LG U+ Why Not SW 캠프 6기
Python 데이터 분석 I

Logistic Regression
(2-class)

조준우
metamath@gmail.com

분류 모델

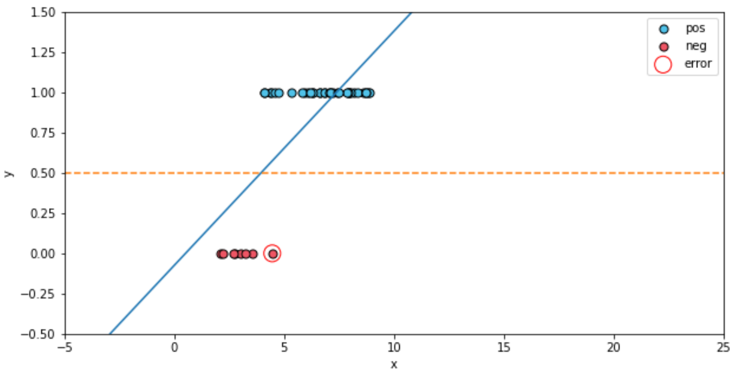
- 이산 타겟 변수와 입력의 관계를 찾는 모델

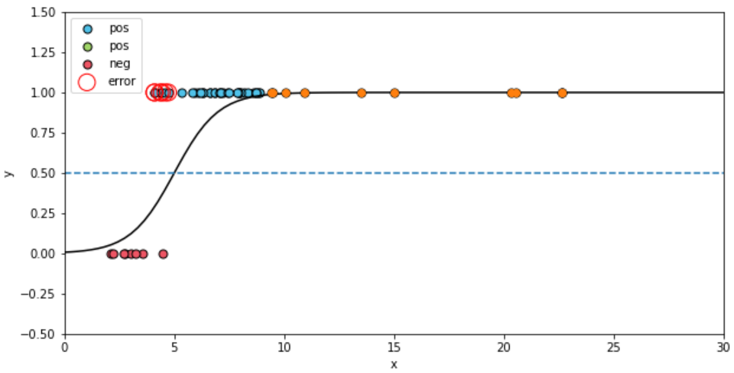
- 대표적인 알고리즘
  - 로지스틱 회귀
  - 결정 트리Decision Tree
  - 나이브 베이즈Naïve Bayes

- 응용분야
  - 스팸 필터링
  - 불량 검출
  - 개체 인식

두가지 관점

- 선형회귀에서 시작



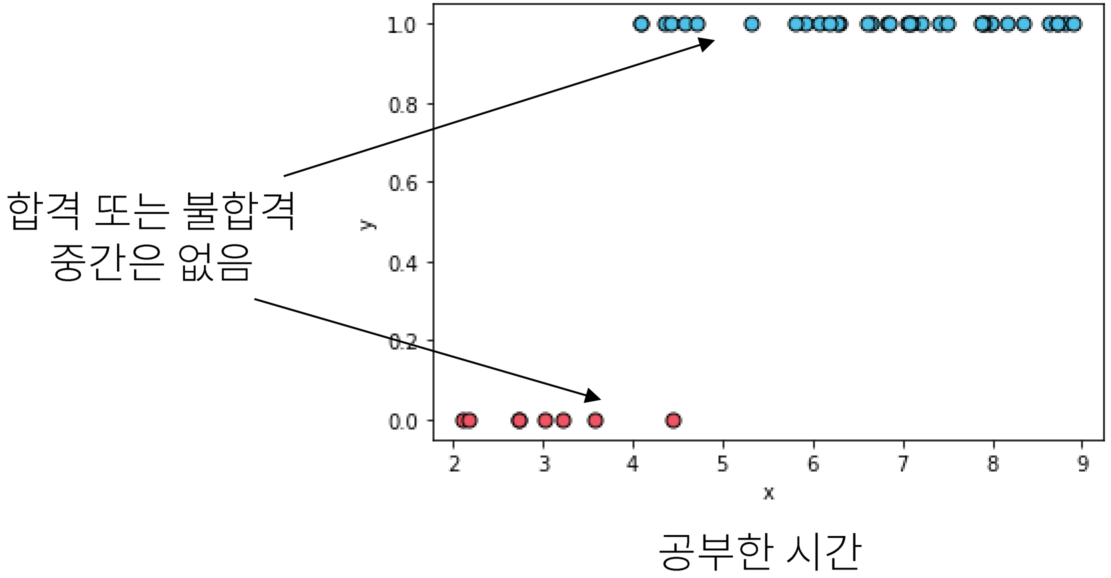


- 확률적 관점

$$p(C_1 \mid x) = \frac{p(x \mid C_1)p(C_1)}{p(x \mid C_1)p(C_1) + p(x \mid C_2)p(C_2)}$$

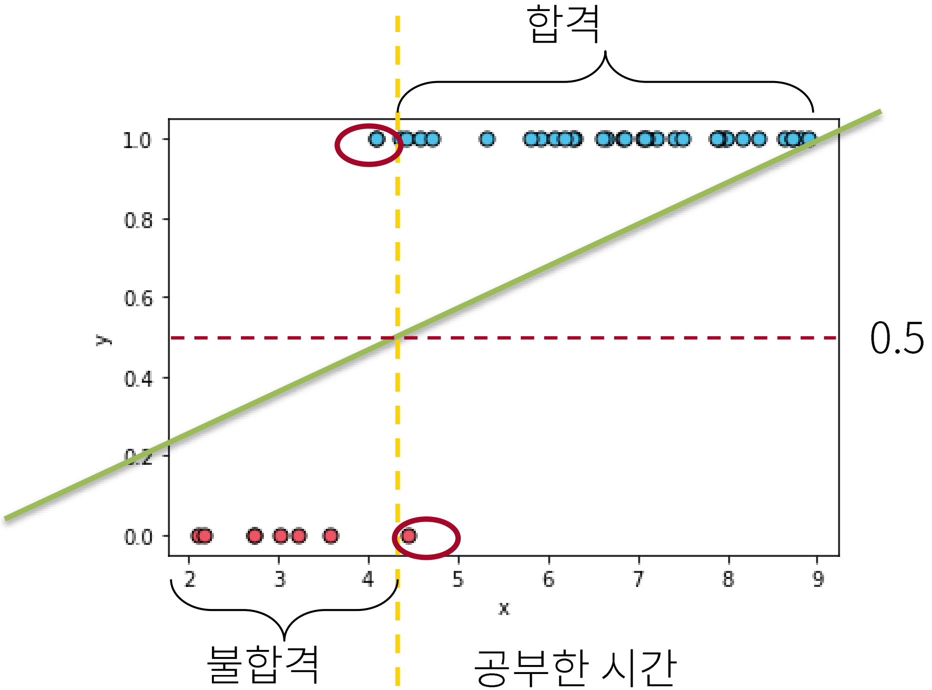
선형회귀에서 시작

- 다음 데이터에 대해……
  - 입력: 공부한 시간
  - 출력: 합격 여부



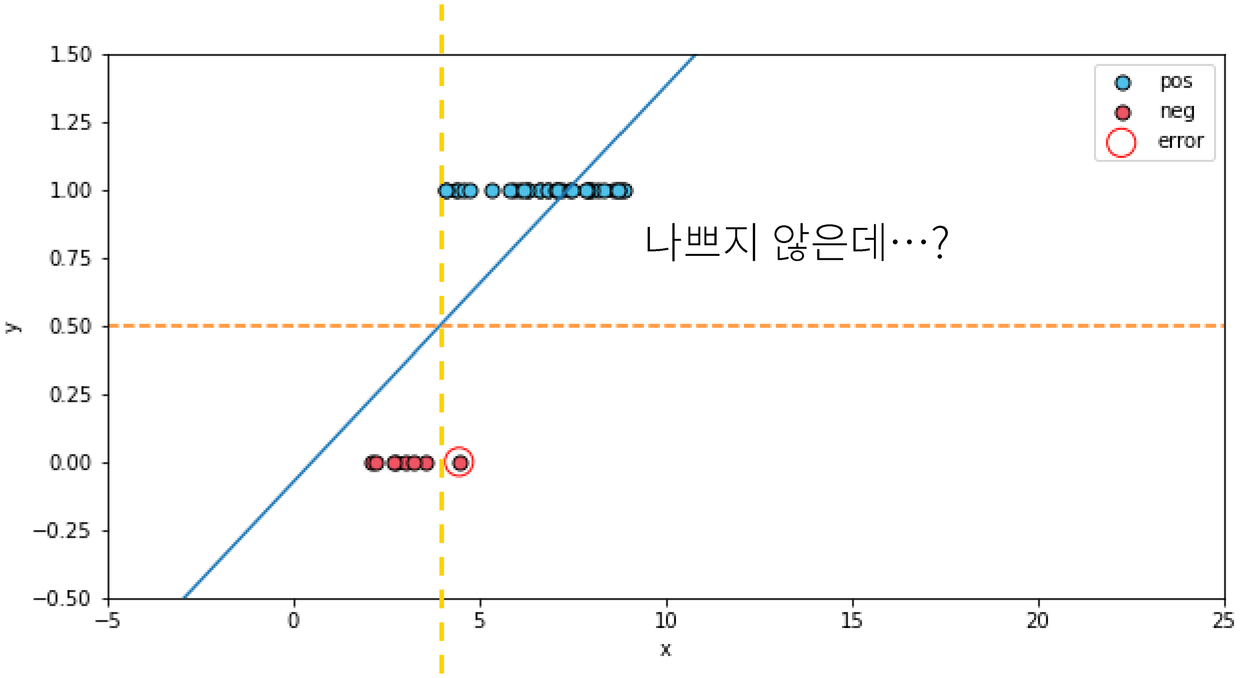
선형회귀에서 시작

- 선형회귀 후 0.5를 기준으로 결정



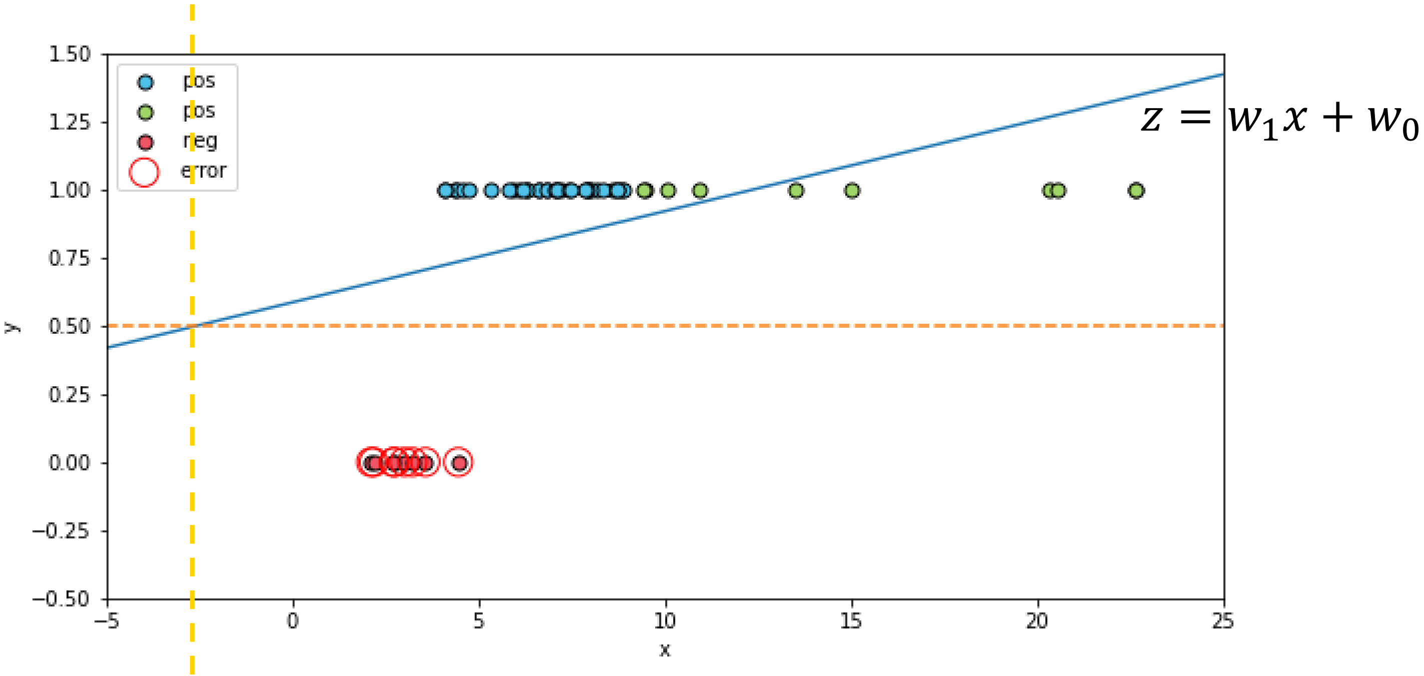
선형회귀로 실험

- 괜찮은 결과



선형회귀의 한계

- 공부를 열심히 한 학생이 많으면



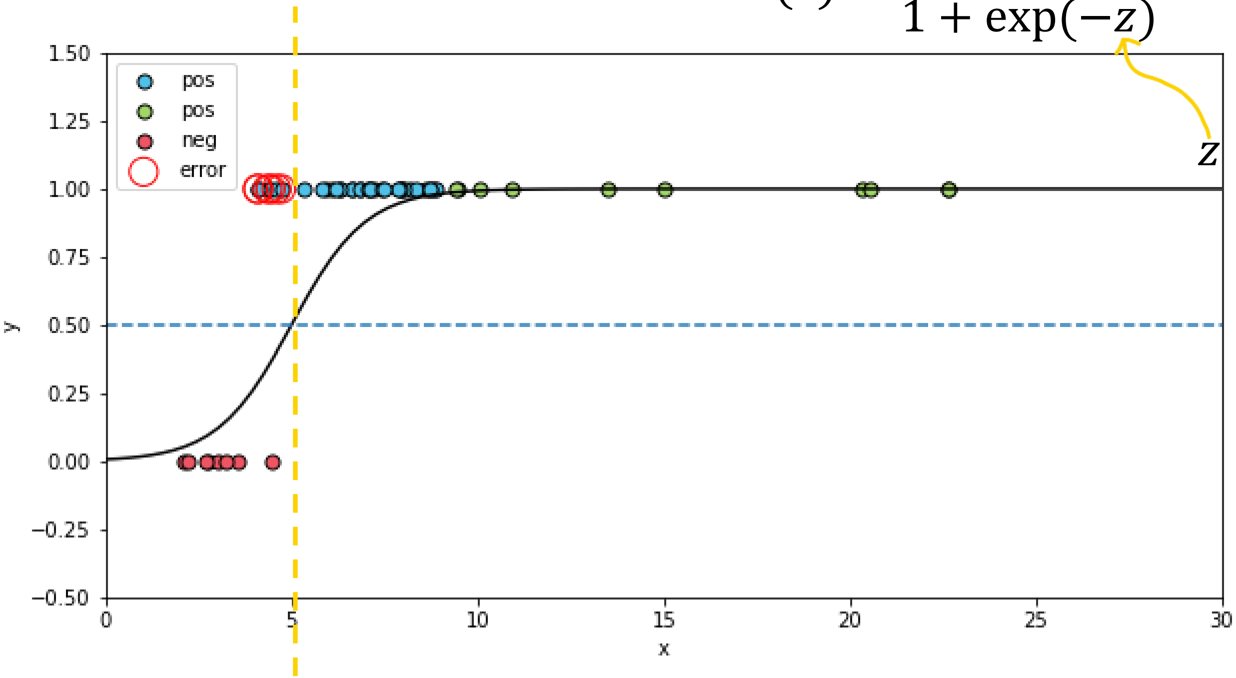
S자 곡선 형태

- 함수의 형태로 부터 결과를 계선

$$\sigma(z) = \frac{1}{1 + \exp(-z)}$$

$$z = w_1 x + w_0$$

여전히
learning param. 두개



목적함수

- $\sigma(z_i)$ 가 점과 가까우면 되니까

- 문제는?

$$\sigma(z_i) = \frac{1}{1 + \exp(-z_i)}$$

$$J(\mathbf{w}) = \frac{1}{2}\sum_{i=1}^{N}\{\sigma(z_i(x_i, \mathbf{w})) - y_i\}^2$$

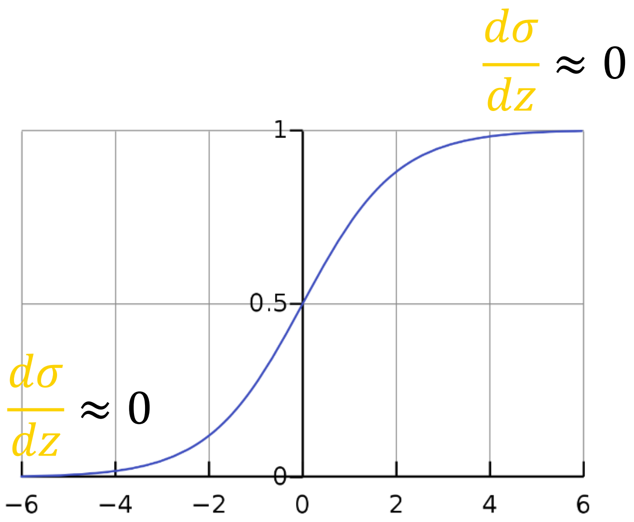
$$z_i = w_1 x_i + w_0$$

목적함수 단점

- 미분 과정에서 미분계수가 사라진다.

$$\frac{\partial}{\partial w_i} J(\mathbf{w}) = \frac{1}{2} \sum_{i=1}^{N} \{ \underbrace{\sigma(z_i(x_i, \mathbf{w}))}_{\text{합성함수→체인룰}} - y_i \}^2$$

$$\frac{d\sigma}{dz} \frac{\partial z}{\partial w_i}$$

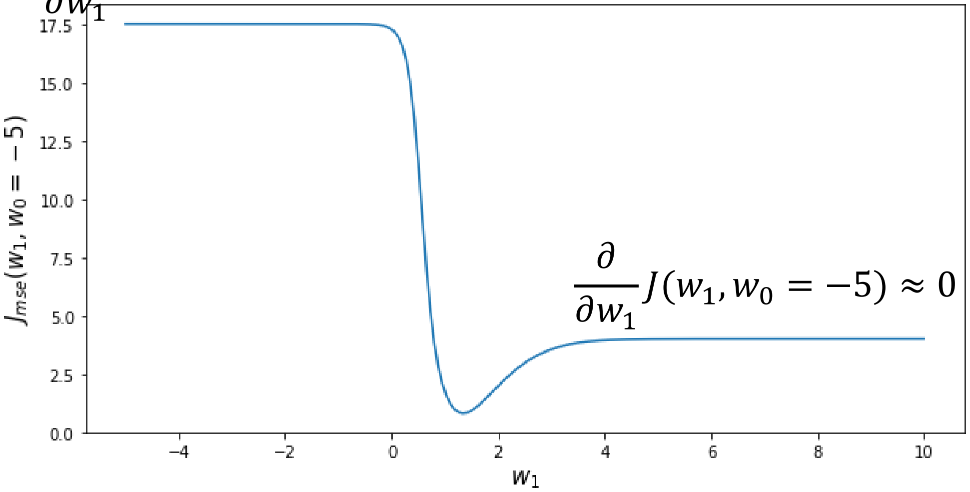


목적함수 단점

- 미분 과정에서 미분계수가 사라진다.

$$J(\mathbf{w}) = \frac{1}{2}\sum_{i=1}^{N}\{\sigma(z_i(x_i,\mathbf{w})) - y_i\}^2$$

$$\frac{\partial}{\partial w_1}J(w_1, w_0 = -5) \approx 0$$



$$\frac{\partial}{\partial w_1}J(w_1, w_0 = -5) \approx 0$$

초간단 분류기: 지수,로그함수활용

- 머신러닝 분류 문제에 있어서 평가 함수로 사용

0

1

$\mathbf{x} \in \mathbb{R}^n$

w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_9, w_{10}

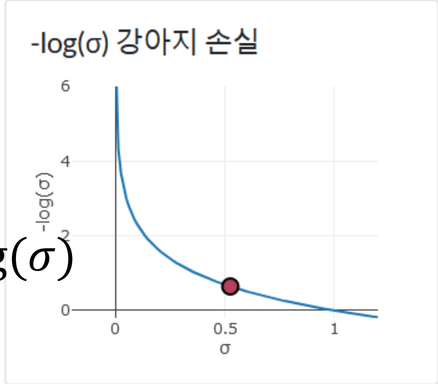
$\mathbf{w}$

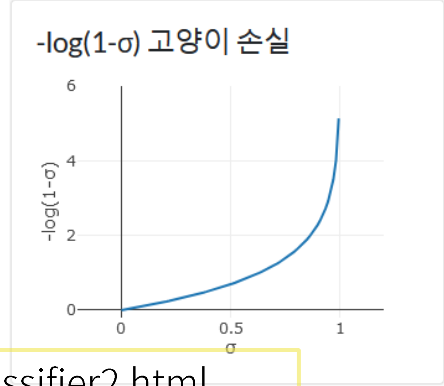
$$z: \mathbb{R}^{10} \to \mathbb{R} \Rightarrow \sigma(z) \in (0,1) \Rightarrow -\log(\sigma)$$

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

강아지

강아지 확률:0.525, 고양이 확률: 0.475





JS metamath1.github.io/noviceml/toyclassifier2.html

초간단 분류기: 지수,로그함수활용

- 머신러닝 분류 문제에 있어서 평가 함수로 사용

0

1

$\mathbf{x} \in \mathbb{R}^n$

w_1, $\boxed{w_2 = 2}$, w_3, w_4, w_5, w_6, w_7, w_8, w_9, w_{10}

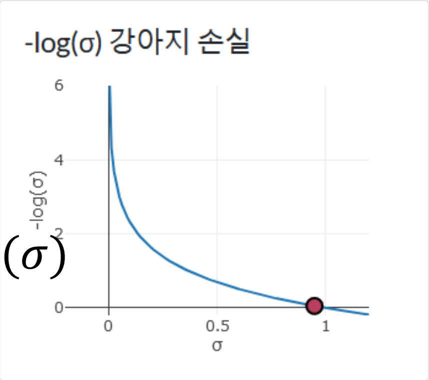
$\mathbf{w}$

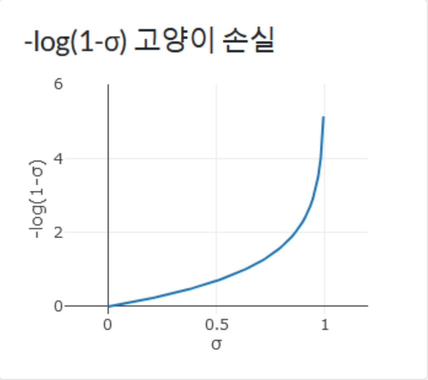
$z: \mathbb{R}^n \to \mathbb{R} \Rightarrow \sigma(z) \in (0,1) \Rightarrow -\log(\sigma)$

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

강아지

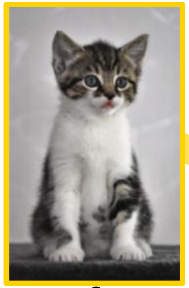
강아지 확률:0.949, 고양이 확률: 0.051





초간단 분류기: 지수,로그함수활용

- 머신러닝 분류 문제에 있어서 평가 함수로 사용



0

1

$\mathbf{x} \in \mathbb{R}^n$

w_1

$\boxed{w_2 = 2}$

w_3 w_4 w_5 w_6 w_7 w_8 w_9 w_{10}

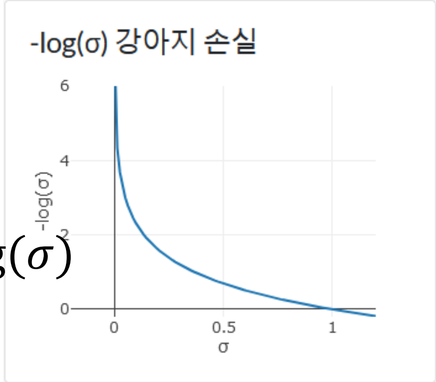
$\mathbf{w}$

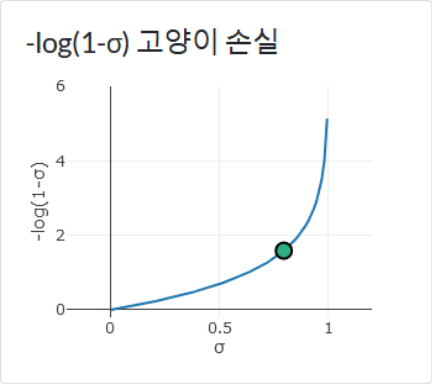
$$z: \mathbb{R}^n \to \mathbb{R} \Rightarrow \sigma(z) \in (0,1) \Rightarrow -\log(\sigma)$$

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

강아지

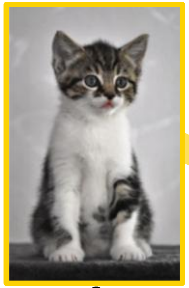
강아지 확률:0.796, 고양이 확률: 0.204





초간단 분류기: 지수,로그함수활용

- 머신러닝 분류 문제에 있어서 평가 함수로 사용



0

1

$\mathbf{x} \in \mathbb{R}^n$

w_1

$\boxed{w_2 = 2}$

w_3

w_4

w_5

w_6

w_7

$\boxed{w_8 = -4}$

w_9

w_{10}

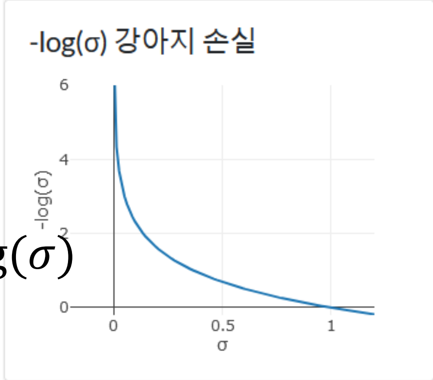
$\mathbf{w}$

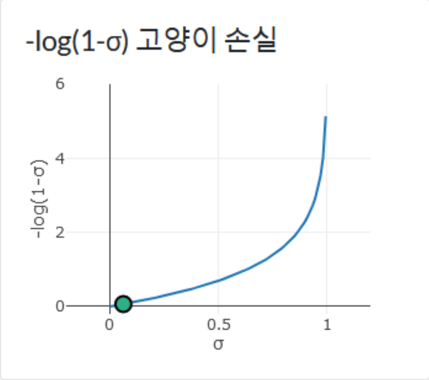
$z: \mathbb{R}^n \rightarrow \mathbb{R} \Rightarrow \sigma(z) \in (0,1) \Rightarrow -\log(\sigma)$

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

강아지

강아지 확률:0.949, 고양이 확률: 0.051





새로운 목적함수

- $\log(\cdot)$ 특성 이용
- y_i 는 0 또는 1

$$J(\mathbf{w}) = -\sum_{i=1}^{N}[y_i \log \sigma(z_i(x_i, \mathbf{w})) + (1 - y_i)\log(1 - \sigma(z_i(x_i, \mathbf{w})))]$$

$$z_i = w_1 x_i + w_0$$

개선된 목적함수

- 지속적으로 감소하거나 상승하는 형태

$$J(\mathbf{w}) = -\sum_{i=1}^{N} [y_i \log \sigma(z_i(x_i, \mathbf{w})) + (1 - y_i) \log(1 - \sigma(z_i(x_i, \mathbf{w})))]$$