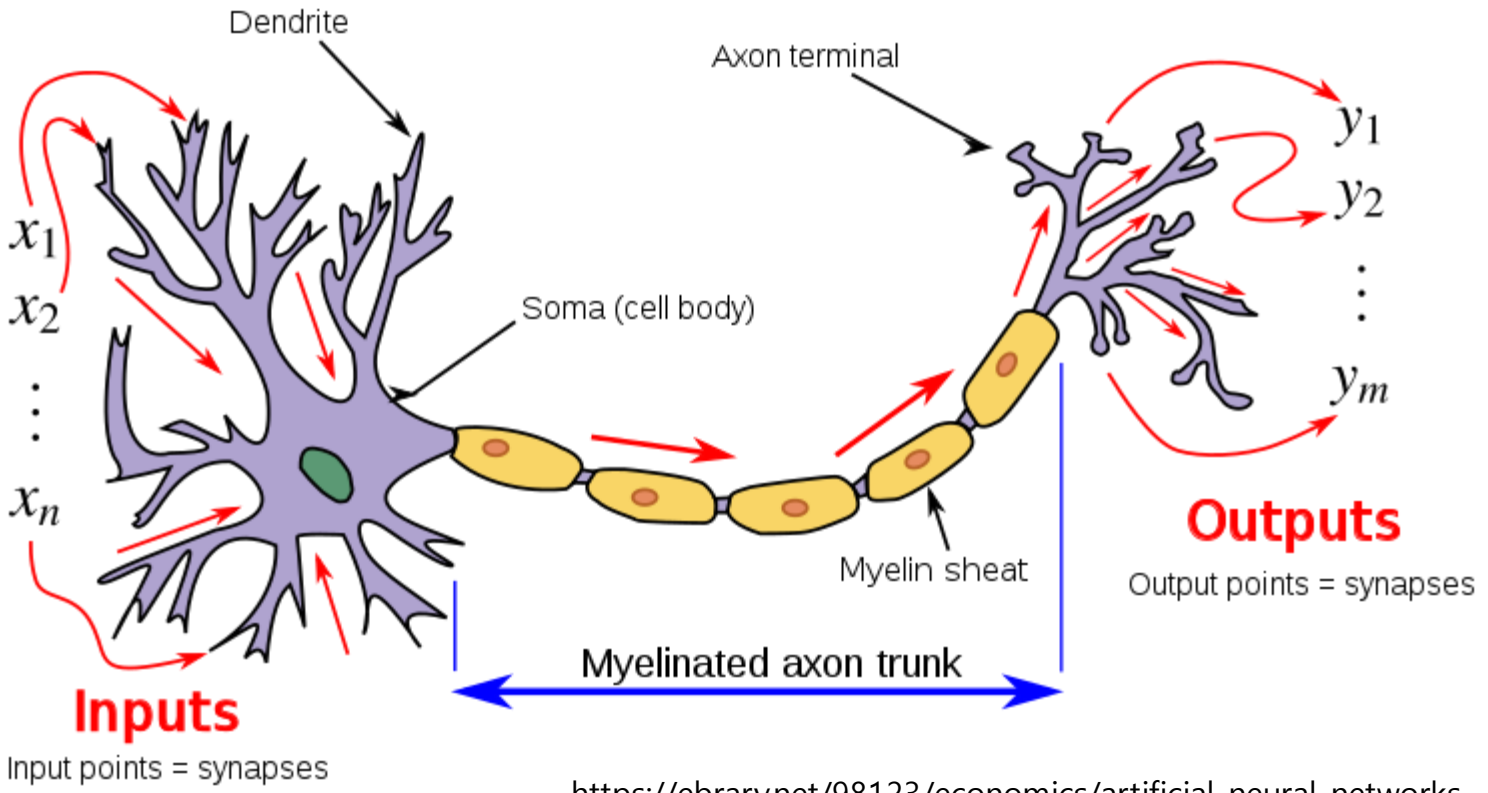
Neural Network

Neural Network

Prof. Jaewook Lee

Blockchain @ SNU



https://ebrary.net/98123/economics/artificial_neural_networks

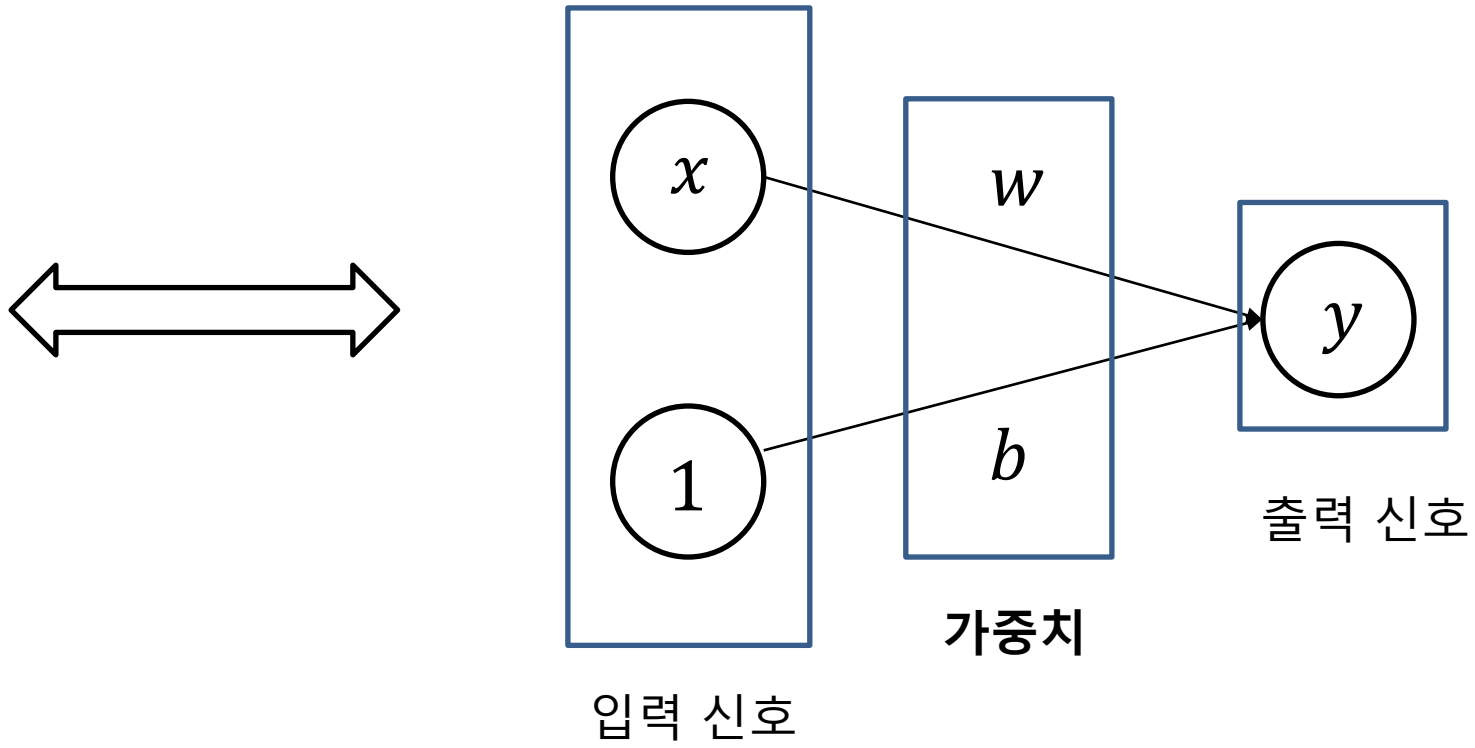
NEURAL NETWORK

선형회귀와 신경망

Source: https://blockchain.berkeley.edu/decal/

- 퍼셉트론
  - 다수의 신호를 입력 받아 하나의 신호를 출력
  - 선형회귀는 하나의 퍼셉트론으로 구현 가능

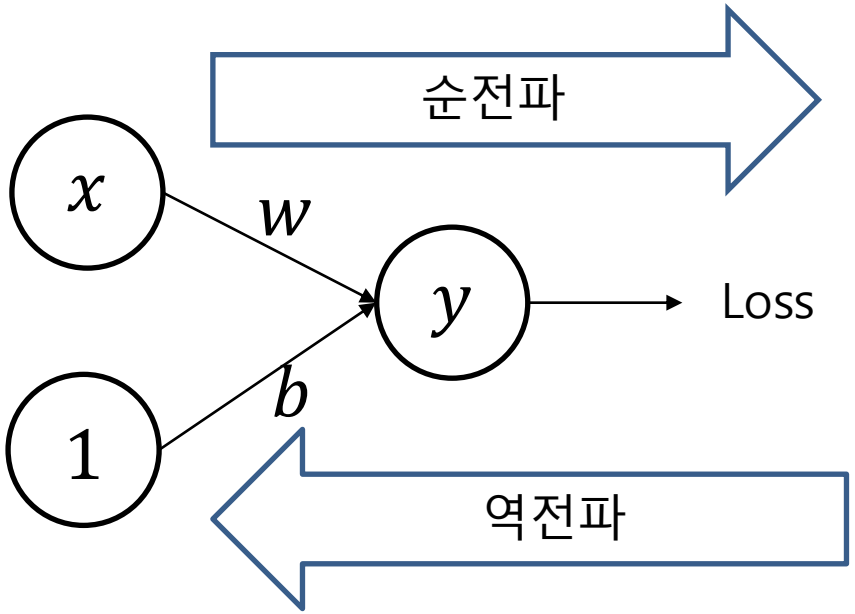
$$wx + b = y$$



역전파와 순전파

목표: 역전파를 통해 좋은 가중치를 찾자!

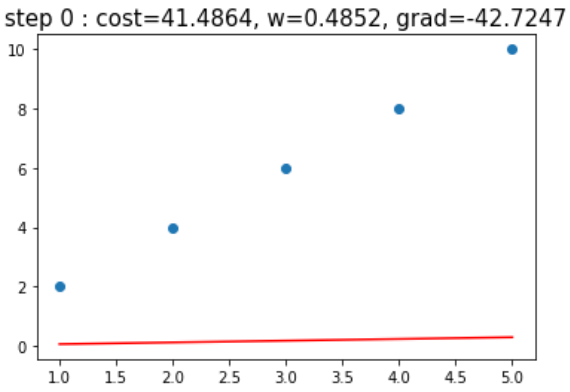
- 순전파
  - 데이터처리 → 모델 구현 → 예측값 도출 → 손실함수 계산
- 역전파
  - 기울기계산 → 개선방향 구하기 → **가중치 개선**

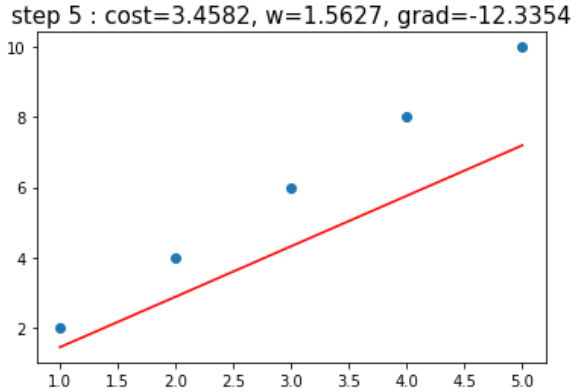


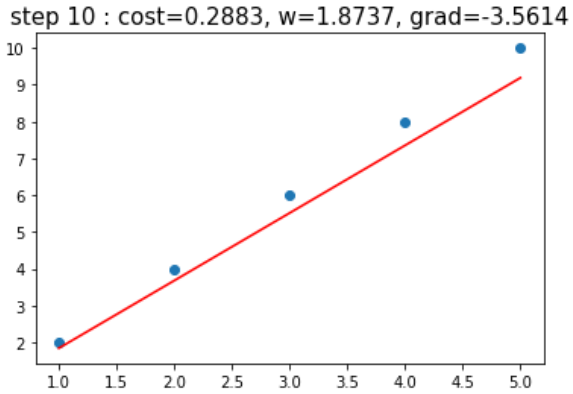
역전파와 순전파

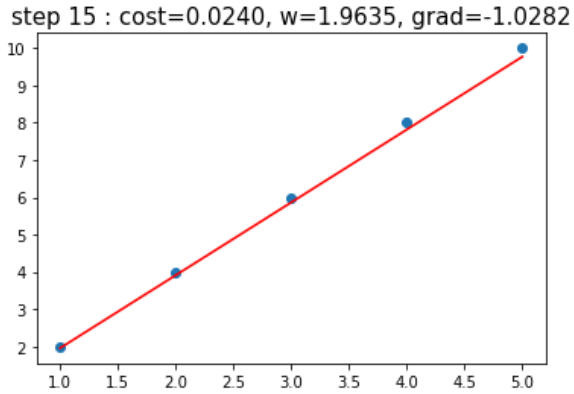
- 실습 1

```
lr = 0.01

for step in range(20):
    pre = w*x
    cost = ((pre - y) ** 2).sum() / len(x)
    #(wx-y)^2 미분 시 2(wx-y)*x
    grad = 2*(pre-y).view(5).dot(x.view(5))/len(x)
    w -= lr*grad

    if step % 5 == 0 :
        plt.scatter(x.numpy(), y.numpy())
        plt.plot(x.numpy(), pre.numpy(), 'r-')
        # w.size() = 1*1, grad.size() = 1
        plt.title('step %d : cost=%.4f, w=%.4f, grad=%.4f' % (step, cost.item(), w.item(
        plt.show()
```

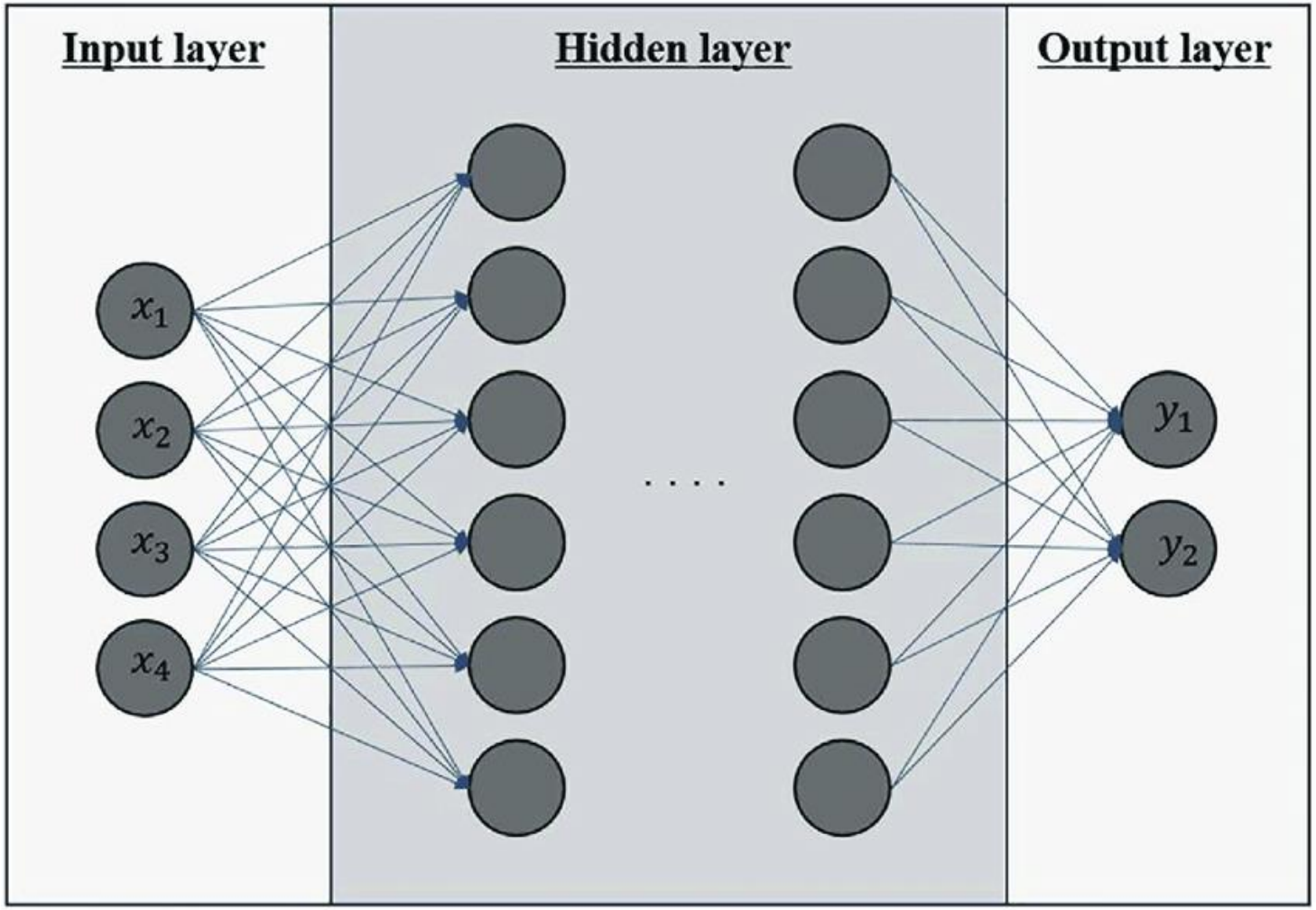








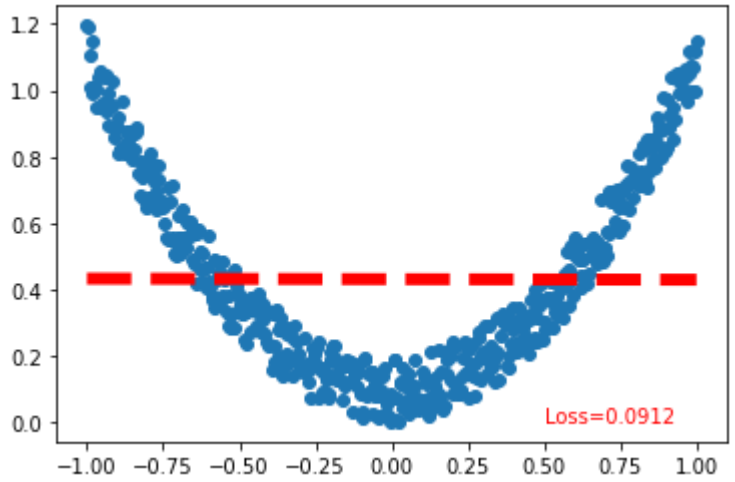
선형 신경망



모든 퍼셉트론이 선형이라면?

선형 신경망

- 실습 2
  - $y = x^2$ 그래프를 linear layer 2개를 사용한 모델로 예측하면?

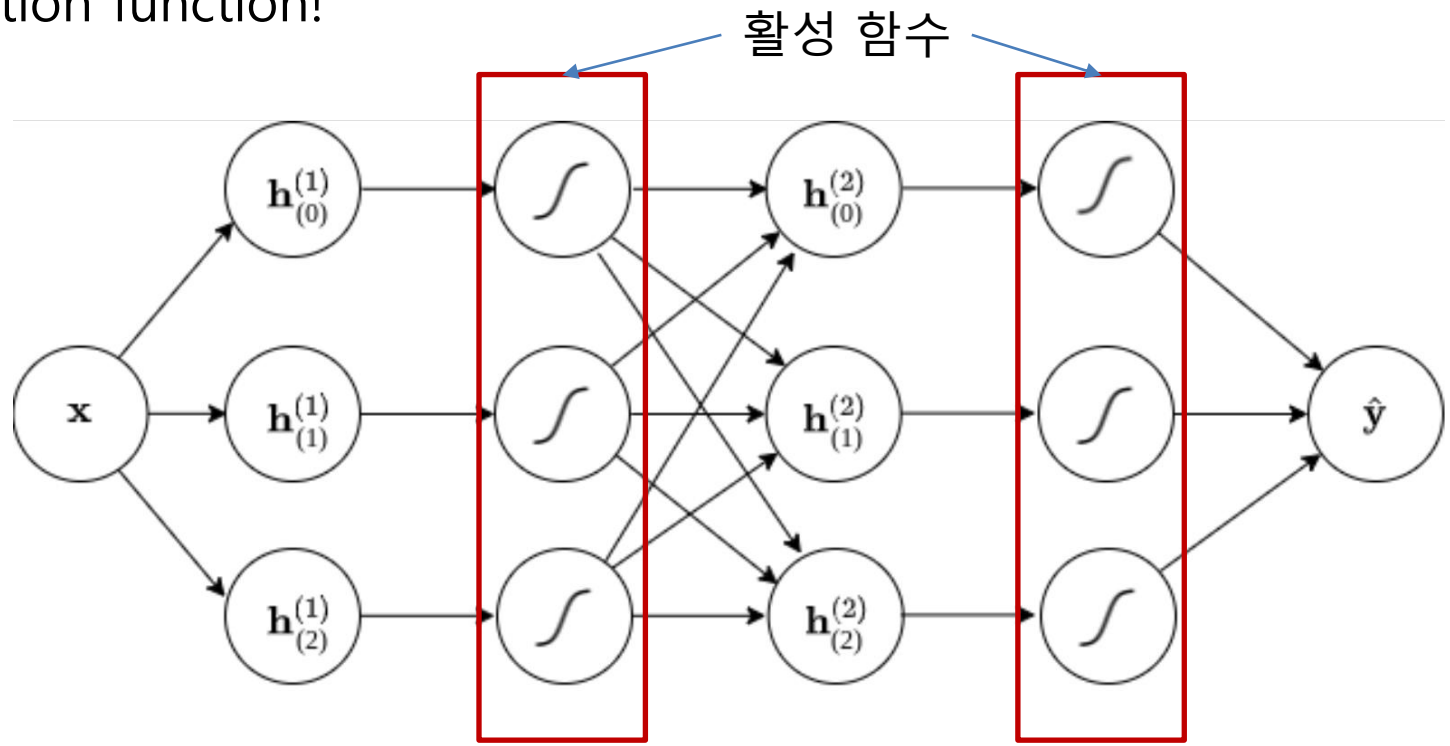


→ Linear layer를 여러 개 쌓는 것은 의미가 없다!

활성함수

- 비선형성
  - 선형결합의 선형결합은 선형결합이기 때문에, linear layer를 여러 번 쌓은 것은 의미가 없음
  - 즉, 비선형성을 갖는 함수가 필요함

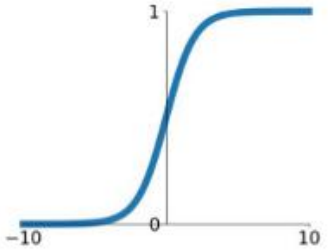
    ➔ Activation function!



활성함수

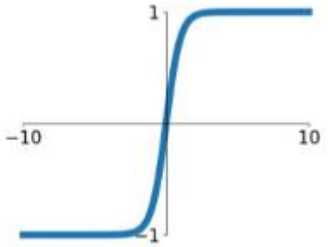
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



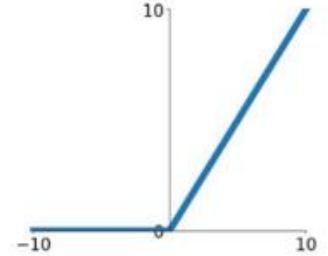
tanh

$$\tanh(x)$$



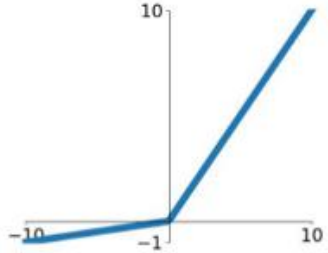
ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

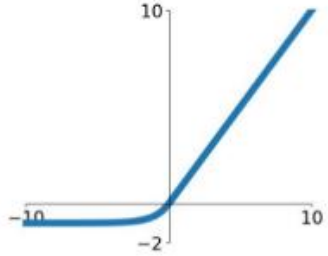


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ELU

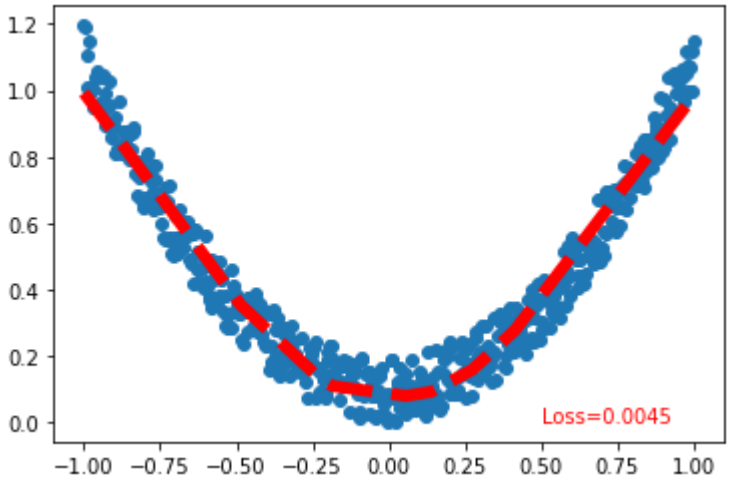
$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



일반적으로 ReLU를 많이 사용함

활성함수

- 실습 3
  - Linear layer와 activation function을 사용해서 $y = x^2$ 그래프를 예측



→ Activation function을 사용하면 비선형성을 나타낼 수 있다.

실습