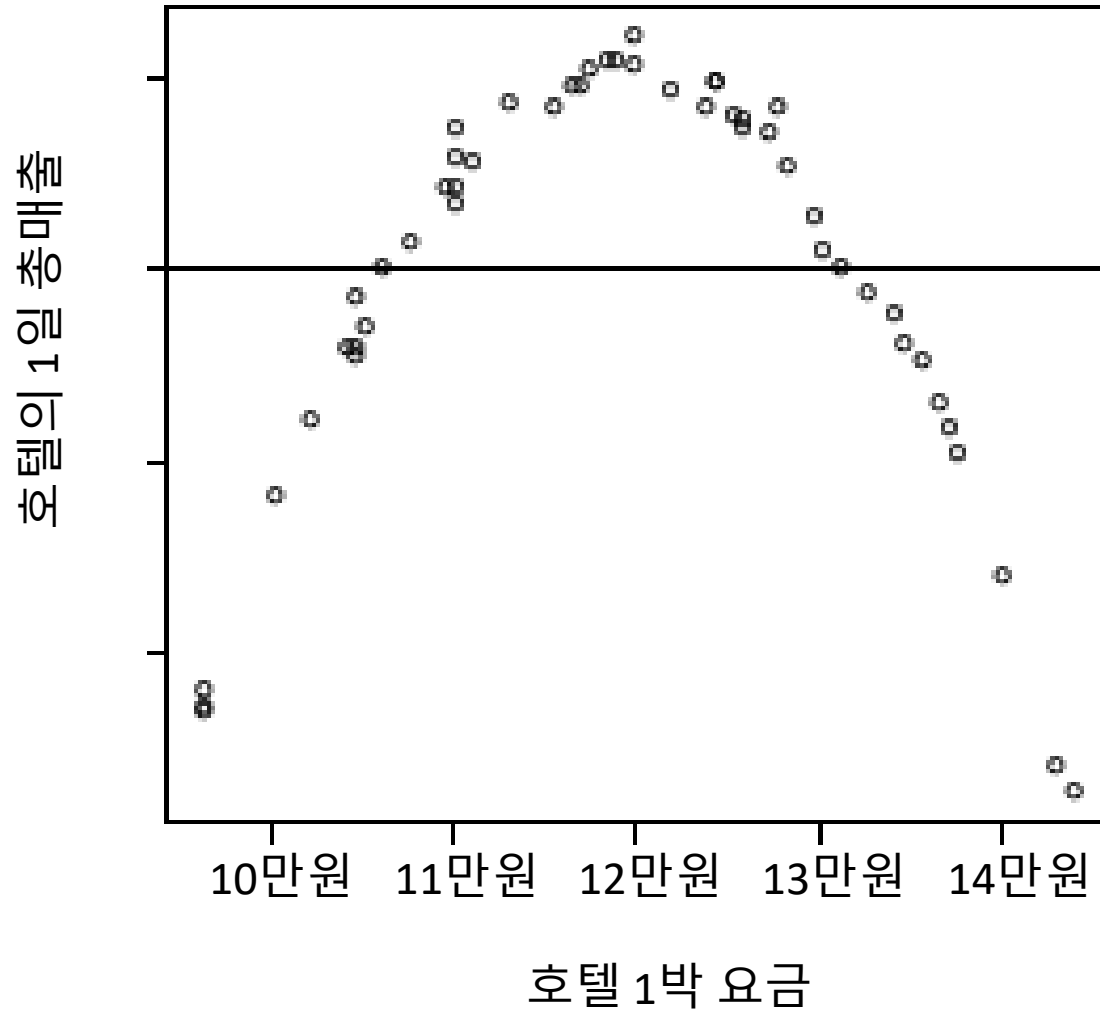


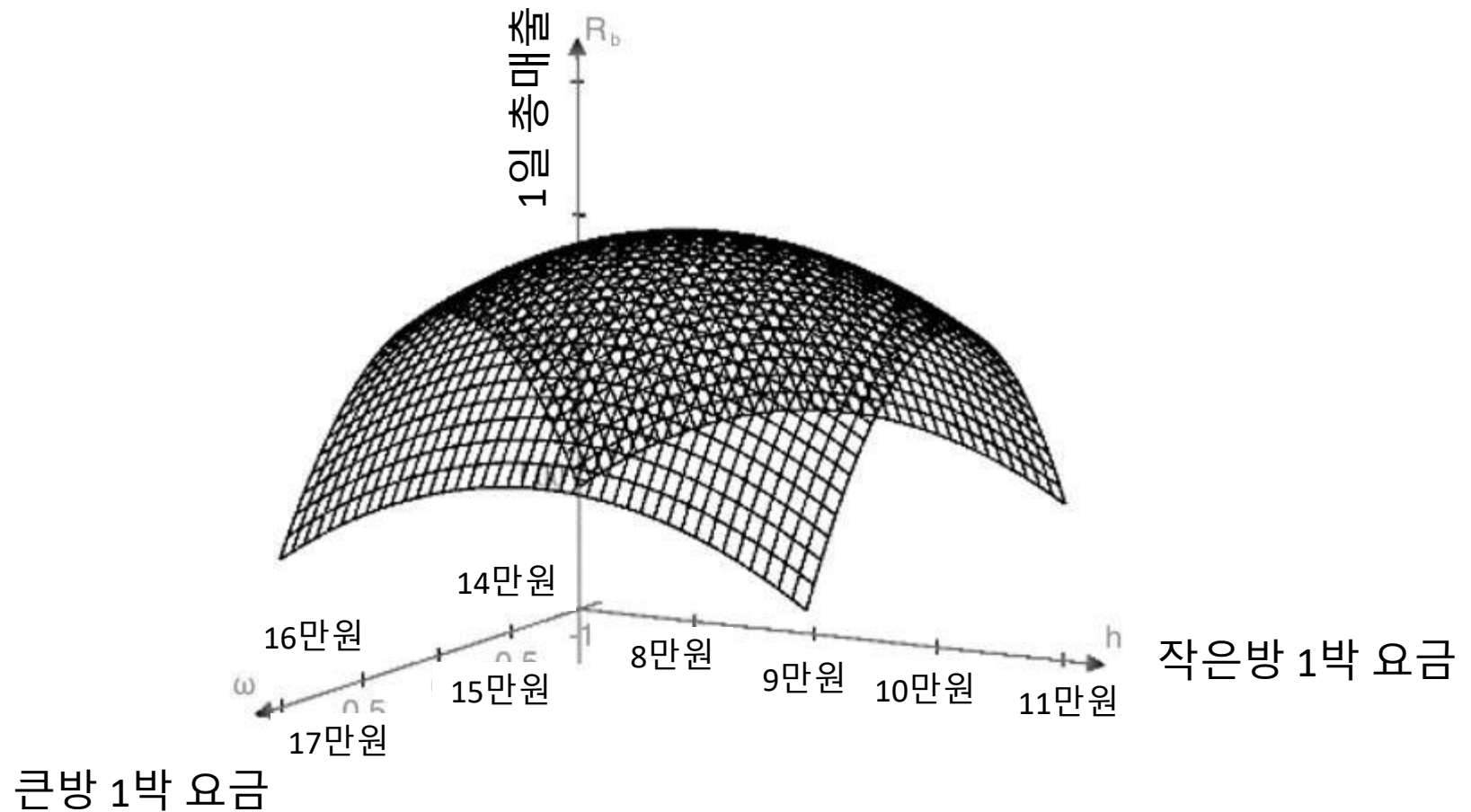
MLP(Multi Layer Perceptron) 완전정복

이택민, Ph.D

호텔 총매출 모델 (1차원)



호텔 총매출 모델 (2차원, 2 - 1 차원)



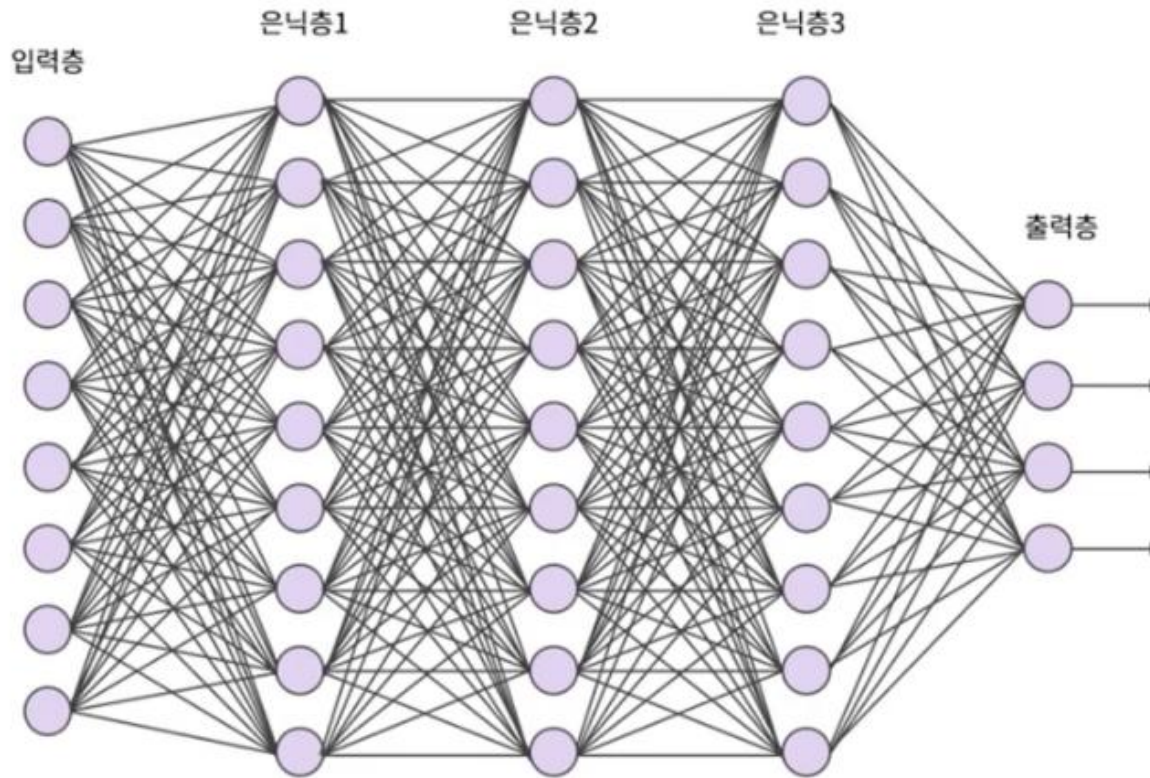
호텔 총매출 모델 (14 x 3)



날짜	디럭스 더블	패밀리 스위트	더앤스 위트	스파온 온천	온천수 영장	조식 뷔페	더나이트 메뉴	카바 나	날씨	주말	휴일	쏟비 치	브리 드	총매 출	수영장 매출	더 나이 트 매출
2022년 06 월 01일	10	15	20	2	6	2.5	7	20	3	0	1	25	22	3100	930	837
2022년 06 월 02일	10.1	15.1	20.1	2	6.1	7.1	7	20	2	0	0	15	12	3400	1020	918
2022년 06 월 03일	10.2	15.2	20.2	2	6.2	7.2	7	20	1	0	0	15	12	3700	1110	999
2022년 06 월 04일	10.3	15.3	20.3	2	6.3	7.3	7	20	0	1	0	25	22	8000	2400	2160
2022년 06 월 05일	10.4	15.4	20.4	2	6.4	7.4	7	20	0	1	0	25	22	8000	2400	2160
2022년 06 월 06일	10.5	15.5	20.5	2	6.5	7.5	7	20	0	0	1	25	22	4000	1200	1080
2022년 06 월 07일	10.6	15.6	20.6	2	6.6	7.6	7	20	0	0	0	20	17	4000	1200	1080
2022년 06 월 08일	10.7	15.7	20.7	2	6.7	7.7	7	20	0	0	0	15	12	4000	1200	1080

MLP (Multi-Layer Perceptron)구조

- ✓ DNN 의 가장 간단한 종류의 하나
- ✓ Tabular 데이터 입력으로 Tabular 데이터 출력을 예측



Terminology

- ✓ Layer
- ✓ Node
- ✓ Neuron
- ✓ Weight
- ✓ Bias
- ✓ Activation Function
- ✓ Nonlinear Function
- ✓ Forward Propagation
- ✓ Backpropagation
- ✓ Gradient Descent
- ✓ Epoch
- ✓ Batch
- ✓ Loss Function

MLP 구조 설계 (Input - Output 선정)

X



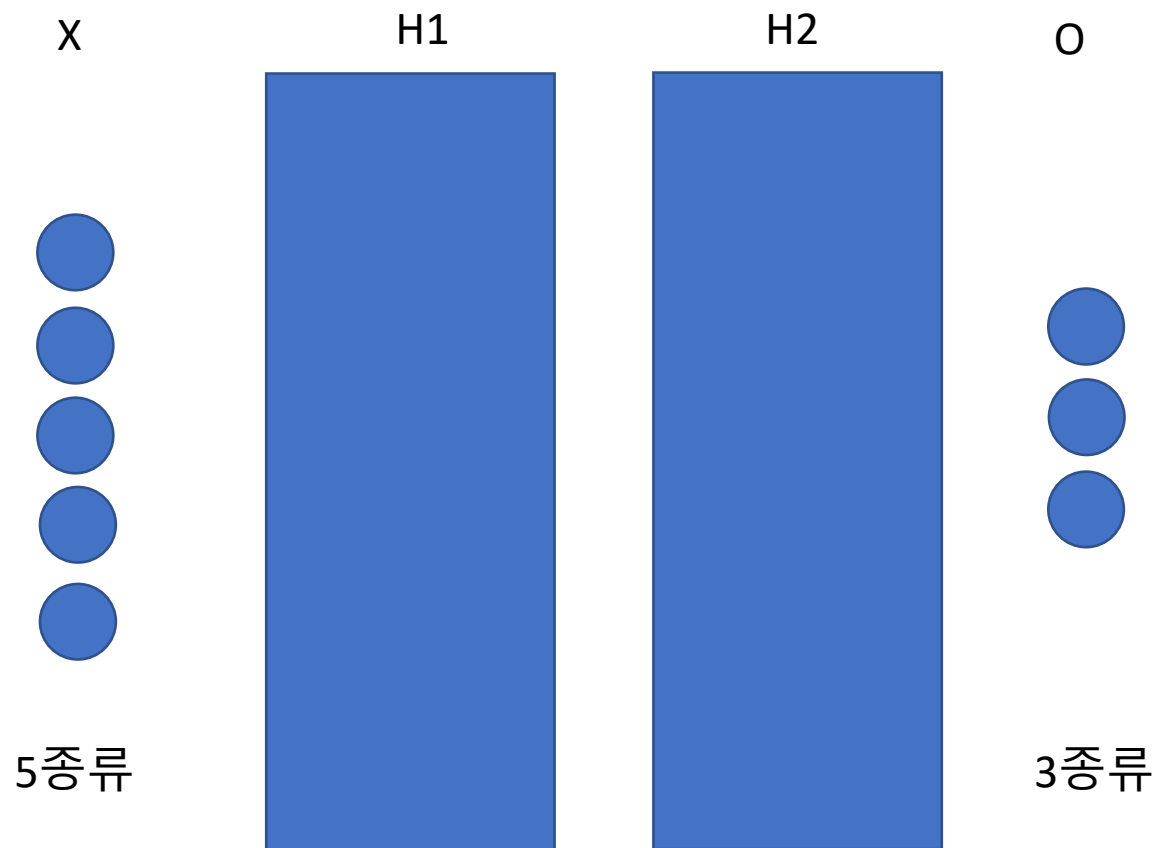
5종류

O

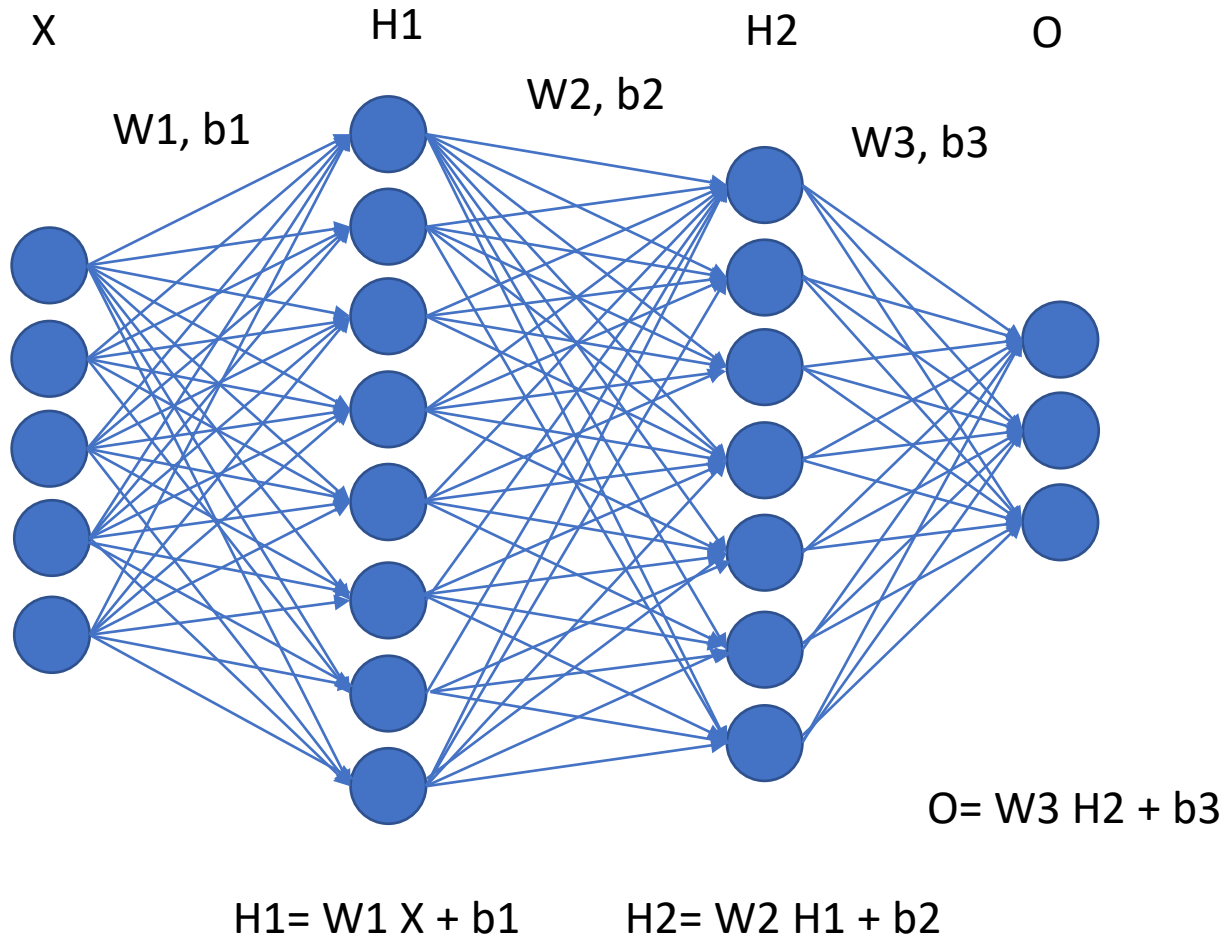


3종류

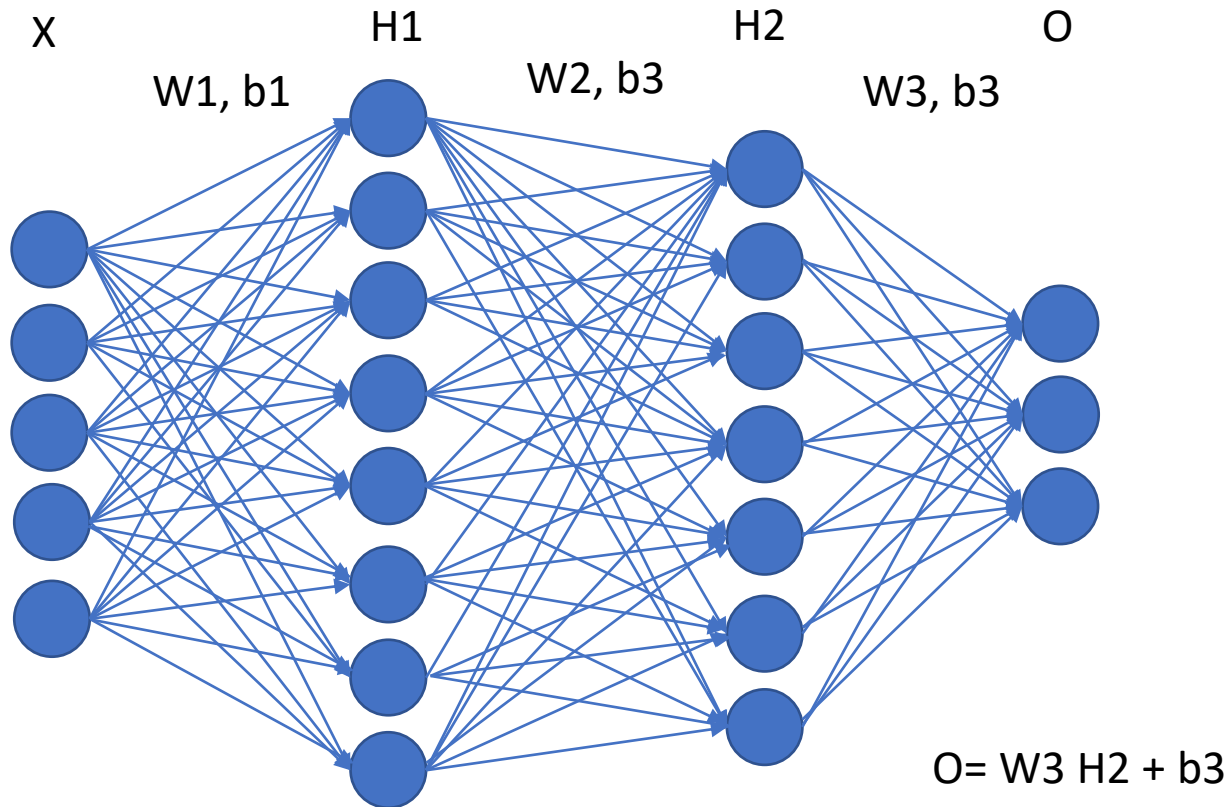
MLP 구조 설계 (Hidden Layer 개수 선정)



MLP 구조 설계 (Node 개수 선정)



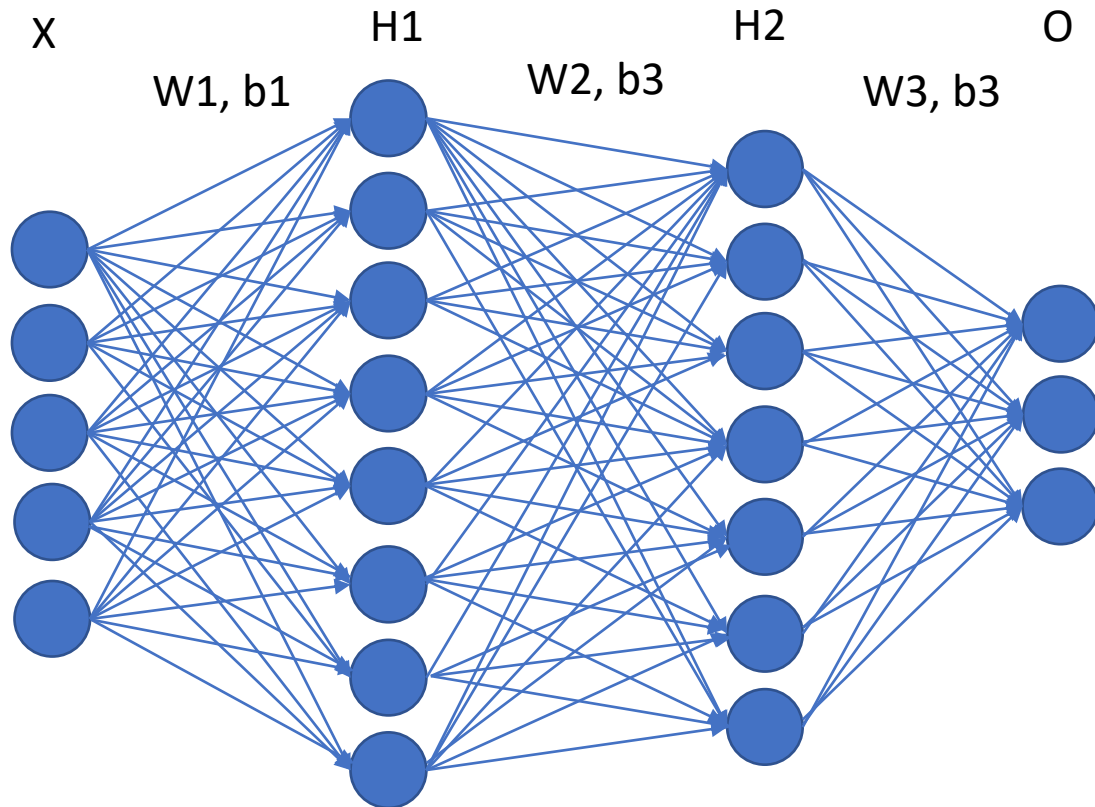
MLP 구조 설계 (Nonlinear 함수 선정 및 은닉층합성)



$$H1 = \sigma(W1 X + b1)$$

$$H2 = \sigma(W2 H1 + b2)$$

MLP 구조 설계 (Nonlinear 함수 선정 및 은닉층합성)

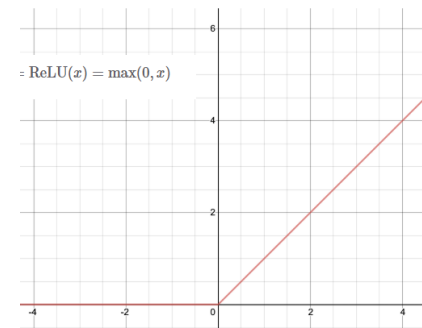


$$H1 = \sigma(W1 X + b1)$$

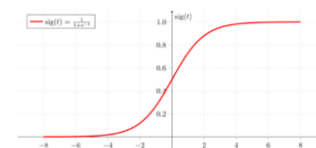
$$H2 = \sigma(W2 H1 + b2)$$

$\sigma(X)$

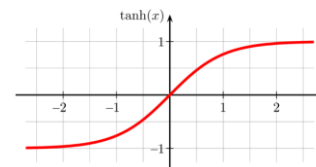
Relu 함수 ($x > 0$ 이면 x , 그외 0)



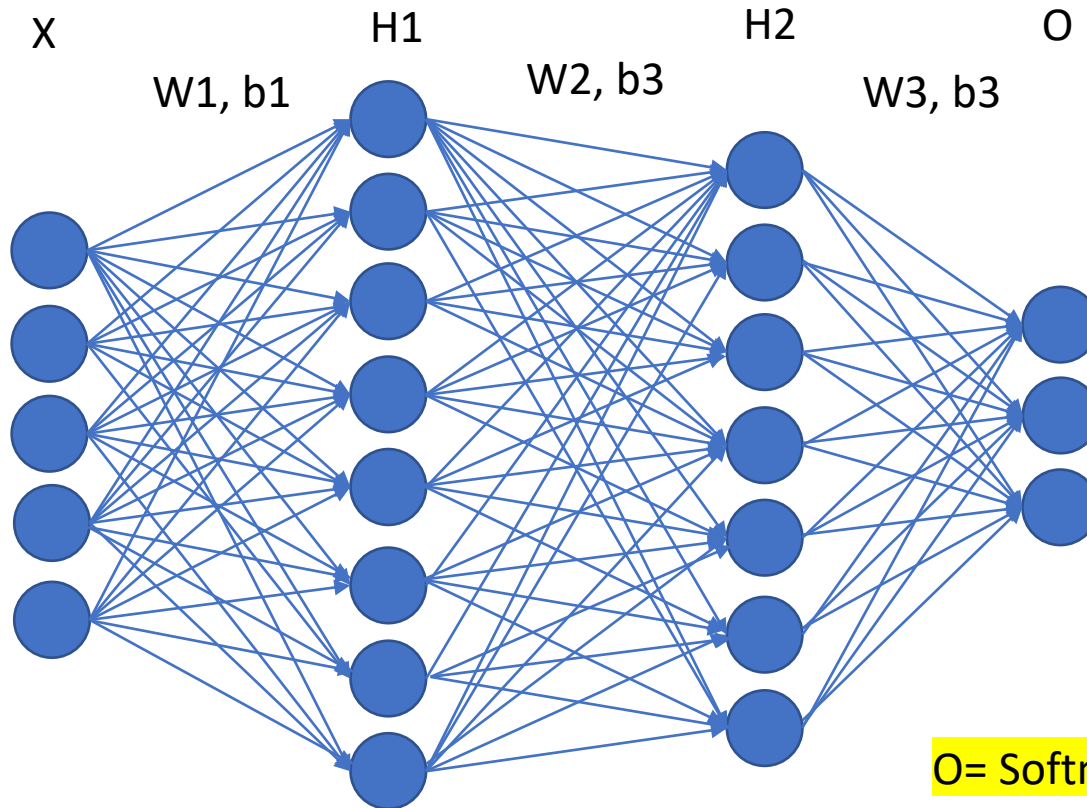
Sigmoid 함수 $f(x) = \frac{1}{1 + e^{-x}}$



Tanh 함수 $f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$



MLP 구조 설계 (출력층에 Activation Function 합성)



$$H1 = \sigma(W1 X + b1)$$

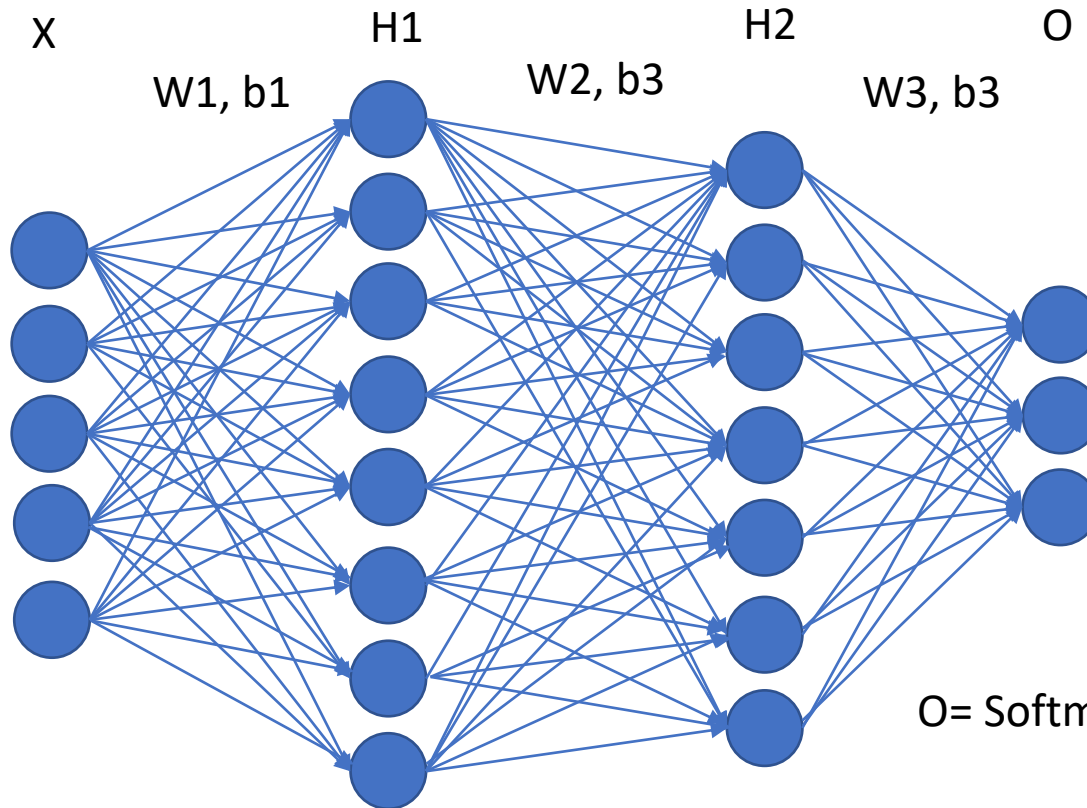
$$H2 = \sigma(W2 H1 + b2)$$

$$O = \text{Softmax}(W3 H2 (\sigma(W2 (\sigma(W1 X + b1)) + b2) + b3))$$

MLP 구조 설계 (Train – Validation- Test data set 나누기)

70:15:15

80:20



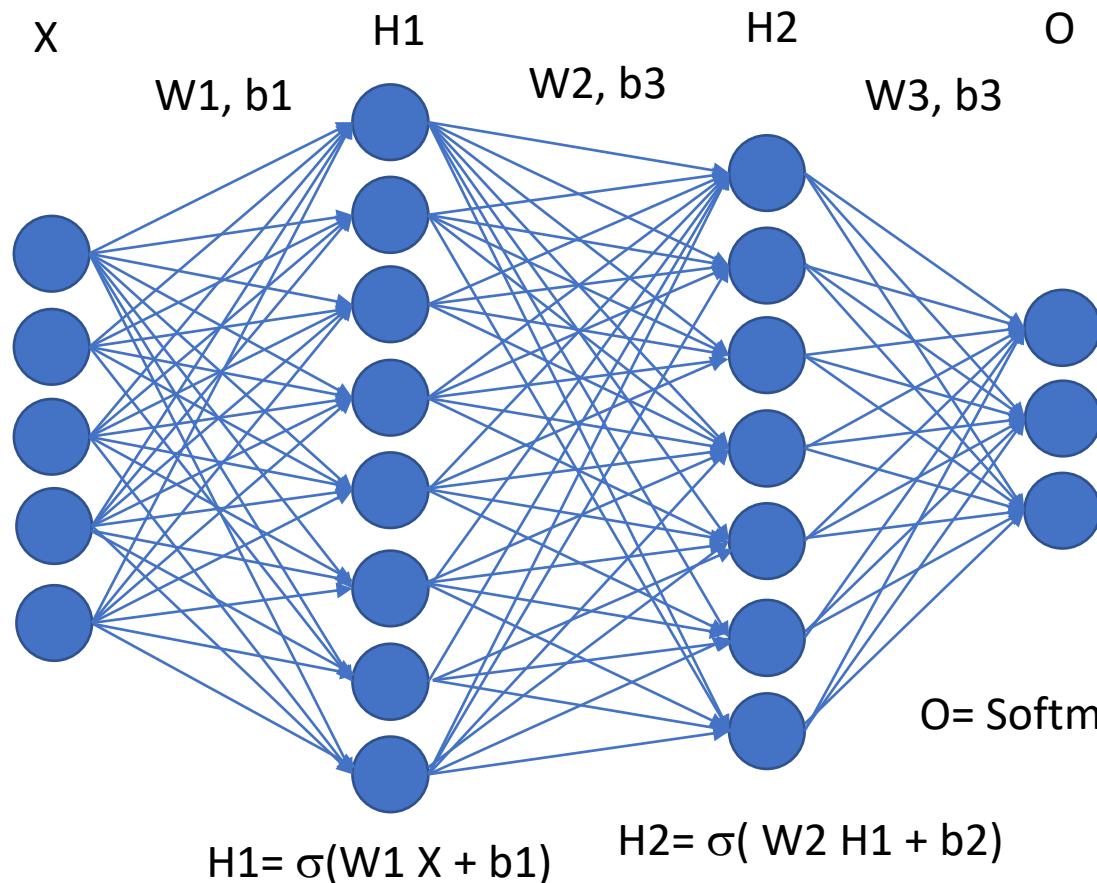
$$O = \text{Softmax}(W3 H2 + b3)$$

$$H1 = \sigma(W1 X + b1) \quad H2 = \sigma(W2 H1 + b2)$$

MLP 구조 설계 (Train data : Forward Propagation)

현재의 W, b 를 이용해 배치 데이터를 입력하여 모델의 출력 \hat{y} 를 계산한다.

$$z = Wx + b, \quad \hat{y} = f(z)$$



Loss Function 으로 손실 계산

Mean Squared Error (MSE)

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

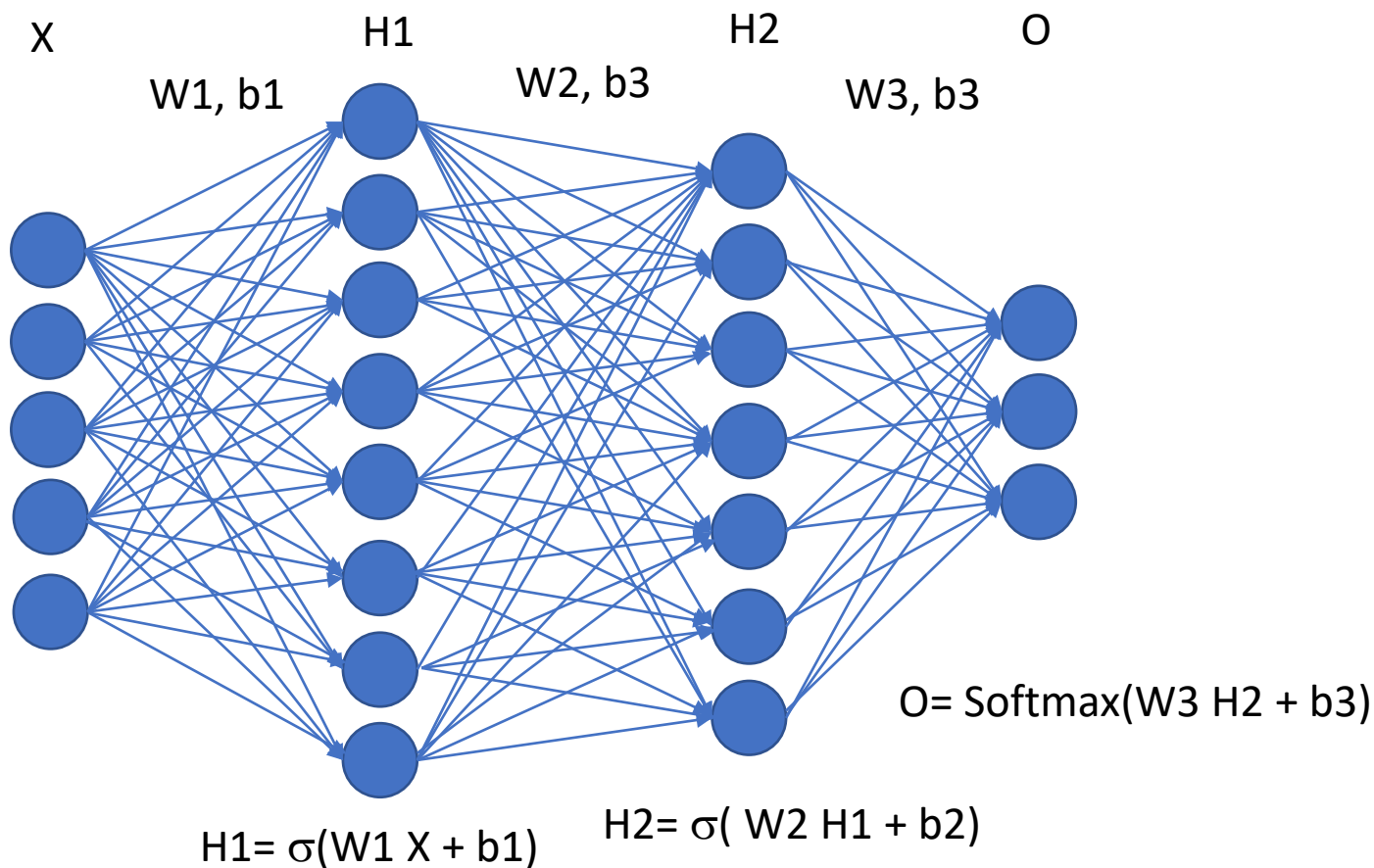
오차의 제곱 평균

$$O = \text{Softmax}(W3 H2 + b3)$$

출력 \hat{y} 와 정답 y 사이의 손실 $L(\hat{y}, y)$ 을 계산한다.

MLP 구조 설계 (Train data : Backpropagation)

손실 함수의 결과를 바탕으로 $\frac{\partial L}{\partial W}$, $\frac{\partial L}{\partial b}$ 를 계산한다.



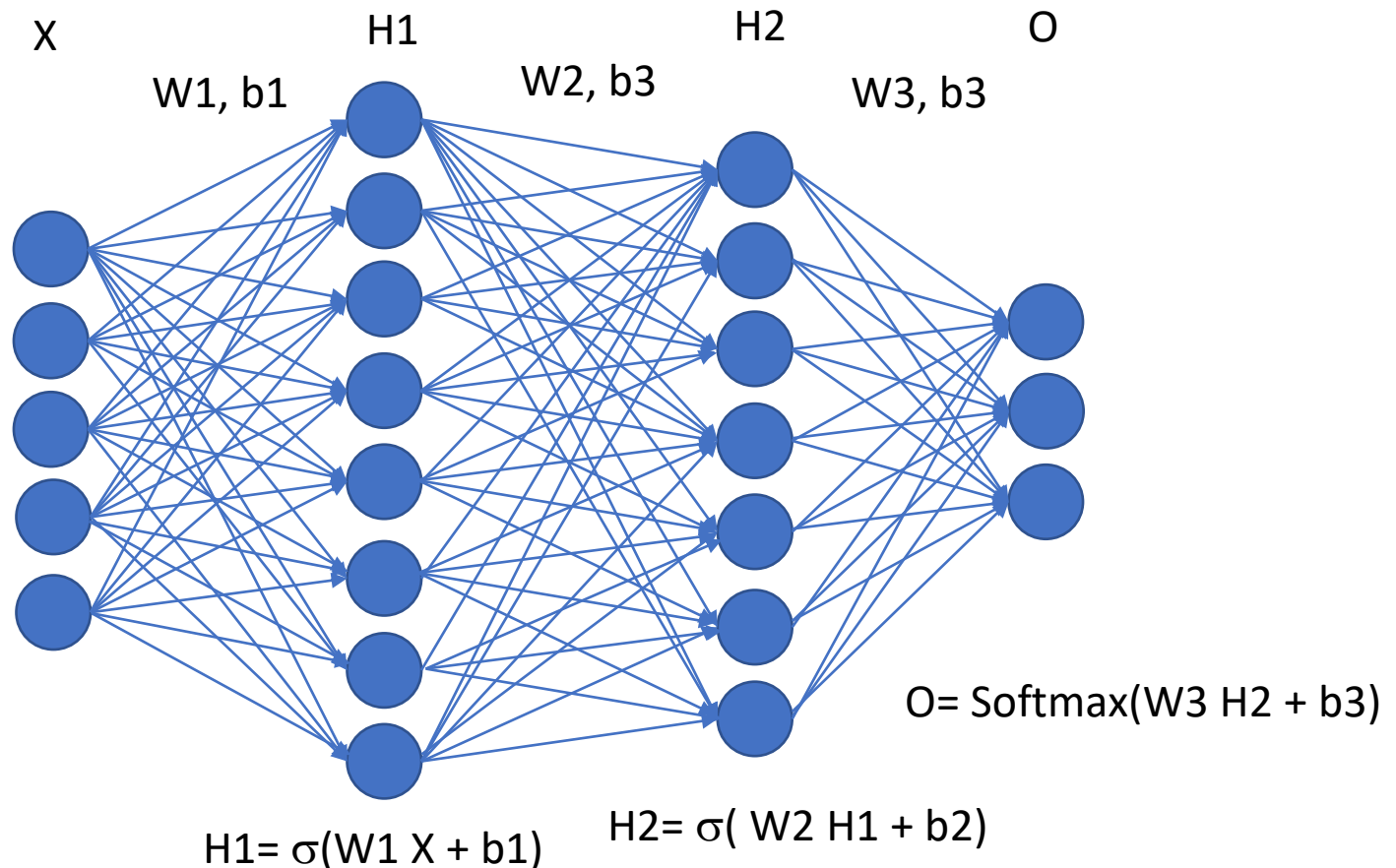
MLP 구조 설계 (Train data : Backpropagation)

옵티마이저가 기울기를 바탕으로 W, b 를 업데이트한다: $W := W - \eta \cdot \frac{\partial L}{\partial W}$

여기서 η 는 학습률(learning rate)이다.

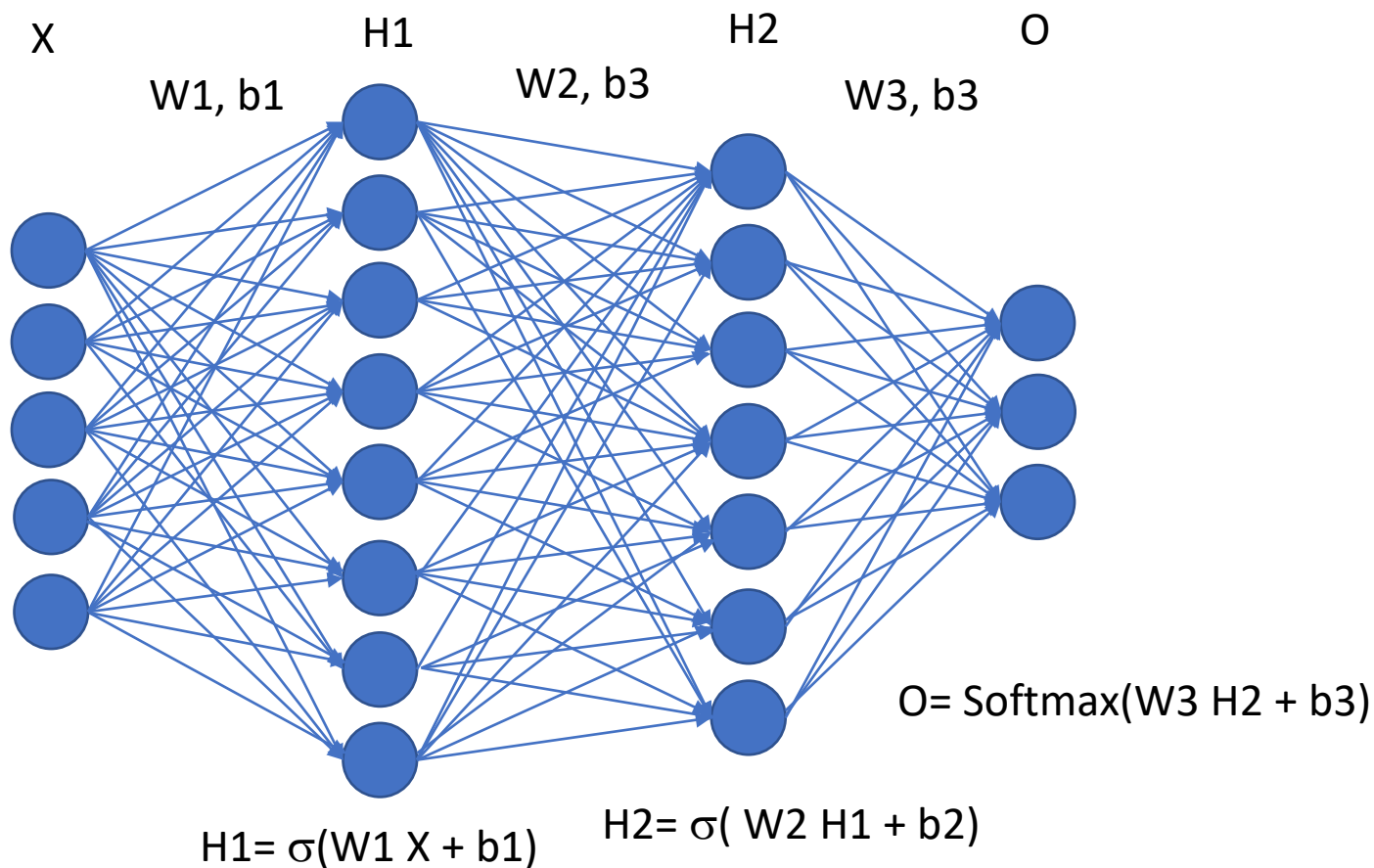
$$b := b - \eta \cdot \frac{\partial L}{\partial b}$$

초기 학습률은 설정해 주고, 다음부터는 Optimizer가 조절한다.



MLP 구조 설계 (Validation data : Loss)

1 epoch이 끝날 때마다 validation 데이터를 이용해 손실함수로 모델의 성능을 평가한다.



감사합니다.