

FastCampus Pytorch

Ch2. Basics for Model Training

HARRY KIM

Lecture Content

- 1 Model Training
- 2 Linear Regression
- 3 Neural Network
- 4 Logistic Regression

Model Training

Linear Regression

Neural Network

Logistic Regression

■ 강의 자료

- Books
 - 밑바닥부터 시작하는 딥러닝 [사이토 고키, 2017]
 - 핸드온 머신러닝 [오렐리앙 제롱, 2018]
 - 머신러닝, 딥러닝 실전개발 입문 [쿠지라 히코우즈쿠에, 2017]
- Online
 - Machine Learning [<https://ko.coursera.org/learn/machine-learning>, 2018]

Model Training

**Linear
Regression**

Neural Network

**Logistic
Regression**

1. Model Training

Model Training

Model Training

Linear
Regression

Neural Network

Logistic
Regression

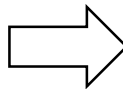
머신러닝(Machine Learning)

- 데이터로부터 규칙을 학습하는 프로그래밍
- 컴퓨터가 학습하는 능력을 갖춘 것 (Arthur Samuel, 1959)
- 작업 T에 대해 경험 E을 학습하여 성능 P를 향상시킨 것 (Tom Mitchell, 1997)

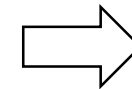
학습(Training)



데이터(DATA)



모델(MODEL)



[5, 0, 4]

데이터(DATA)

Model Training

Linear
Regression

Neural Network

Logistic
Regression

■ 모델 학습 방법

- 지도/교사 학습(Supervised Learning)
 - 데이터와 정답을 제공하여 학습하게 하는 알고리즘 (ex. 회귀/분류)
 - 비지도/비교사 학습(Unsupervised Learning)
 - 데이터만을 제공하여 학습하게 하는 알고리즘 (ex. 클러스터링)
-
- 강화학습(Reinforcement Learning)
 - 주어진 리워드를 최대화하도록 학습하는 알고리즘

Model Training

Linear
Regression

Neural Network

Logistic
Regression

■ 모델 학습 방법

- 지도/교사 학습(Supervised Learning)
 - 데이터와 정답을 제공하여 학습하게 하는 알고리즘 (ex. 회귀/분류)
 - 비지도/비교사 학습(Unsupervised Learning)
 - 데이터만을 제공하여 학습하게 하는 알고리즘 (ex. 클러스터링)
-
- 강화학습(Reinforcement Learning)
 - 주어진 리워드를 최대화하도록 학습하는 알고리즘

Model Training

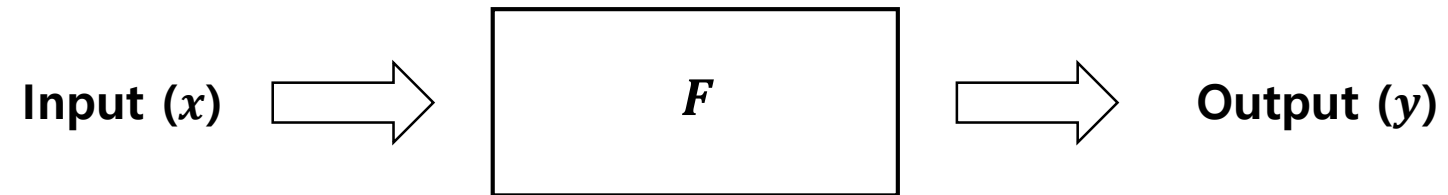
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- **교사 학습(Supervised Learning)**
 - 데이터와 정답을 제공하여 컴퓨터를 학습시키는 알고리즘
 - 이상(Ideal) 모델



Model Training

Model Training

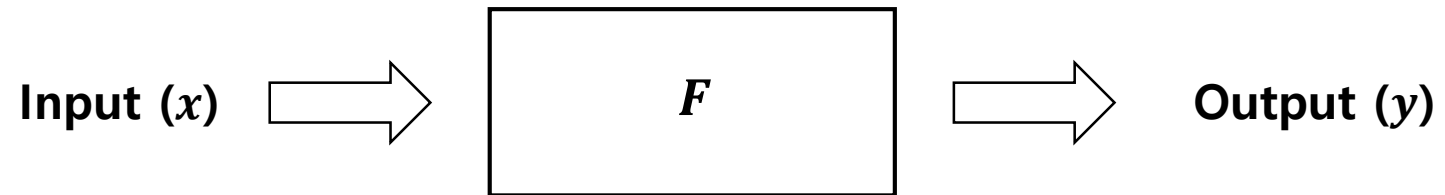
Linear
Regression

Neural Network

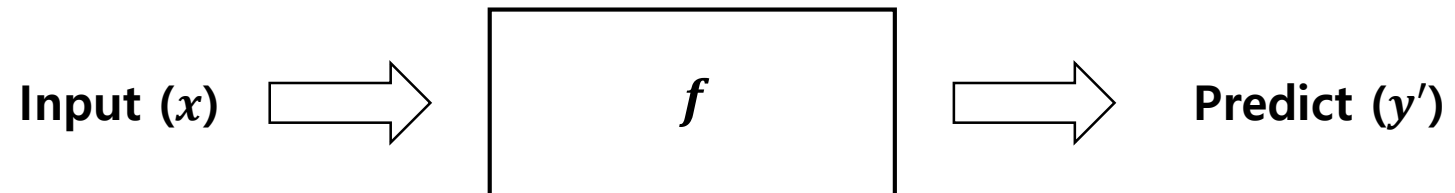
Logistic
Regression

■ 교사 학습(Supervised Learning)

- 데이터와 정답을 제공하여 컴퓨터를 학습시키는 알고리즘
 - 이상(Ideal) 모델



- 현실(Real) 모델



Model Training

Model Training

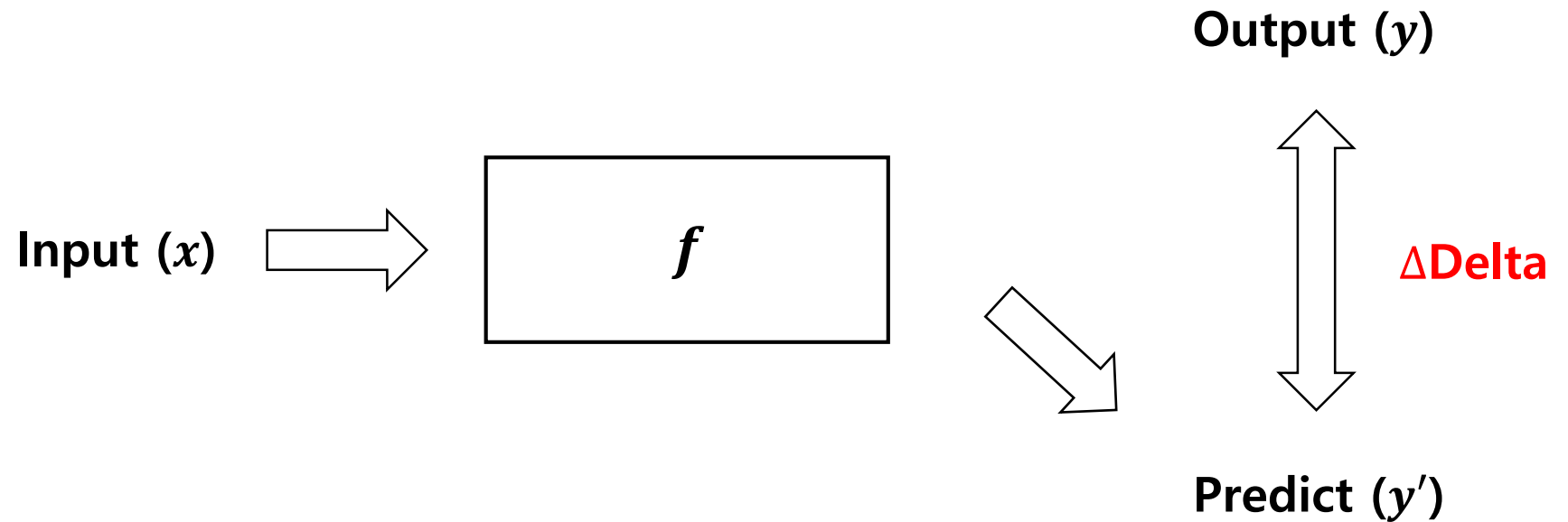
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습(Supervised Learning)

- 실제와 예측의 차이를 줄이는 것이 목표
- 차이는 어떻게 정의할 것인가?
 - 절댓값? 제곱값?



Model Training

Model Training

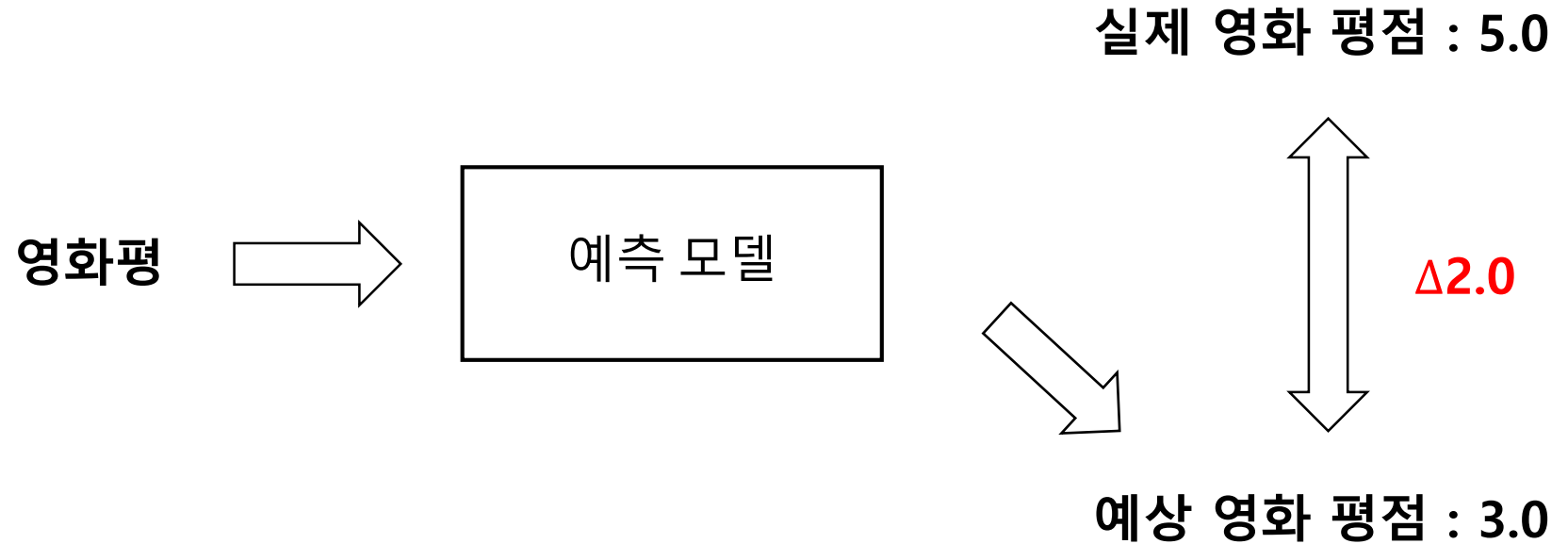
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습(Supervised Learning)

- 실제와 예측의 차이를 줄이는 것이 목표
- 차이는 어떻게 정의할 것인가?
 - 절댓값? 제곱값?



Model Training

Model Training

Linear
Regression

Neural Network

Logistic
Regression

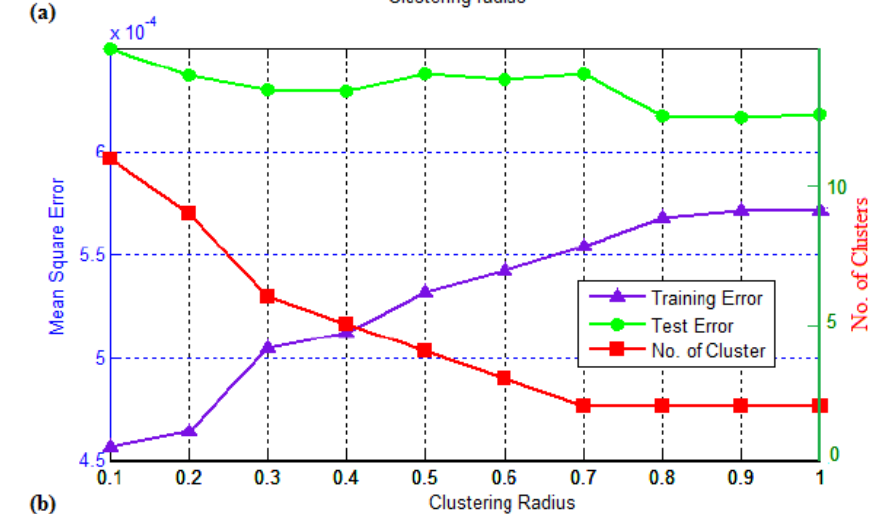
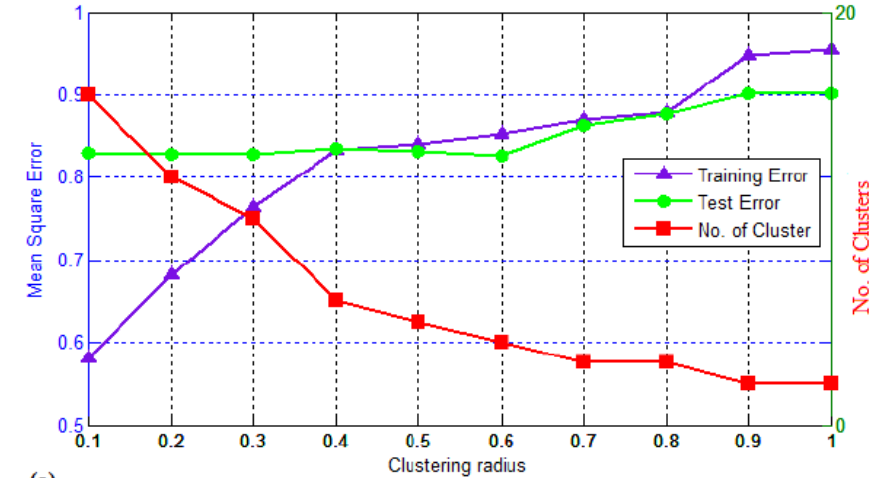
손실 함수(Loss Function)

- 최적의 모델을 찾기 위한 목적 함수
- 평균 제곱 오차(Mean Squared Error, MSE)
 - 회귀에서 자주 사용되는 손실 함수
 - 오차의 제곱의 합의 평균

$$MSE = \frac{\sum err^2}{n}$$

- 교차 엔트로피 오차(Cross Entropy Error, CEE)
 - 분류에서 자주 사용되는 손실 함수
 - 정답과 오차의 엔트로피의 합

$$CEE = -\sum t_k \log y_k$$



https://www.researchgate.net/figure/Graphs-showing-the-mean-square-error-MSE-performance-of-Takagi-Sugeno-fuzzy-inference_fig6_281662470

Model Training

Model Training

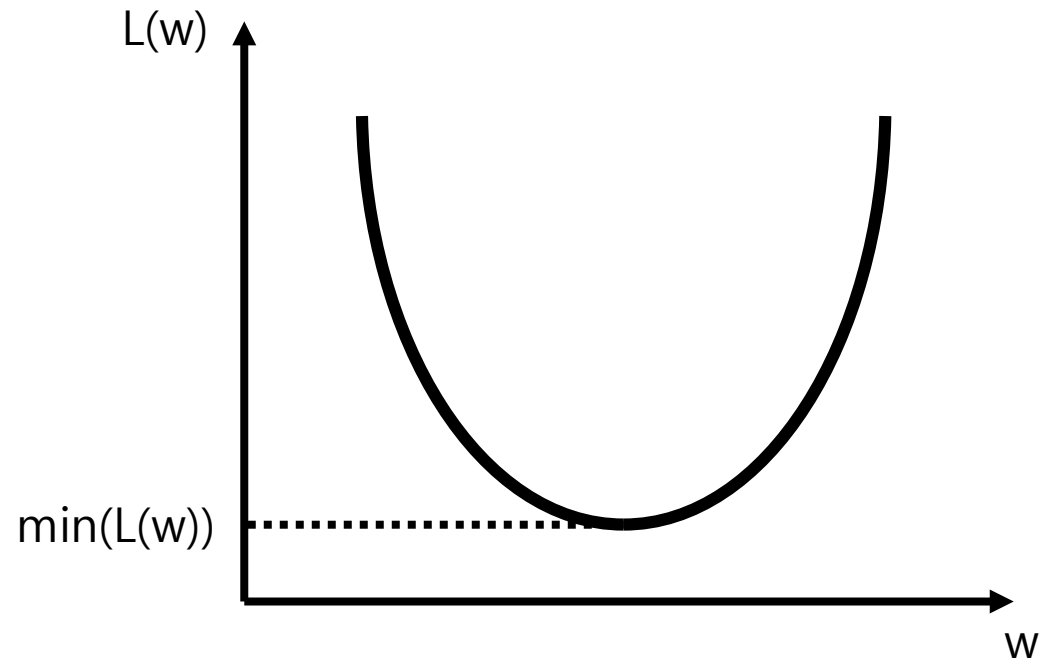
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 목표

- 최대한 정확한 값을 예측하자
- = 손실 함수를 최소한으로 줄이자
- = $\min(\text{Loss Function})$



Model Training

Model Training

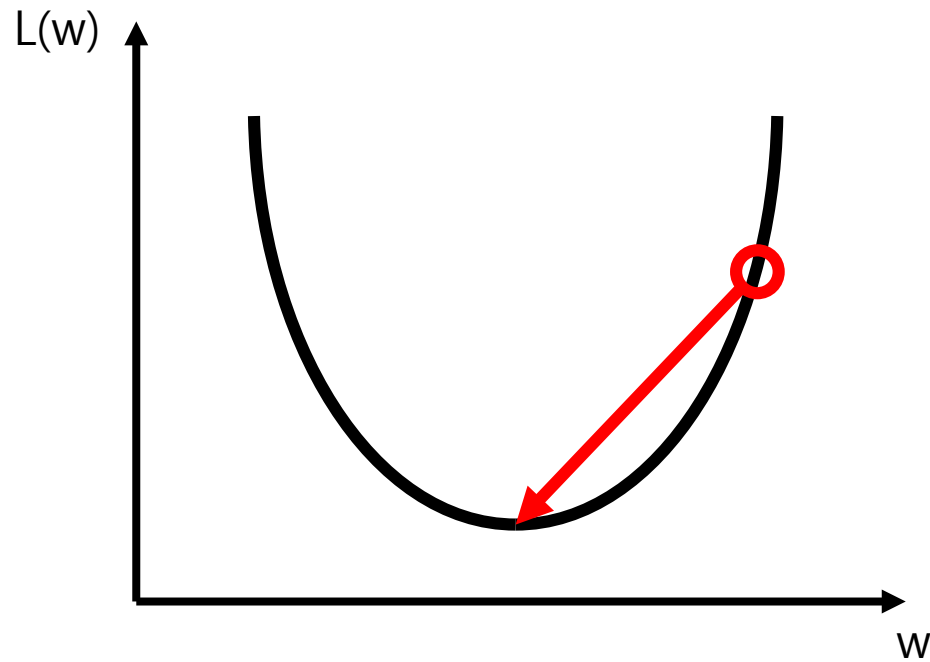
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 목표

- 어떻게 손실 함수를 줄일 것인가?
- 정규방정식(Normal Equation)
- 경사하강법(GD)/확률적경사하강법(SGD)...



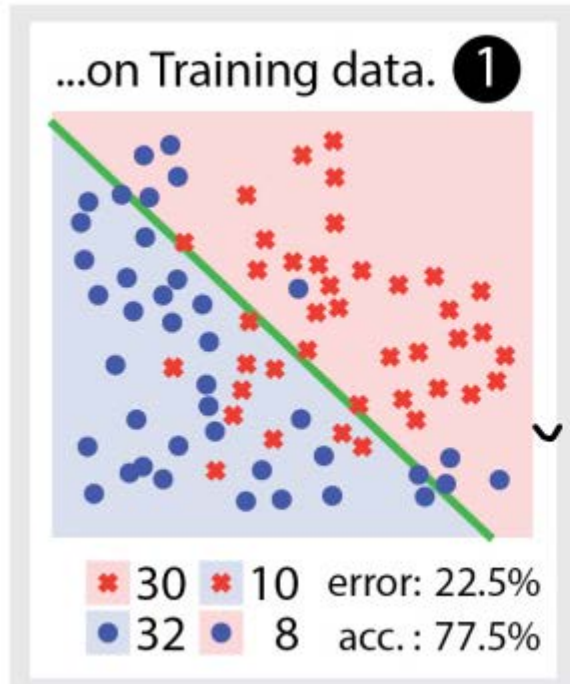
Model Training

Model Training

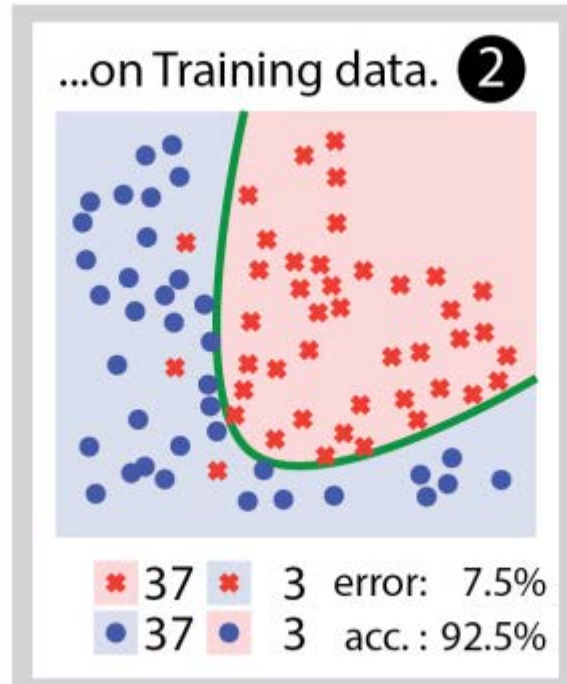
교사 학습의 목표

- 그렇다면 무작정 손실 함수를 줄이면 되는 것인가?

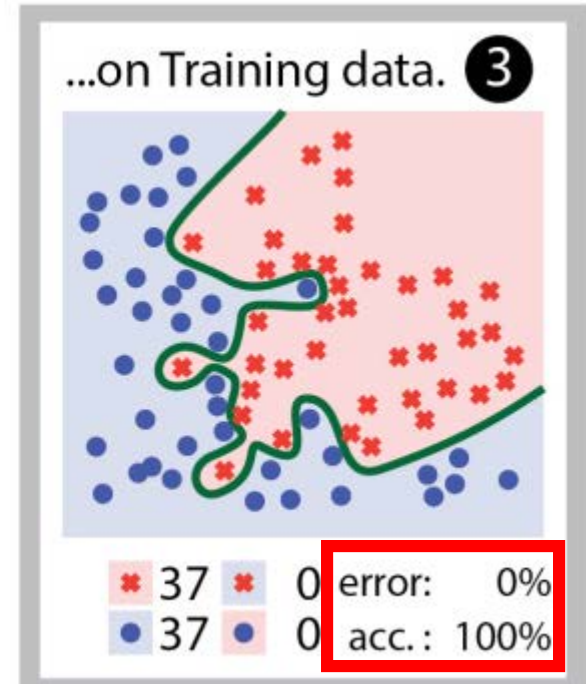
Model 1...



Model 2...



Model 3...



<https://sigmoidal.io/machine-learning-terminology-explained-top-8-must-know-concepts/overfitting/>

Linear
Regression

Neural Network

Logistic
Regression

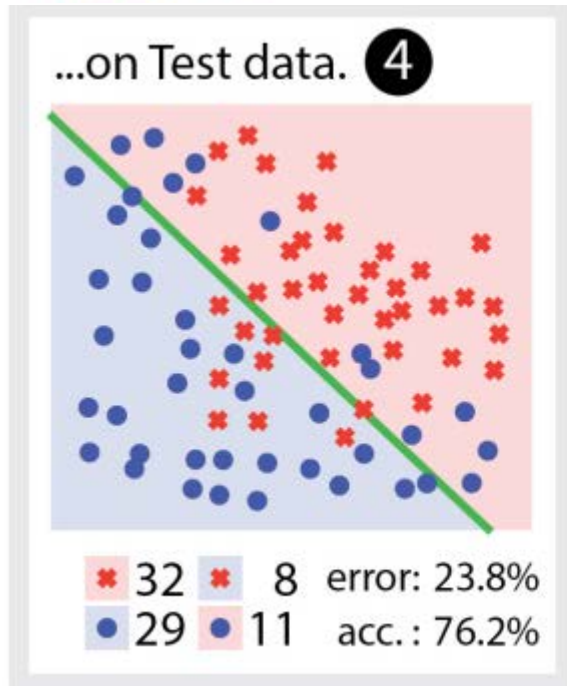
Model Training

Model Training

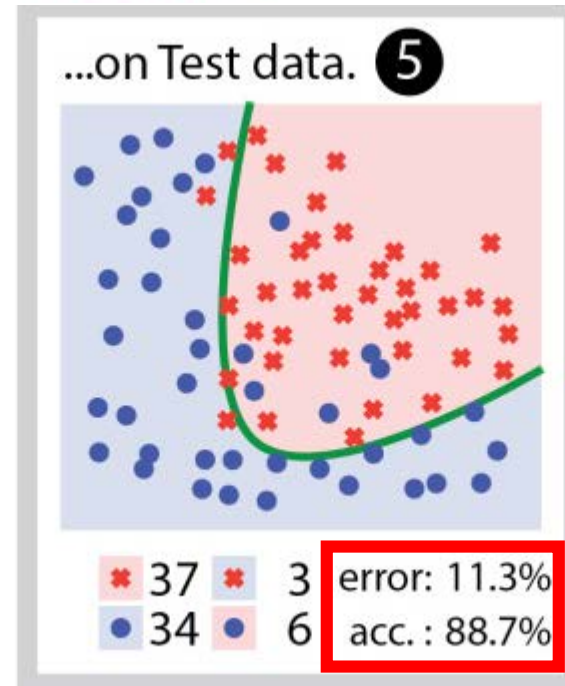
교사 학습의 목표

- 주어진 데이터에 대해 손실 함수를 최소화하는 것은 오히려 **독**이 될 수 있다.

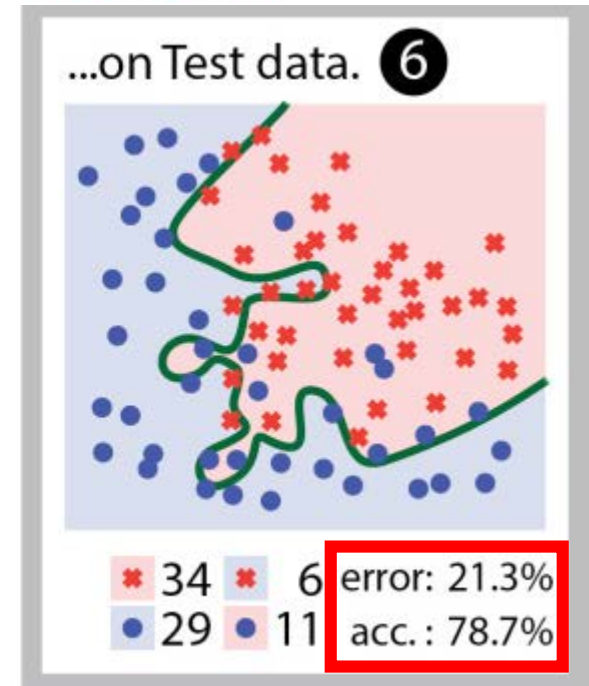
Model 1...



Model 2...



Model 3...



<https://sigmoidal.io/machine-learning-terminology-explained-top-8-must-know-concepts/overfitting/>

Linear
Regression

Neural Network

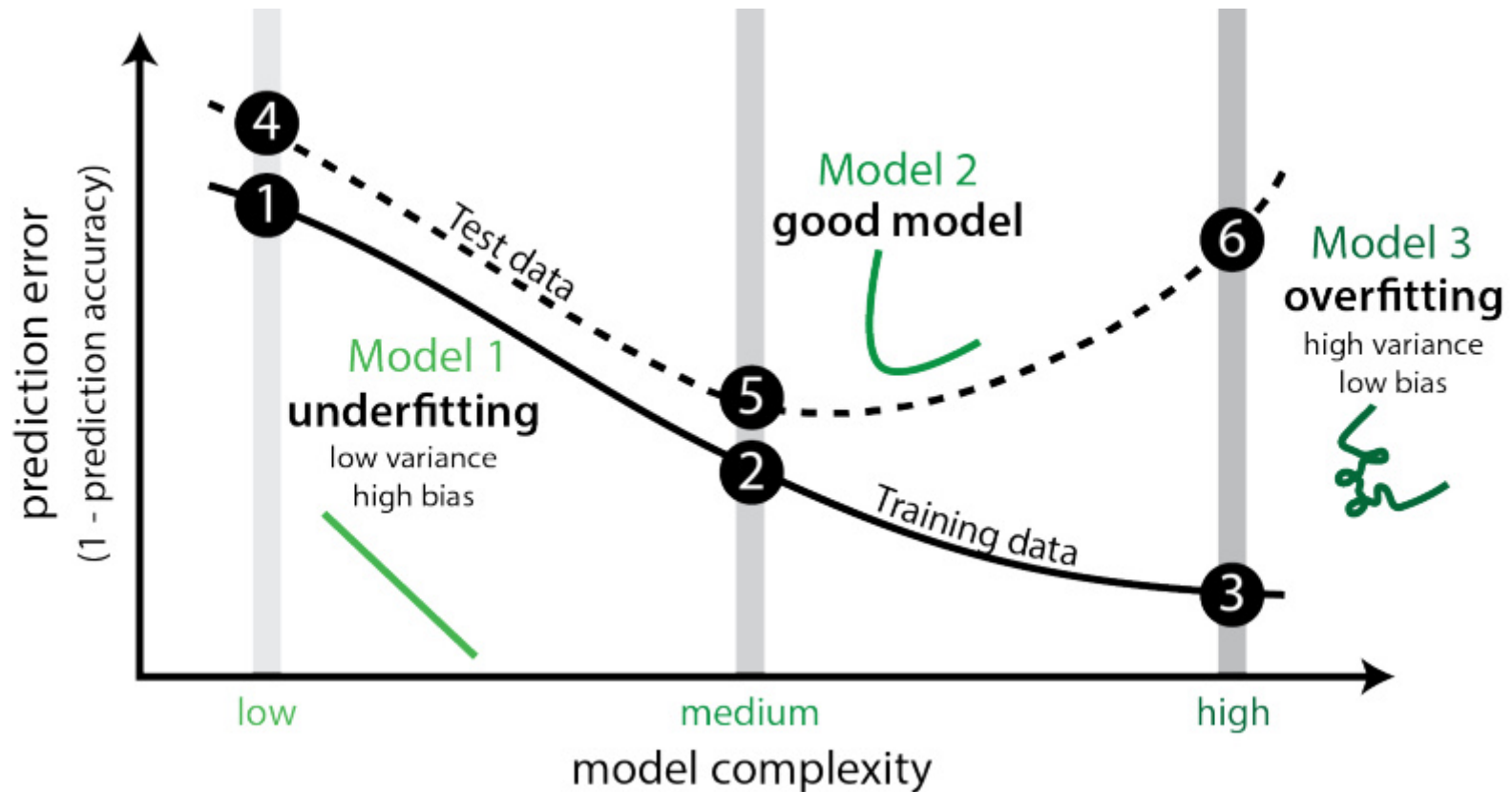
Logistic
Regression

Model Training

Model Training

■ 교사 학습의 목표

- 주어진 데이터에 대해 손실 함수를 최소화하는 것은 오히려 독이 될 수 있다.



Model Training

Model Training

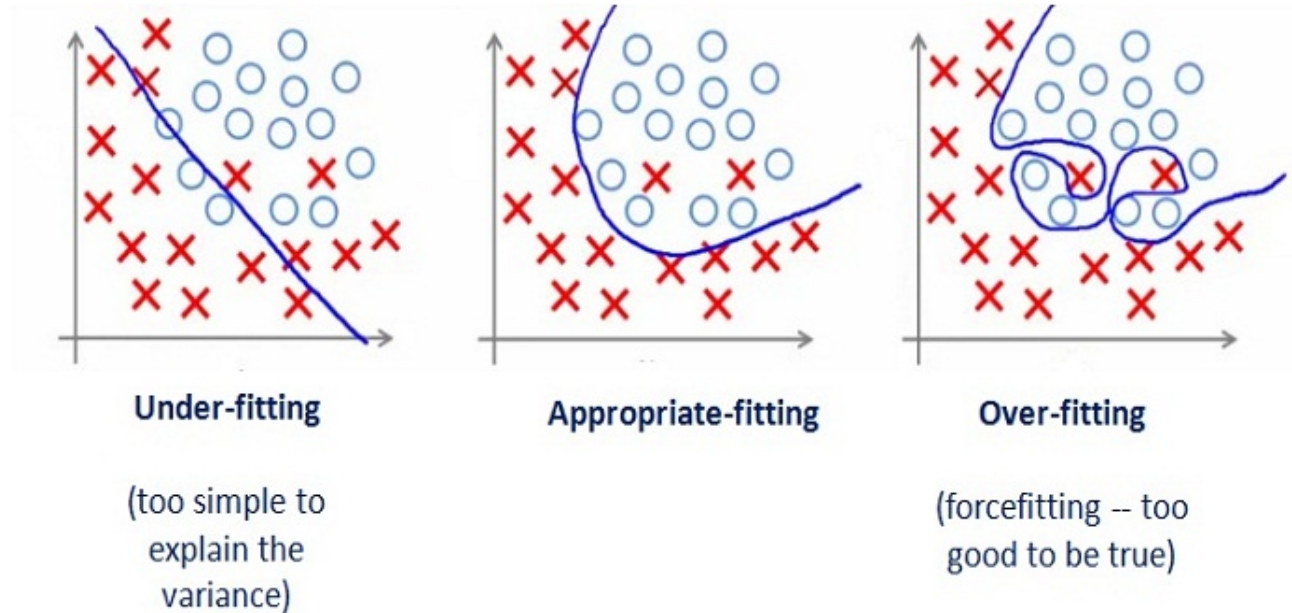
Linear
Regression

Neural Network

Logistic
Regression

교사 학습의 목표

- 과적합(Overfitting)



- 이를 방지하기 위해 데이터를 **Train/Validation/Test**로 나누어 진행

Model Training

Model Training

Linear
Regression

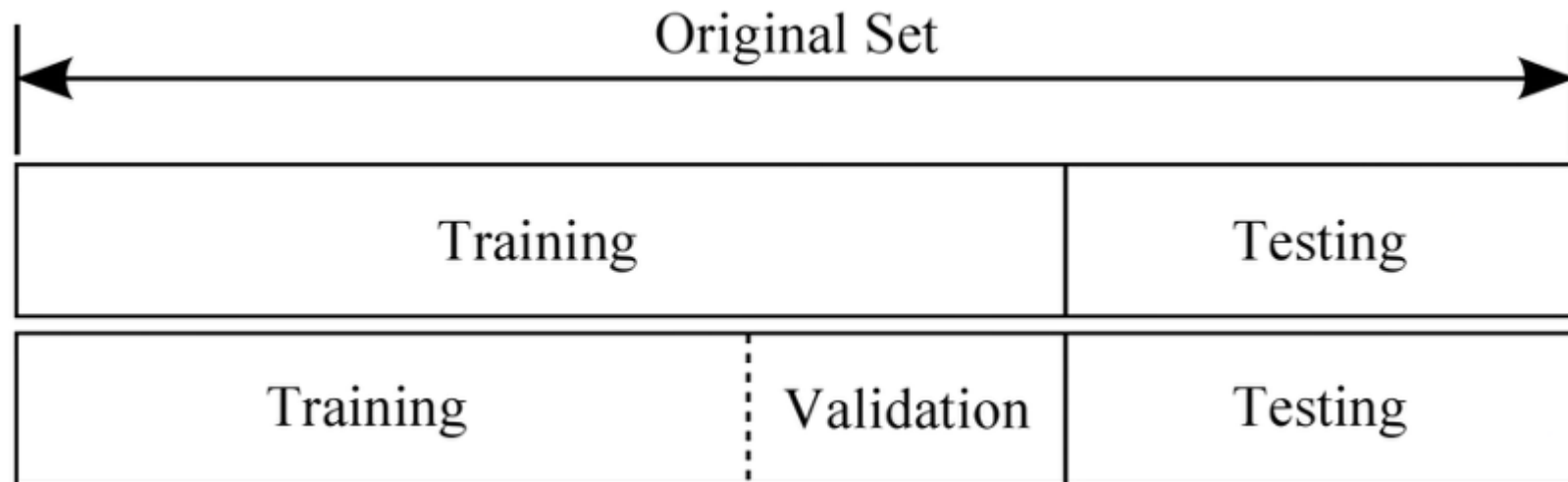
Neural Network

Logistic
Regression

■ 교사 학습의 목표

■ Train/Validation/Test

- **Train** : 학습을 위한 데이터
- **Validation** : 모델 성능 검증을 위한 데이터
- **Test** : 실제 최종 도출된 모델을 적용할 데이터 (모델 학습에 관여 안함)



Model Training

Model Training

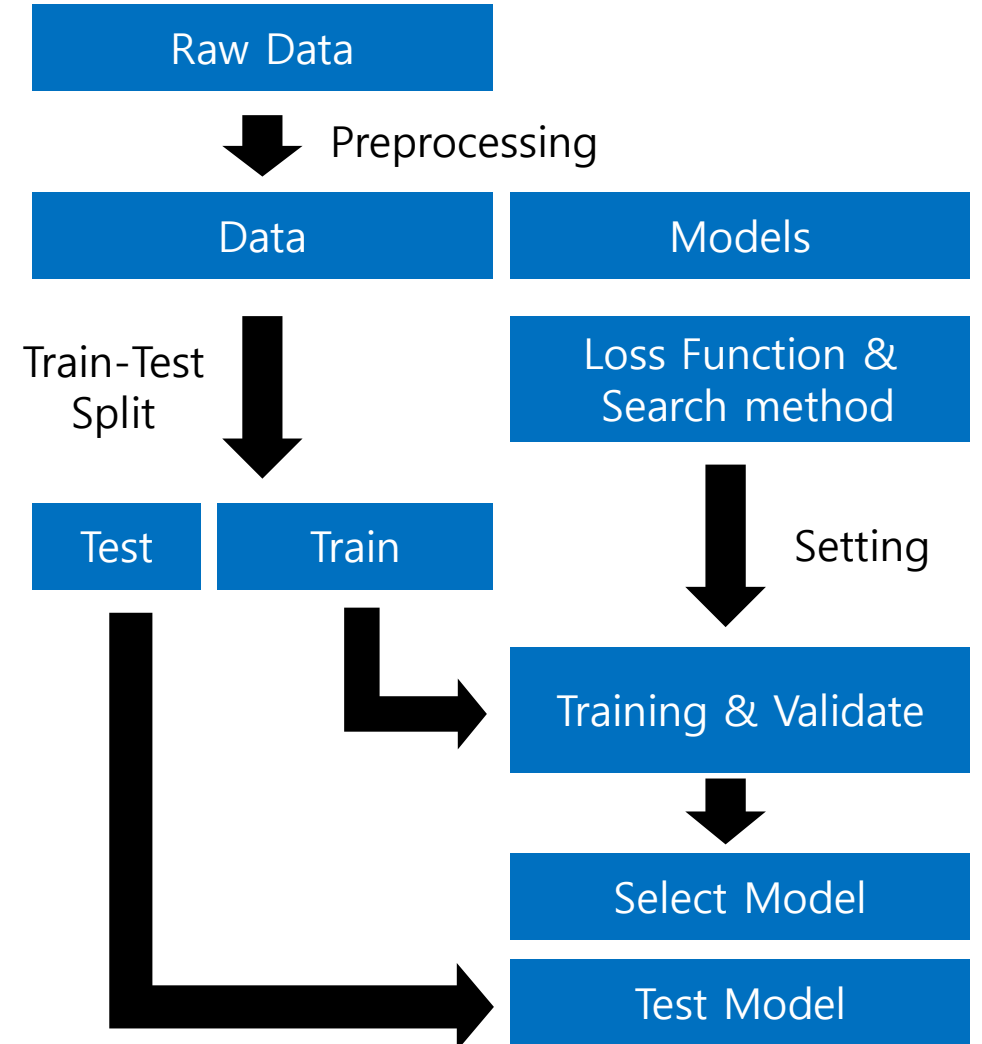
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 모델 학습 과정

1. 데이터 수집 및 분석/전처리
2. Training/Test Set로 데이터 분할
3. 모델 후보군 설정
4. 손실 함수 및 탐색 방법 설정
5. Training Set를 통한 모델 학습
(* 필요시 Validation Set을 통한 모델 성능 비교)
6. Test Set을 통한 성능/과적합 확인
7. (1-6) 반복



Model Training

Model Training

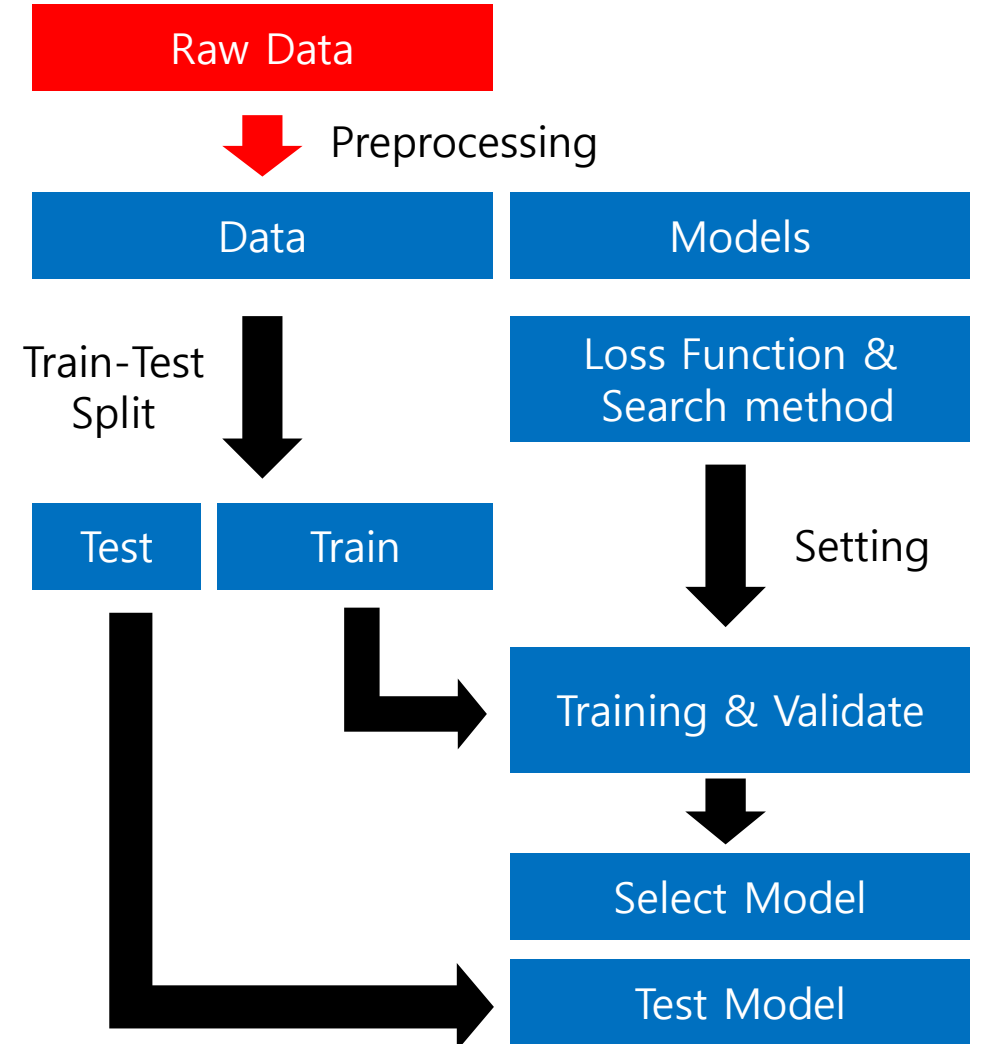
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 모델 학습 과정

1. 데이터 수집 및 분석/전처리
2. Training/Test Set로 데이터 분할
3. 모델 후보군 설정
4. 손실 함수 및 탐색 방법 설정
5. Training Set를 통한 모델 학습
(* 필요시 Validation Set을 통한 모델 성능 비교)
6. Test Set을 통한 성능/과적합 확인
7. (1-6) 반복



Model Training

Model Training

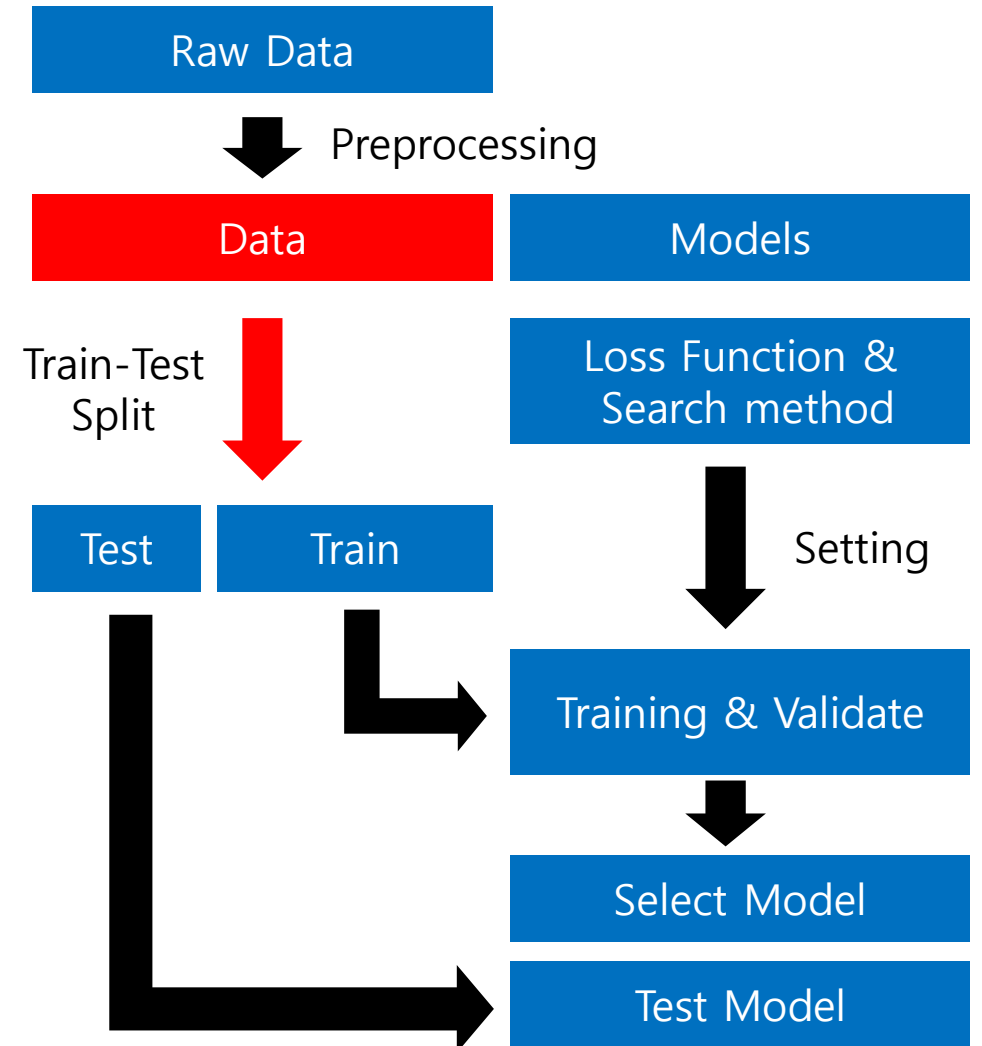
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 모델 학습 과정

1. 데이터 수집 및 분석/전처리
2. **Training/Test Set로 데이터 분할**
3. 모델 후보군 설정
4. 손실 함수 및 탐색 방법 설정
5. Training Set를 통한 모델 학습
(* 필요시 Validation Set을 통한 모델 성능 비교)
6. Test Set을 통한 성능/과적합 확인
7. (1-6) 반복



Model Training

Model Training

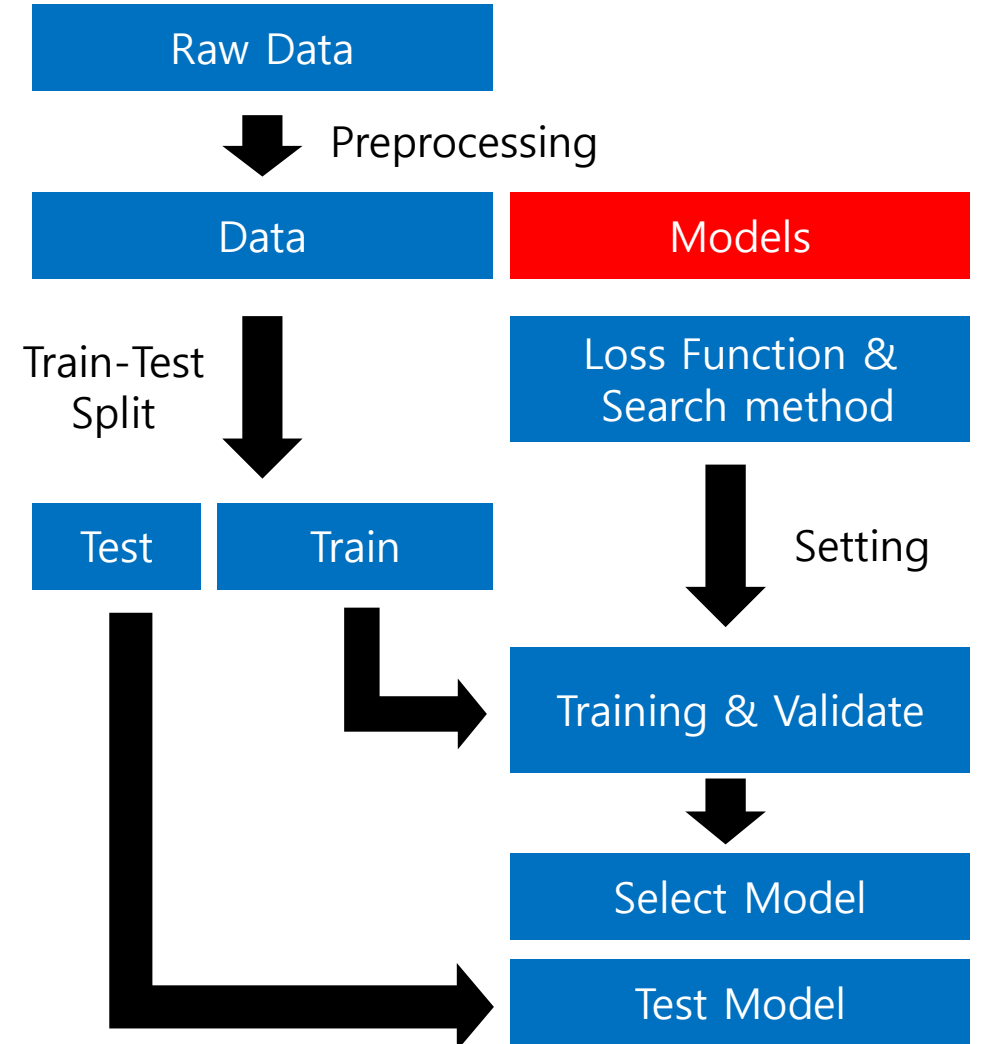
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 모델 학습 과정

1. 데이터 수집 및 분석/전처리
2. Training/Test Set로 데이터 분할
- 3. 모델 후보군 설정**
4. 손실 함수 및 탐색 방법 설정
5. Training Set를 통한 모델 학습
(* 필요시 Validation Set을 통한 모델 성능 비교)
6. Test Set을 통한 성능/과적합 확인
7. (1-6) 반복



Model Training

Model Training

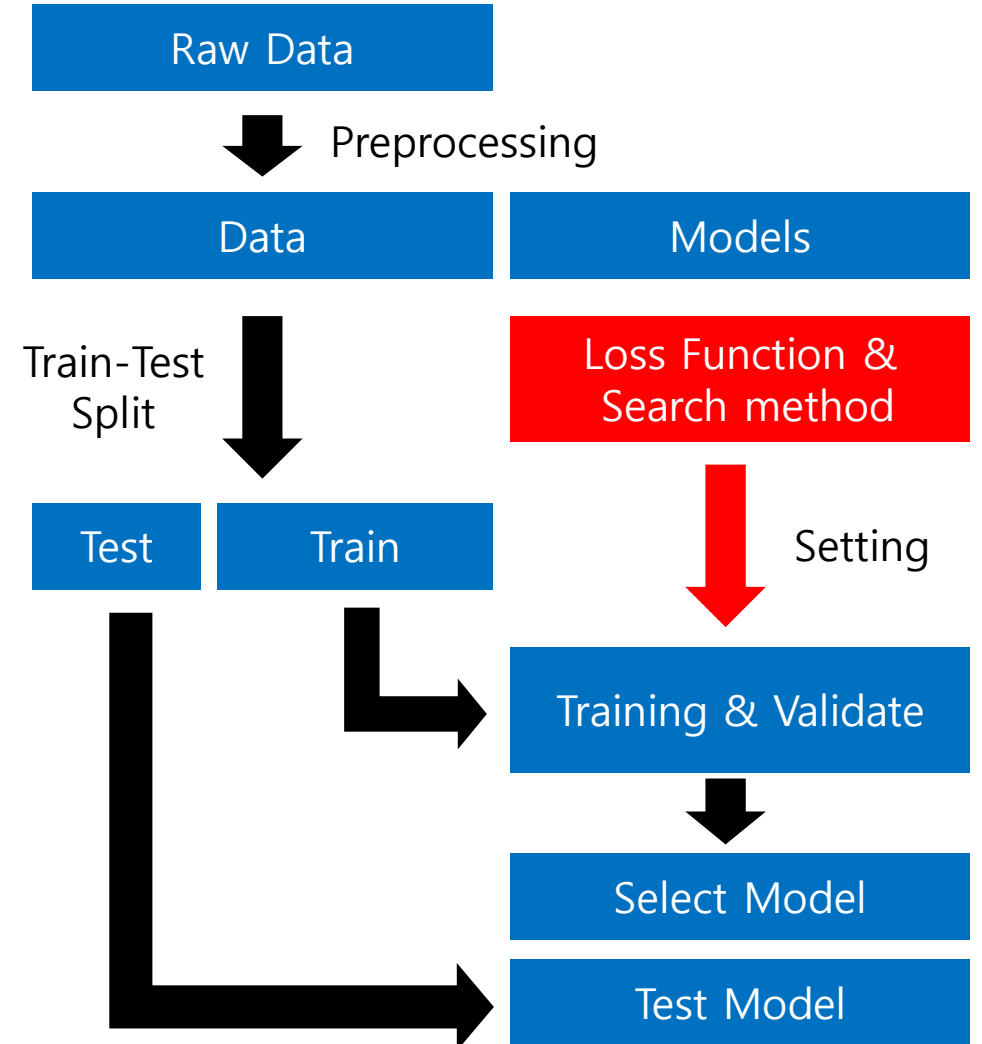
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 모델 학습 과정

1. 데이터 수집 및 분석/전처리
2. Training/Test Set로 데이터 분할
3. 모델 후보군 설정
- 4. 손실 함수 및 탐색 방법 설정**
5. Training Set를 통한 모델 학습
(* 필요시 Validation Set을 통한 모델 성능 비교)
6. Test Set을 통한 성능/과적합 확인
7. (1-6) 반복



Model Training

Model Training

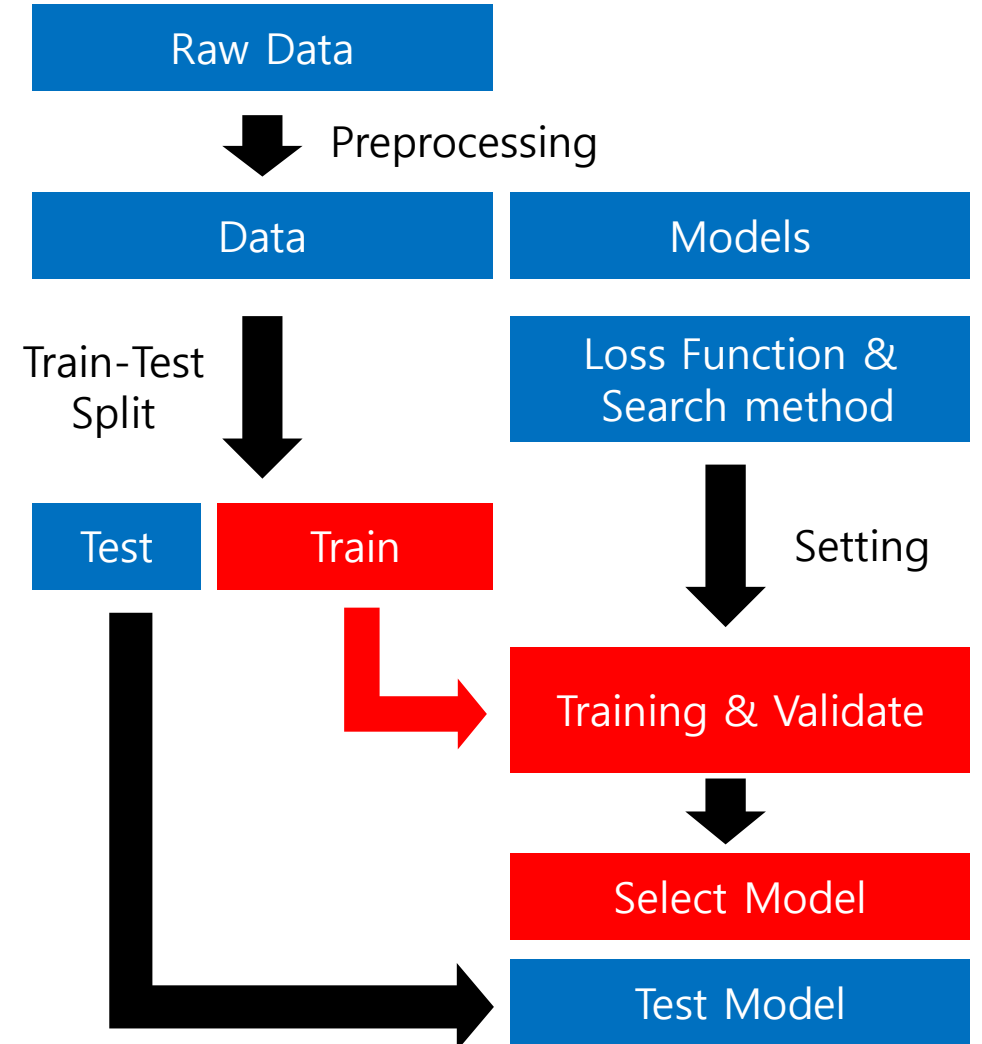
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 모델 학습 과정

1. 데이터 수집 및 분석/전처리
2. Training/Test Set로 데이터 분할
3. 모델 후보군 설정
4. 손실 함수 및 탐색 방법 설정
- 5. Training Set를 통한 모델 학습**
(* 필요시 Validation Set을 통한 모델 성능 비교)
6. Test Set을 통한 성능/과적합 확인
7. (1-6) 반복



Model Training

Model Training

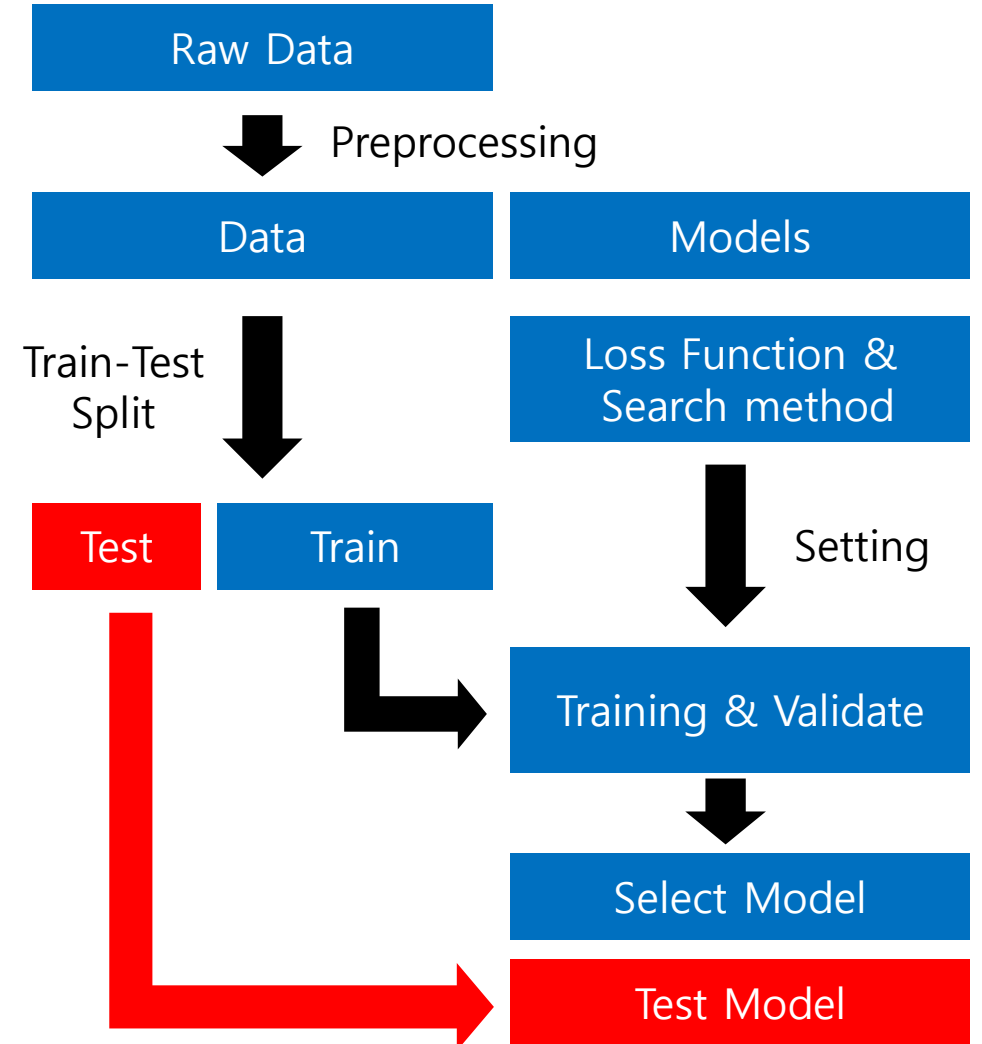
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 모델 학습 과정

1. 데이터 수집 및 분석/전처리
2. Training/Test Set로 데이터 분할
3. 모델 후보군 설정
4. 손실 함수 및 탐색 방법 설정
5. Training Set를 통한 모델 학습
(* 필요시 Validation Set을 통한 모델 성능 비교)
- 6. Test Set을 통한 성능/과적합 확인**
7. (1-6) 반복



Model Training

Model Training

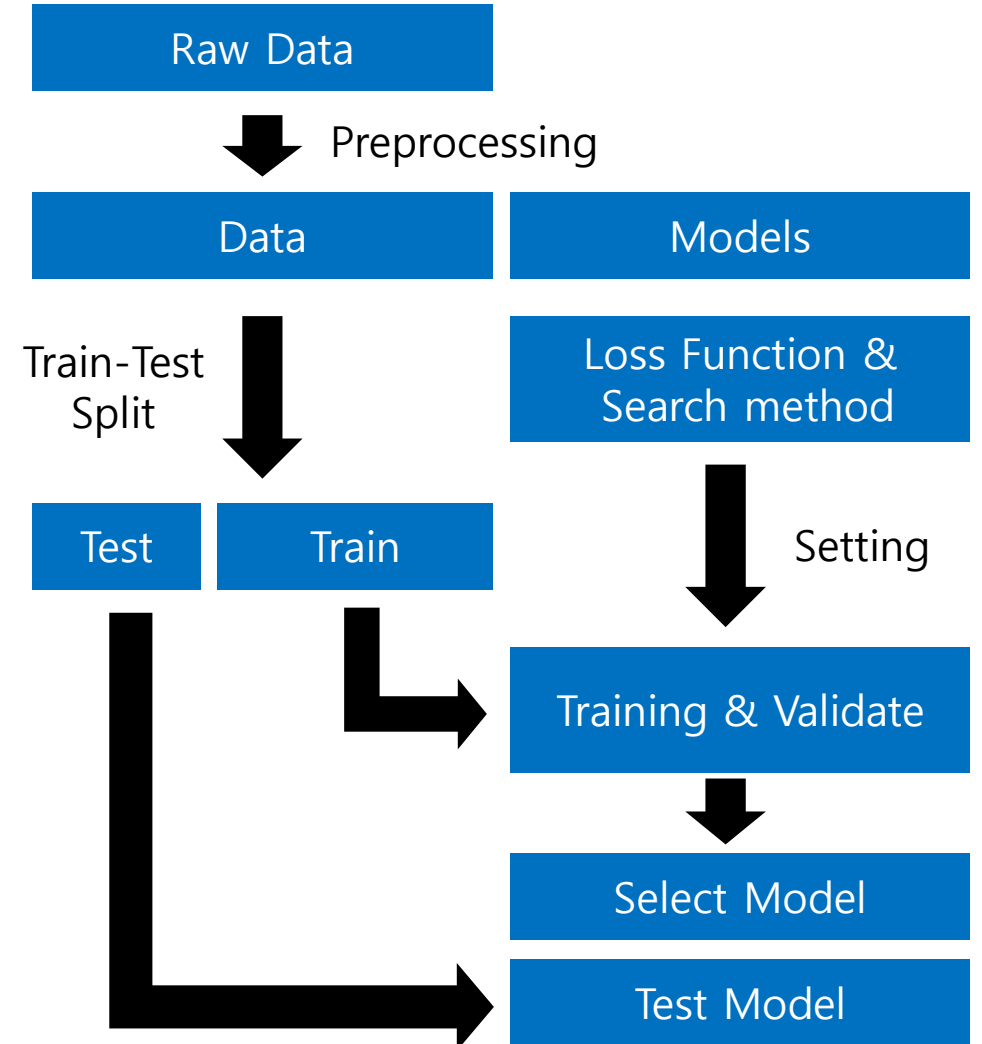
Linear
Regression

Neural Network

Logistic
Regression

■ 교사 학습의 모델 학습 과정

1. 데이터 수집 및 분석/전처리
2. Training/Test Set로 데이터 분할
3. 모델 후보군 설정
4. 손실 함수 및 탐색 방법 설정
5. Training Set를 통한 모델 학습
(* 필요시 Validation Set을 통한 모델 성능 비교)
6. Test Set을 통한 성능/과적합 확인
7. (1-6) 반복

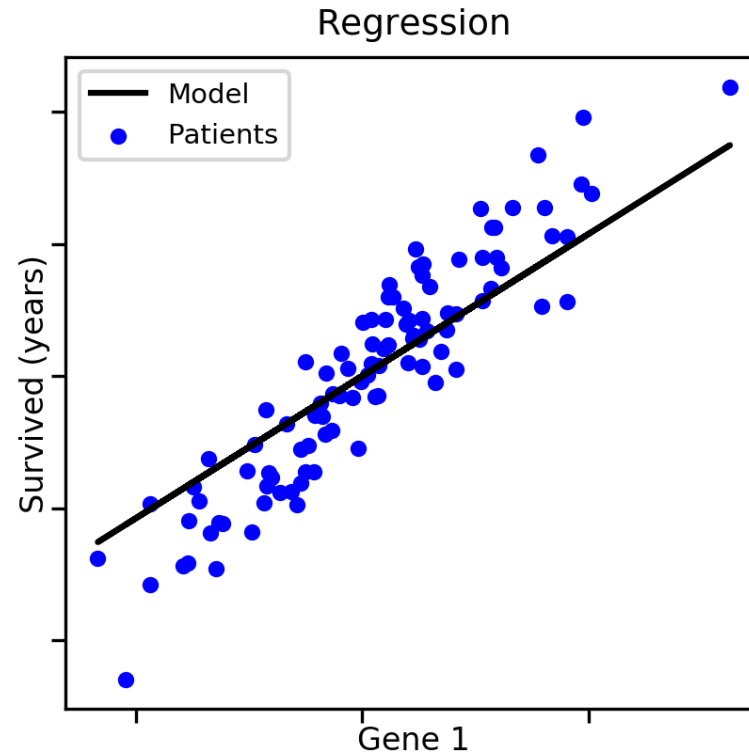


Model Training

Model Training

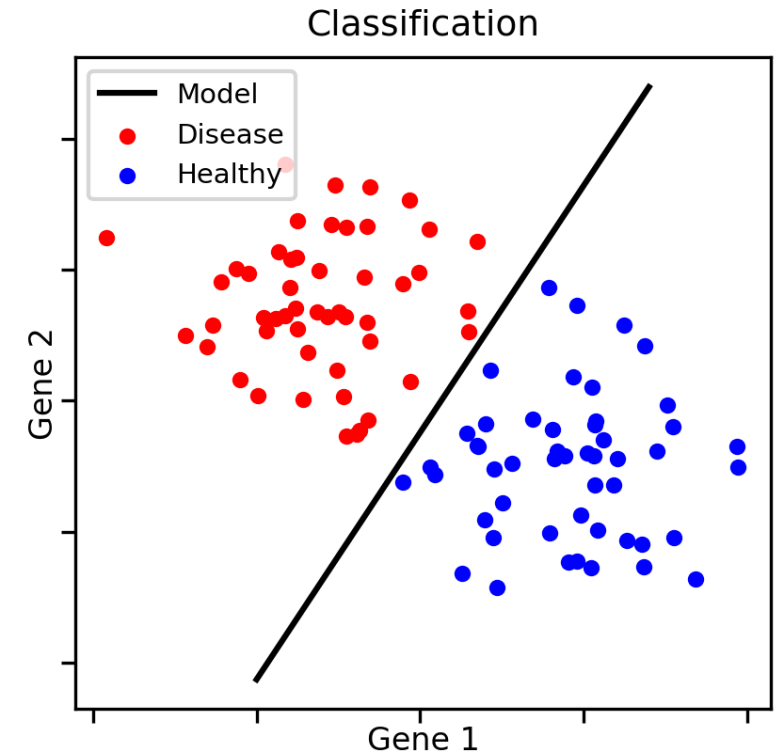
■ 교사 학습의 분류

- 회귀(Regression) : 연속적인 값을 예측
 - EX) 집의 가격 예측



<https://aldro61.github.io/microbiome-summer-school-2017/sections/basics/>

- 분류(Classification) : 이산적인 값을 예측
 - EX) 암 발병 예측



<https://aldro61.github.io/microbiome-summer-school-2017/sections/basics/>

Linear
Regression

Logistic
Regression

Neural Network

Model Training

Linear
Regression

Neural Network

Logistic
Regression

2. Linear Regression

Linear Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 교사 학습의 분류
 - 할아버지의 재산 + 어머니의 연락처 수 \rightarrow 아이의 성적

Linear Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

- **교사 학습의 분류**
 - 할아버지의 재산 + 어머니의 연락처 수 \rightarrow 아이의 성적
- 정답이 존재하므로 **교사 학습(Supervised Learning)**에 해당
 - Input(X) : 할아버지의 재산, 어머니의 연락처 수
 - Output(Y) : 아이의 성적

Linear Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

- **교사 학습의 분류**
 - 할아버지의 재산 + 어머니의 연락처 수 \rightarrow 아이의 성적
- 정답이 존재하므로 **교사 학습(Supervised Learning)**에 해당
 - Input(X) : 할아버지의 재산, 어머니의 연락처 수
 - Output(Y) : 아이의 성적
- 또한 목표 변수가 연속이므로 **회귀(Regression)**에 해당

Linear Regression

Model Training

Linear Regression

Neural Network

Logistic Regression

- **교사 학습의 분류**
 - 할아버지의 재산 + 어머니의 연락처 수 \rightarrow 아이의 성적
- 정답이 존재하므로 **교사 학습(Supervised Learning)**에 해당
 - Input(X) : 할아버지의 재산, 어머니의 연락처 수
 - Output(Y) : 아이의 성적
- 또한 목표 변수가 연속이므로 **회귀(Regression)**에 해당
- 회귀 모델 :
 - **Linear Regression**
 - Multivariate Linear Regression
 - Non-Linear Regression

Linear Regression

Model Training

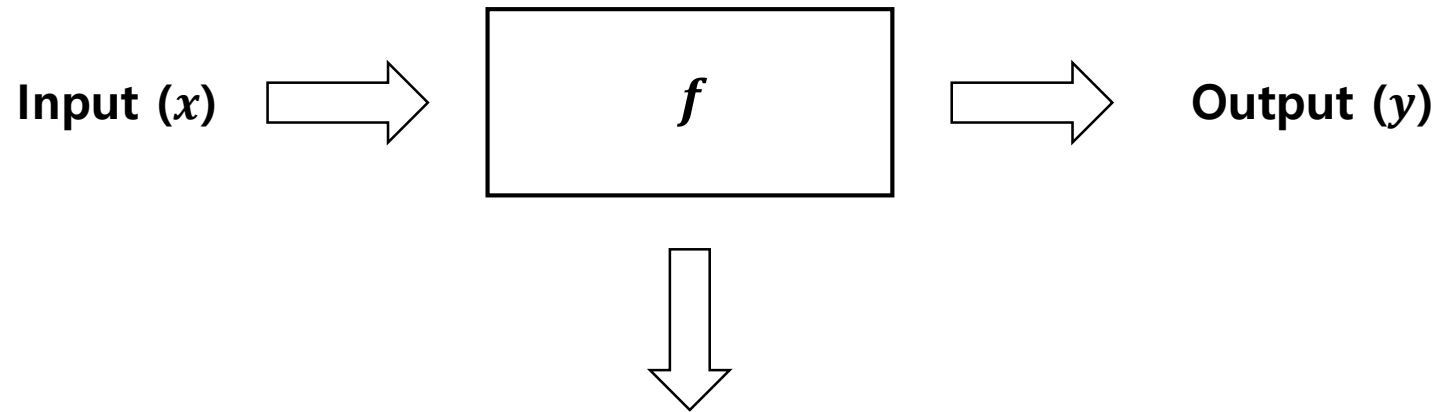
Linear
Regression

Neural Network

Logistic
Regression

■ 선형 회귀(Linear Regression)

- 종속 변수 y 와 한 개의 독립 변수 x 와의 선형 관계를 모델링하는 회귀 분석 기법



$$f(x) = wx + b$$

w : 기울기

b : 절편

Linear Regression

Model Training

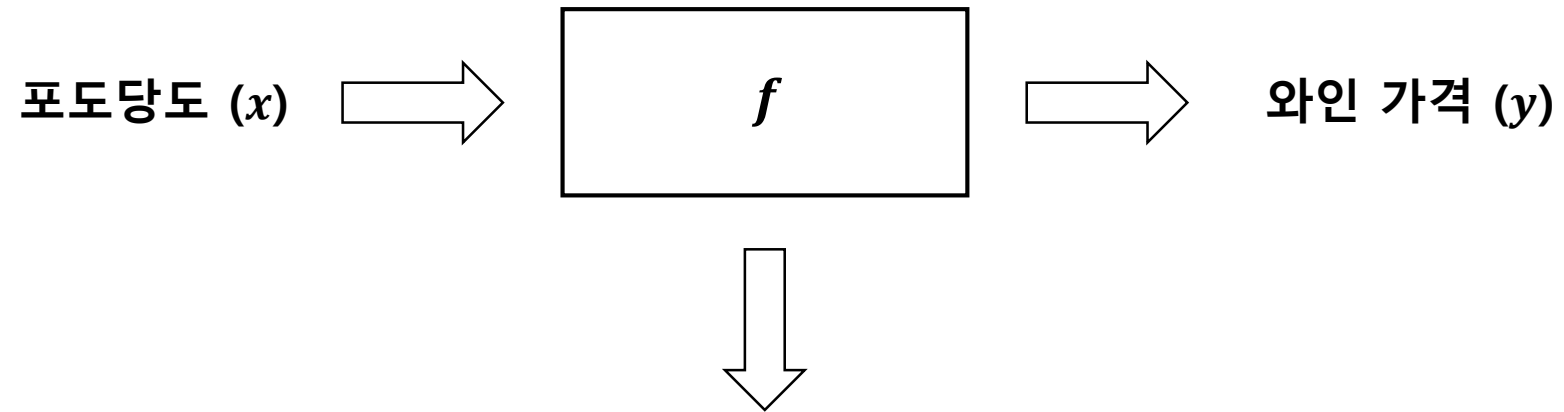
Linear
Regression

Neural Network

Logistic
Regression

■ 선형 회귀(Linear Regression)

- 종속 변수 y 와 한 개의 독립 변수 x 와의 선형 관계를 모델링하는 회귀 분석 기법



$$f(x) = 100x + 50$$

포도당도가 1 증가하면 와인가격은 100 증가

$$\text{와인가격} = \text{포도당도} * 100 + 50$$

Linear Regression

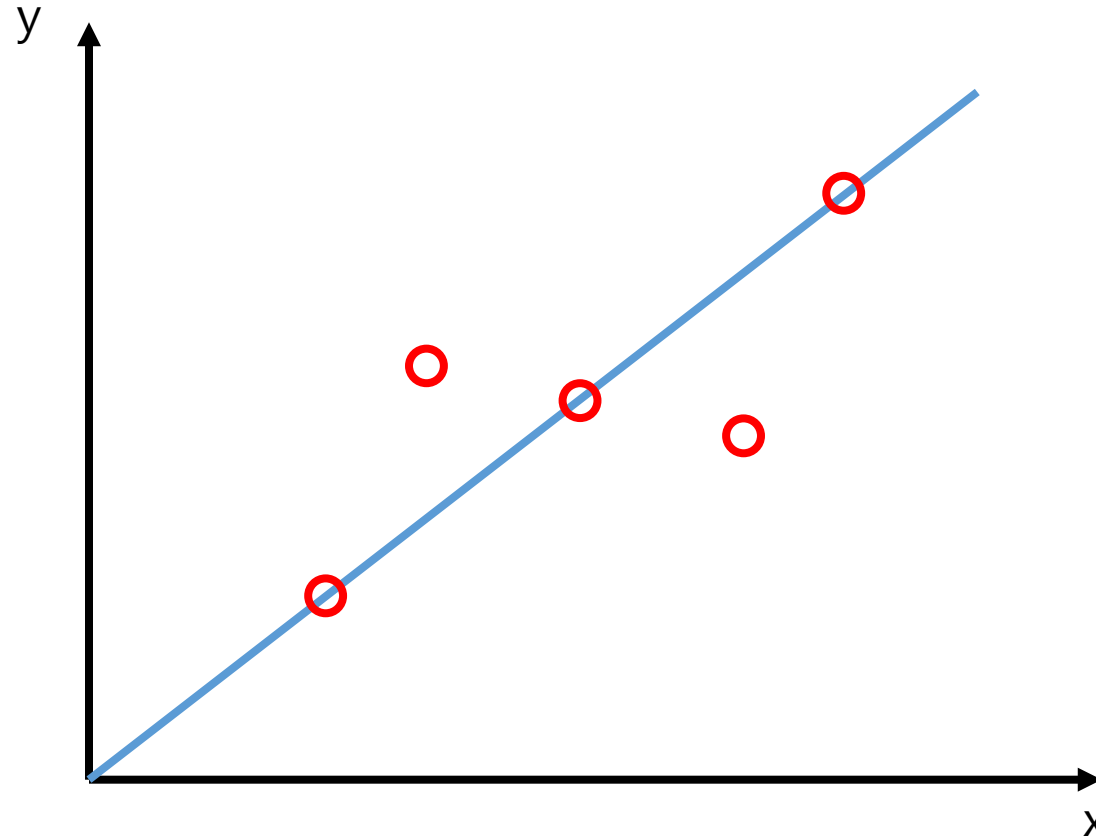
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 선형 회귀(Linear Regression)
 - 주어진 데이터에 가장 적합한 직선을 도출



Linear Regression

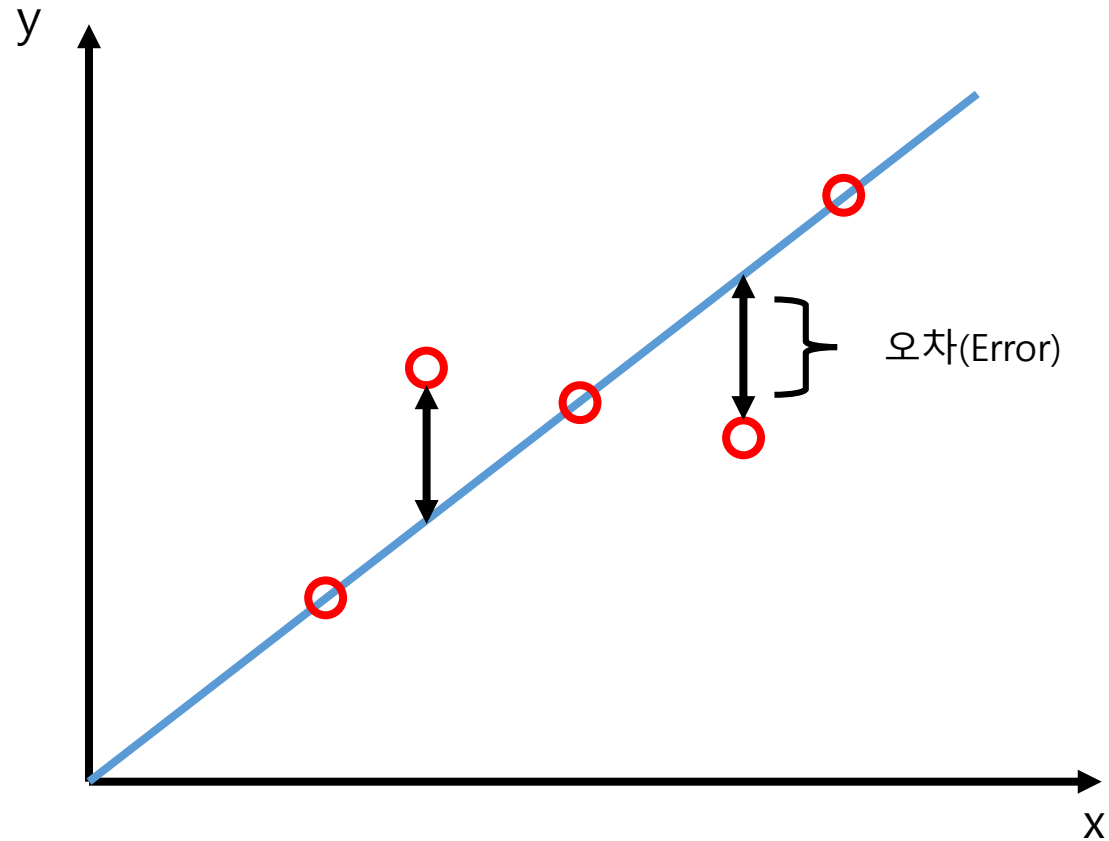
Model Training

Linear Regression

Neural Network

Logistic Regression

- 선형 회귀(Linear Regression)
 - 얼마나 정확한 모델인가?



Linear Regression

Model Training

Linear Regression

Neural Network

Logistic Regression

- 선형 회귀(Linear Regression)의 평가 척도
 - 평균 제곱 오차(Mean Squared Error, MSE) : 선형 회귀에서의 손실 함수
 - 오차의 제곱의 합의 평균
 - $$MSE = \frac{\sum err^2}{n} = \frac{\sum (y - wx - b)^2}{n}$$
 - 결정 계수(Coefficient of Determination, R^2)
 - 학습 모형으로 설명 가능한 부분의 비율
 - $$0 \leq R^2 = \frac{SSR}{SST} \leq 1$$
 - MAPE(Mean Absolute Percentage Error), MAE(Mean Absolute Error)...

Linear Regression

Model Training

Linear Regression

Neural Network

Logistic Regression

- 선형 회귀(Linear Regression)의 평가 척도
 - 평균 제곱 오차(Mean Squared Error, MSE) : 선형 회귀에서의 손실 함수
 - 오차의 제곱의 합의 평균

$$MSE = \frac{\sum err^2}{n} = \frac{\sum (y - wx - b)^2}{n}$$

| 신용 점수 | 민수 | 지예 | 철수 |
|-------|----|----|----|
| 실제값 | 10 | 9 | 8 |
| 예측값 | 9 | 7 | 10 |

Linear Regression

Model Training

Linear Regression

Neural Network

Logistic Regression

- 선형 회귀(Linear Regression)의 평가 척도
 - 평균 제곱 오차(Mean Squared Error, MSE) : 선형 회귀에서의 손실 함수
 - 오차의 제곱의 합의 평균

$$MSE = \frac{\sum err^2}{n} = \frac{\sum (y - wx - b)^2}{n}$$

| 신용 점수 | 민수 | 지예 | 철수 |
|-------|----|----|----|
| 실제값 | 10 | 9 | 8 |
| 예측값 | 9 | 7 | 10 |
| 오차 | 1 | 2 | -2 |

Linear Regression

Model Training

Linear Regression

Neural Network

Logistic Regression

- 선형 회귀(Linear Regression)의 평가 척도
 - 평균 제곱 오차(Mean Squared Error, MSE) : 선형 회귀에서의 손실 함수
 - 오차의 제곱의 합의 평균

$$MSE = \frac{\sum err^2}{n} = \frac{\sum (y - wx - b)^2}{n}$$

| 신용 점수 | 민수 | 지예 | 철수 |
|-------|----|----|----|
| 실제값 | 10 | 9 | 8 |
| 예측값 | 9 | 7 | 10 |
| 오차 | 1 | 2 | -2 |

$$\frac{1^2 + 2^2 + (-2)^2}{3} = 3$$

- 어떻게 줄일 것인가?
 - 정규방정식(Normal Equation)
 - 손실 함수를 최소화하는 해석적인 방법
 - 선형 회귀의 경우 MSE를 최소화하는 식이 알려져 있음 (정규방정식 존재)
 - $y = XA \rightarrow A = (X^T X)^{-1} X^T y$
 - 경사하강법(Gradient Descent)
 - 손실 함수의 기울기(변화)를 통해 손실 함수를 줄이는 방법
 - 정규방정식이 존재하지 않아도 손실 함수만 있다면 가능

Linear Regression

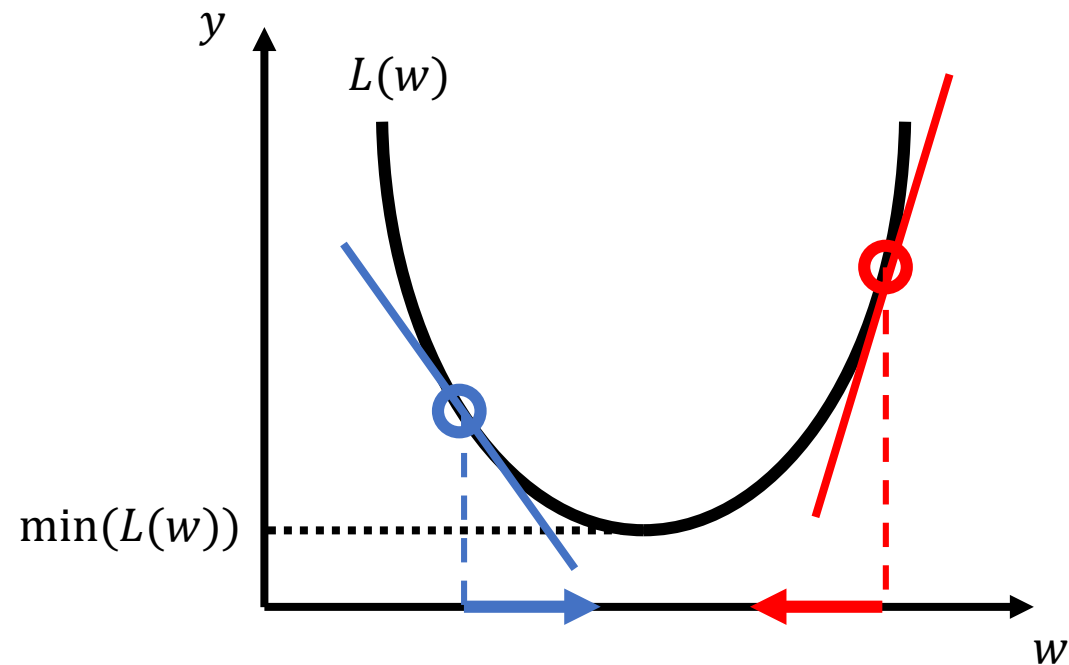
Model Training

Linear Regression

Neural Network

Logistic Regression

- 어떻게 줄일 것인가?
 - 경사하강법(Gradient Descent)
 - 초기점(Initial Point)에서 출발
 - 초기점에서의 기울기를 구하여 기울기의 반대방향으로 움직임
 - 움직이는 정도는 학습계수(Learning Rate)를 통해 조절



Linear Regression

Model Training

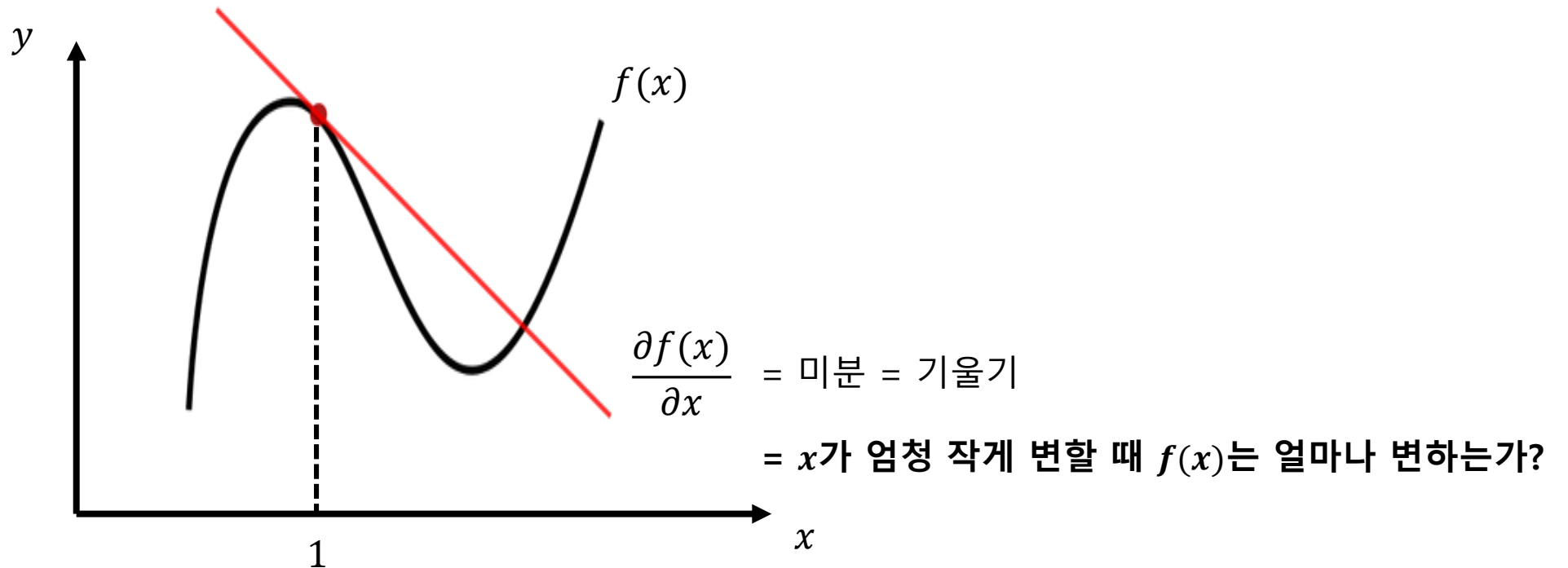
Linear Regression

Neural Network

Logistic Regression

■ 어떻게 줄일 것인가?

- 경사하강법(Gradient Descent)
 - 초기점(Initial Point)에서 출발
 - 초기점에서의 기울기(=미분)를 구하여 기울기의 반대방향으로 움직임
 - 움직이는 정도는 학습계수(Learning Rate)를 통해 조절



Linear Regression

Model Training

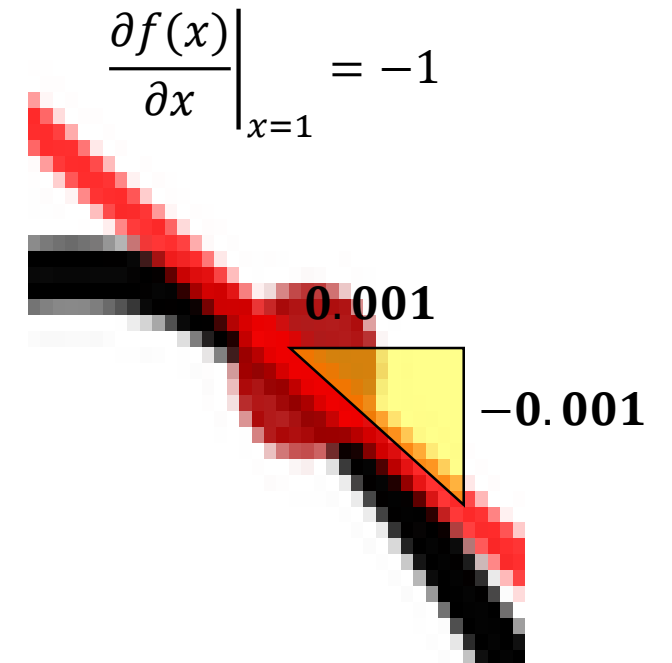
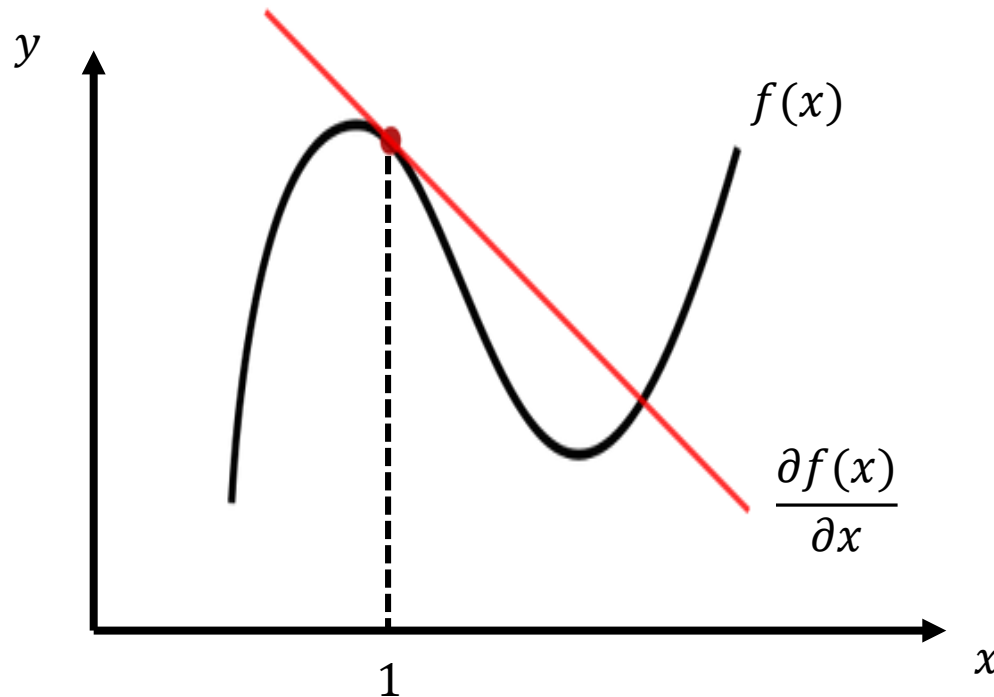
Linear Regression

Neural Network

Logistic Regression

■ 어떻게 줄일 것인가?

- 경사하강법(Gradient Descent)
 - 초기점(Initial Point)에서 출발
 - 초기점에서의 기울기(=미분)를 구하여 기울기의 반대방향으로 움직임
 - 움직이는 정도는 학습계수(Learning Rate)를 통해 조절



Linear Regression

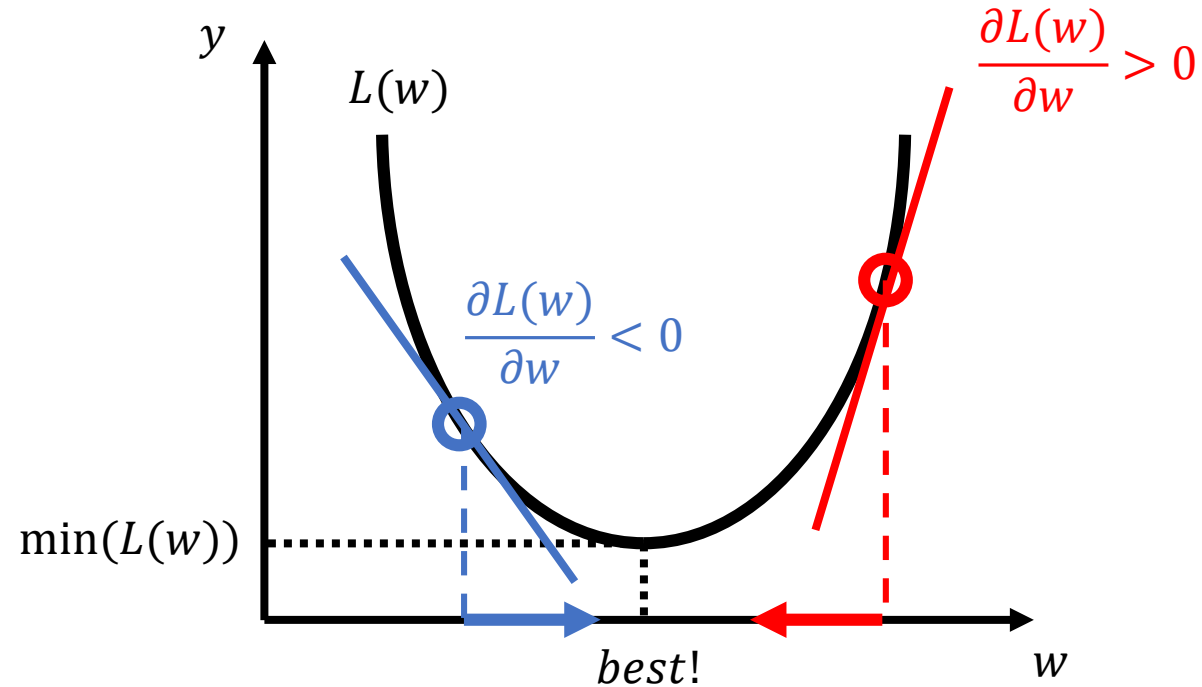
Model Training

Linear Regression

Neural Network

Logistic Regression

- 어떻게 줄일 것인가?
 - 경사하강법(Gradient Descent)
 - 초기점(Initial Point)에서 출발
 - 초기점에서의 기울기를 구하여 기울기의 반대방향으로 움직임
 - 움직이는 정도는 학습계수(Learning Rate)를 통해 조절



Linear Regression

Model Training

Linear Regression

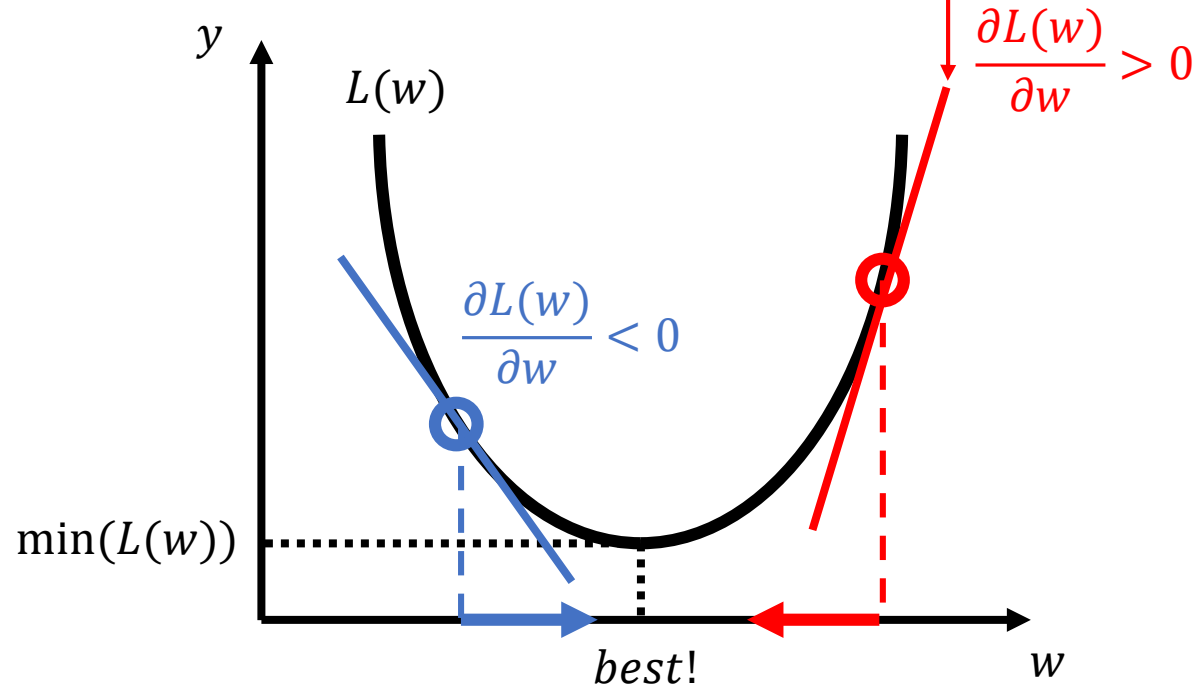
Neural Network

Logistic Regression

- 어떻게 줄일 것인가?
- 경사하강법(Gradient Descent)

- $$w' := w - \alpha * \frac{\partial}{\partial w} L(w)$$

- α : 학습계수



Linear Regression

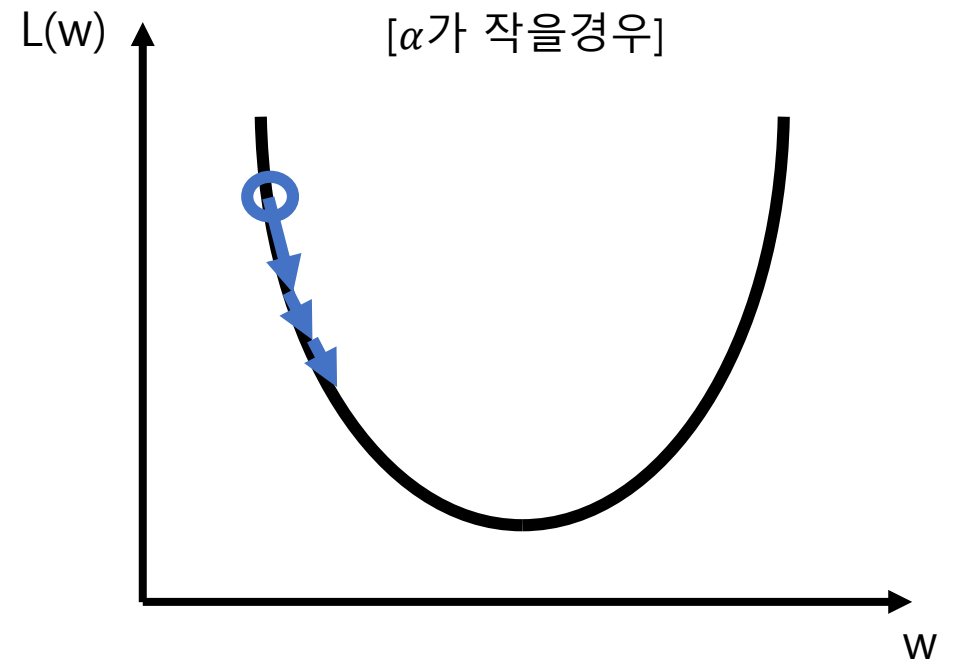
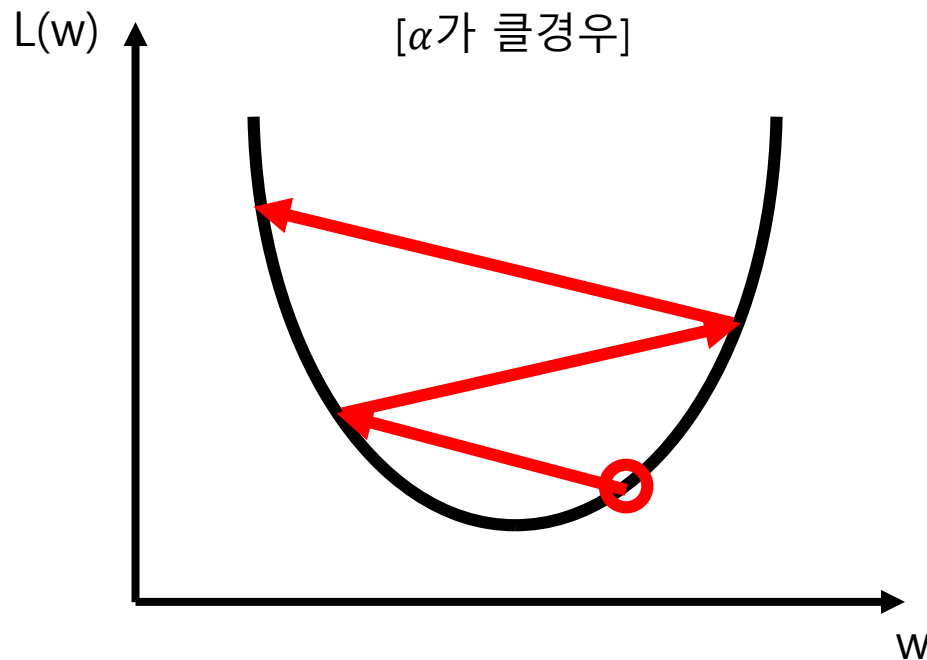
Model Training

Linear Regression

Neural Network

Logistic Regression

- 어떻게 줄일 것인가?
 - 경사하강법(Gradient Descent)
 - $w' := w - \alpha \frac{\partial}{\partial w} L(w)$
 - α : 학습계수



Linear Regression

Model Training

Linear Regression

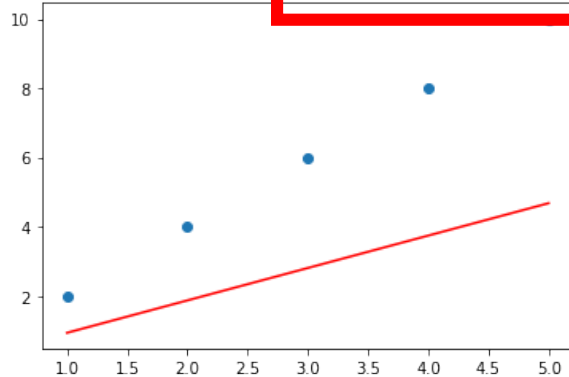
Neural Network

Logistic Regression

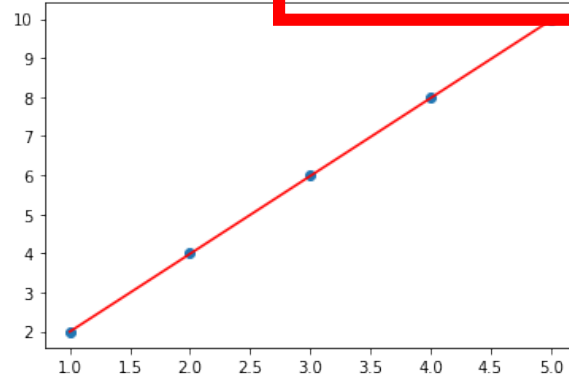
■ 선형회귀에서의 경사하강법

- $y = wx$
- $MSE = \frac{\sum err^2}{n} = \frac{\sum_x (wx - y)^2}{n}$
- $grad = \frac{\partial}{\partial w} MSE = \frac{1}{n} \frac{\partial}{\partial w} \sum_x (wx - y)^2 = \frac{\sum_x 2(wx - y)x}{n}$
- $w' := w - \alpha * grad$

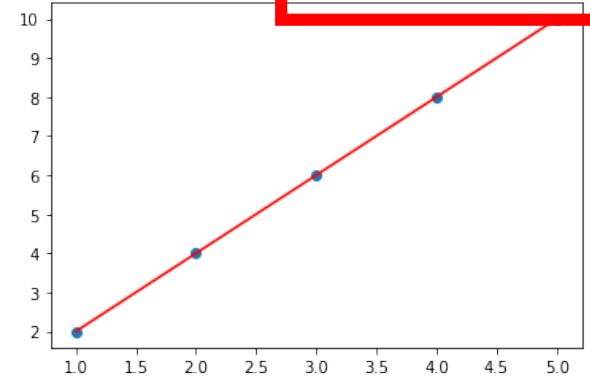
step 0 : cost=12.4594, w=1.1699, grad=-23.4139



step 20 : cost=0.0000, w=1.9942, grad=-0.1627



step 80 : cost=0.0000, w=2.0000, grad=-0.0000



Linear Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

다변량 선형 회귀(Multivariate Linear Regression)

- 종속 변수 y 와 두 개 이상의 독립 변수 x 와의 선형 관계를 모델링하는 회귀 분석 기법

$$f(x) = w_0x_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n \quad (x_0 = 1)$$

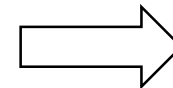
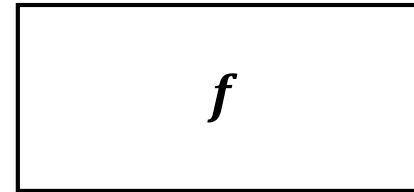
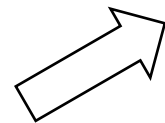
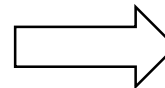
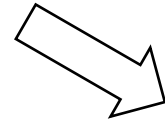
Input (x_1)

⋮

Input (x_m)

⋮

Input (x_n)



Output (y)

Linear Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

다변량 선형 회귀(Multivariate Linear Regression)

$$\text{신용도} = 50 * \text{결제 금액} - 100 * \text{연체 횟수} - 5 * \text{통장 개수}$$

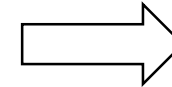
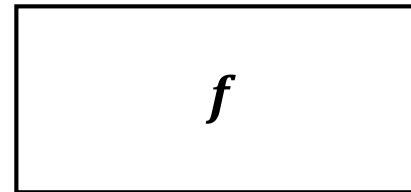
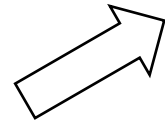
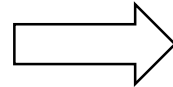
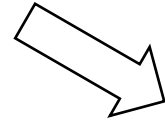
결제 금액 (x_1)

⋮

연체 횟수 (x_m)

⋮

통장 개수 (x_n)



신용도 (y)

- 다변량 선형 회귀(Multivariate Linear Regression)

$$f(x) = w_0x_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n \quad (x_0 = 1)$$

$$= [w_0 \quad w_1 \quad \cdots \quad w_n] \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = w^T x$$

Linear Regression

Model Training

Linear Regression

Neural Network

Logistic Regression

- 다변량 선형 회귀에서의 경사하강법

- $y = w^T x$ (w, x is vector)

- $$MSE = \frac{\sum err^2}{n} = \frac{\sum_x (w^T x - y)^2}{n}$$

$$w = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix} \quad x = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- $$grad = \frac{\partial}{\partial w} MSE = \frac{1}{n} \frac{\partial}{\partial w} \sum_x (w^T x - y)^2 = \frac{\sum_x 2(w^T x - y)x^T}{n}$$

- $$w' := w - \alpha \frac{\partial}{\partial w} L(w) = w - \alpha * grad$$

Model Training

**Linear
Regression**

Neural Network

**Logistic
Regression**

3. Neural Network

Model Training

Linear Regression

Neural Network

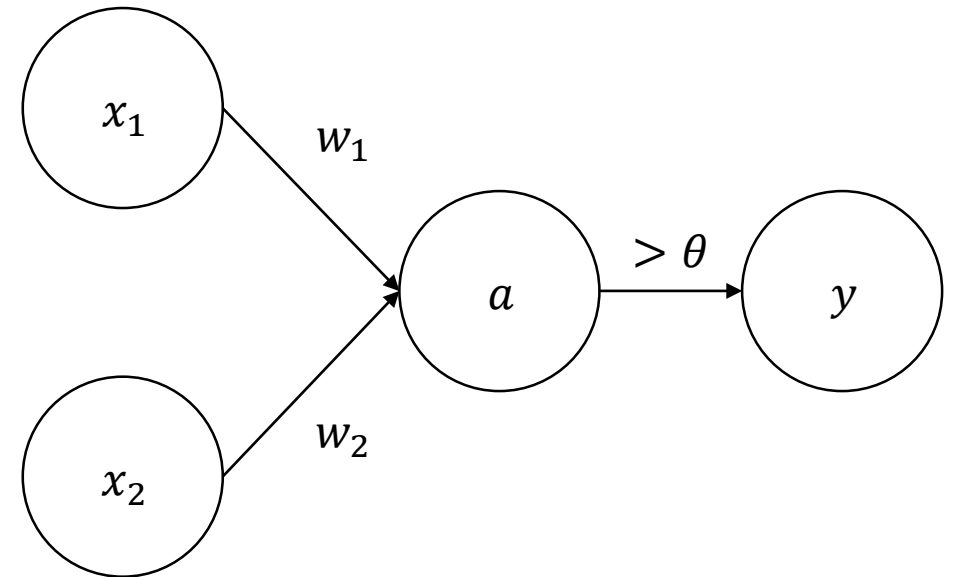
Logistic Regression

■ 선형 회귀(Linear Regression)와 신경망(Neural Network)

- 선형 회귀식과 퍼셉트론은 매우 닮아 있음

■ [Review] : 퍼셉트론(Perceptron)

- 1957년 코넬 항공 연구소의 프랑크 로젠블라트에 의해 고안
- 가장 간단한 형태의 Feed-Forward 네트워크
- 다수의 신호를 입력받아 하나의 신호를 출력
 - x_1, x_2 : 입력 신호
 - w_1, w_2 : 가중치
 - y : 출력 신호
- 일정 임계값(θ)을 넘어설 때만 1을 출력
 - $$y = \begin{cases} 0 & (w_1 * x_1 + w_2 * x_2 \leq \theta) \\ 1 & (w_1 * x_1 + w_2 * x_2 > \theta) \end{cases}$$
 - "선형 분리"



Neural Network

Model Training

Linear
Regression

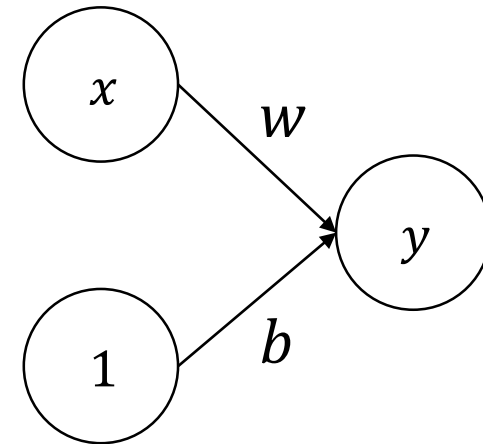
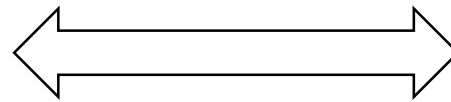
Neural Network

Logistic
Regression

■ 선형 회귀(Linear Regression)와 신경망(Neural Network)

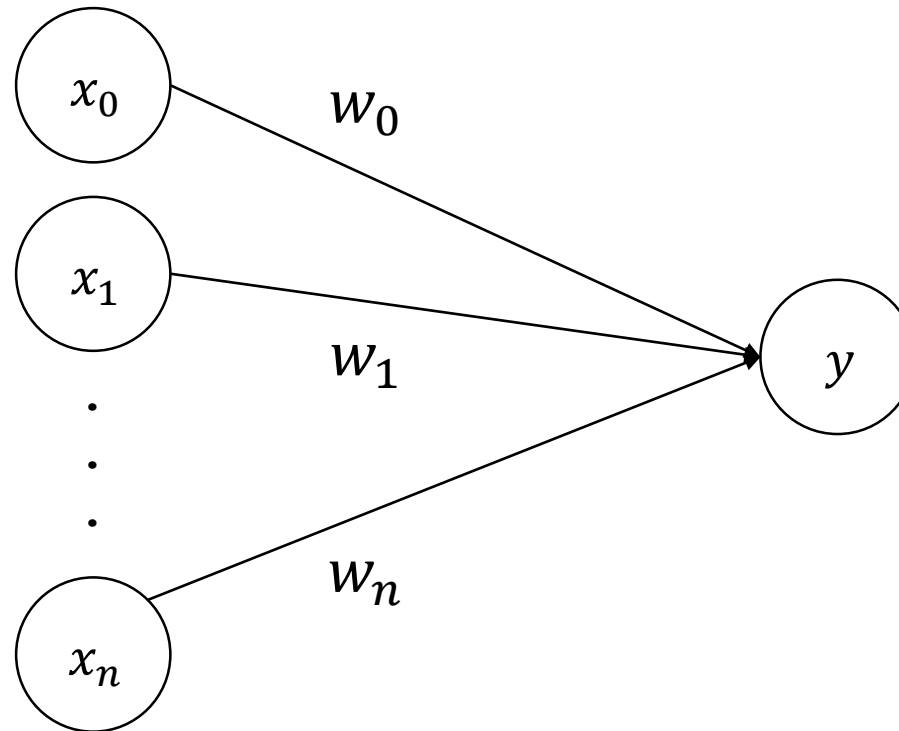
- 선형 회귀식과 퍼셉트론은 매우 닮아 있음
- 퍼셉트론 : 다수의 신호를 입력받아 하나의 신호를 출력
 - $x, 1$: 입력 신호
 - w, b : 가중치
 - y : 출력 신호

$$y = wx + b$$



- 선형 회귀(Linear Regression)와 신경망(Neural Network)
 - 다변량 선형 회귀식 역시 한 층의 퍼셉트론으로도 구현 가능

$$f(x) = w_0x_0 + w_1x_1 + \cdots + w_nx_n \quad (x_0 = 1)$$



Neural Network

Model Training

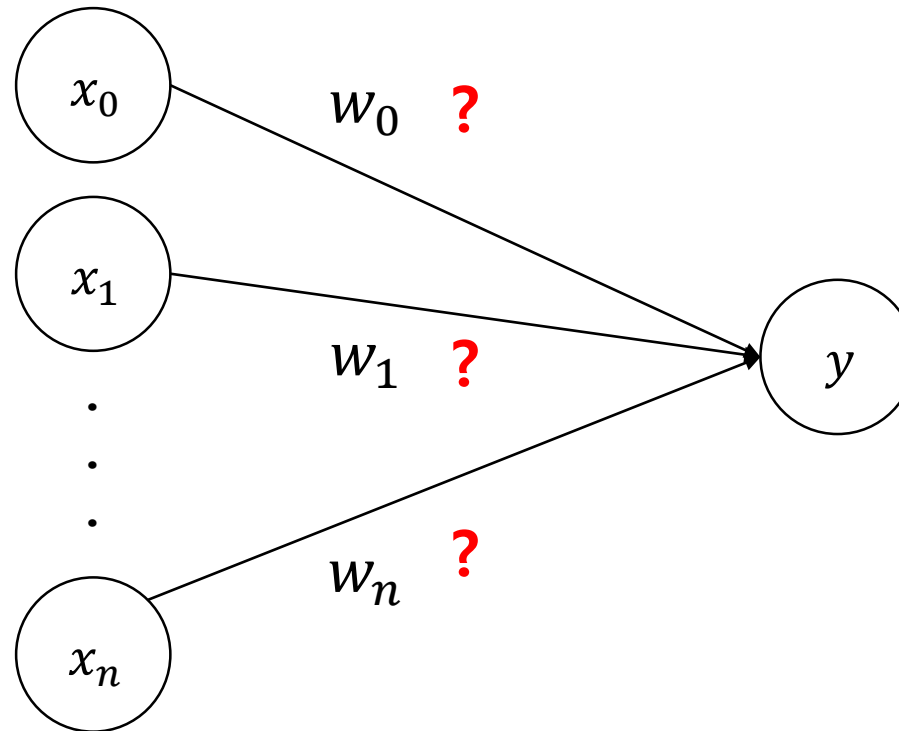
Linear
Regression

Neural Network

Logistic
Regression

- 선형 회귀(Linear Regression)와 신경망(Neural Network)
 - 그렇다면 학습은 어떻게?

$$f(x) = w_0x_0 + w_1x_1 + \cdots + w_nx_n \quad (x_0 = 1)$$



Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 신경망(Neural Network)에서의 학습
 - 문제1) 철수는 1000원짜리 과자를 3개 샀다. 이 때 철수의 지불액은? (단, 소비세 10% 부과)

Model Training

Linear
Regression

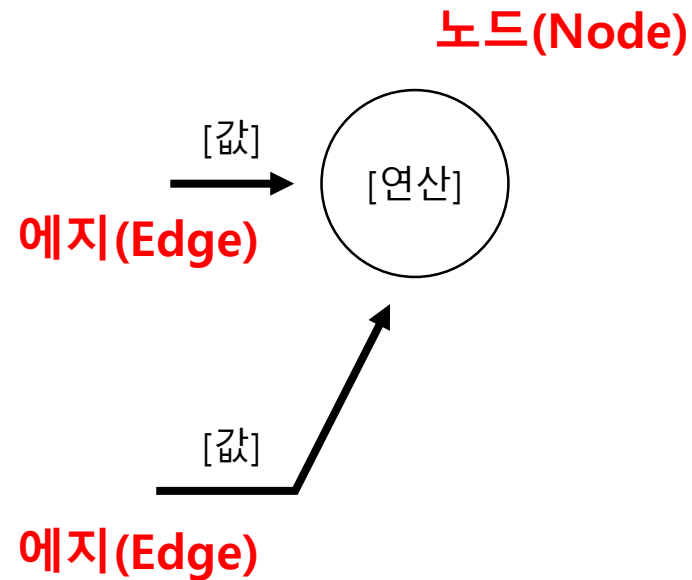
Neural Network

Logistic
Regression

- 신경망(Neural Network)에서의 학습
 - 문제1) 철수는 1000원짜리 과자를 3개 샀다. 이 때 철수의 지불액은? (단, 소비세 10% 부과)
 - $= (1000 \times 3) \times 1.1 = 3300$
 - = 간단한 단순 선형식

■ 신경망(Neural Network)에서의 학습

- 문제 1) 철수는 1000원짜리 과자를 3개 샀다. 이 때 철수의 지불액은? (단, 소비세 10% 부과)
- 계산그래프(Computational Graph) : 노드(Node)와 에지(Edge)로 구성



Model Training

Linear Regression

Neural Network

Logistic Regression

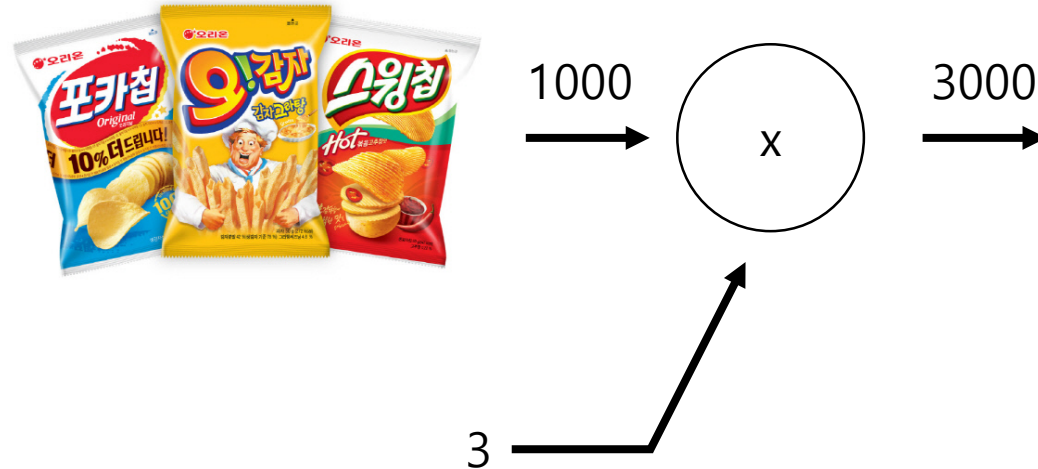
■ 신경망(Neural Network)에서의 학습

- 문제 1) 철수는 1000원짜리 과자를 3개 샀다. 이 때 철수의 지불액은? (단, 소비세 10% 부과)
- 계산그래프(Computational Graph) : 노드(Node)와 에지(Edge)로 구성



■ 신경망(Neural Network)에서의 학습

- 문제 1) 철수는 1000원짜리 과자를 3개 샀다. 이 때 철수의 지불액은? (단, 소비세 10% 부과)
- 계산그래프(Computational Graph) : 노드(Node)와 에지(Edge)로 구성



Neural Network

Model Training

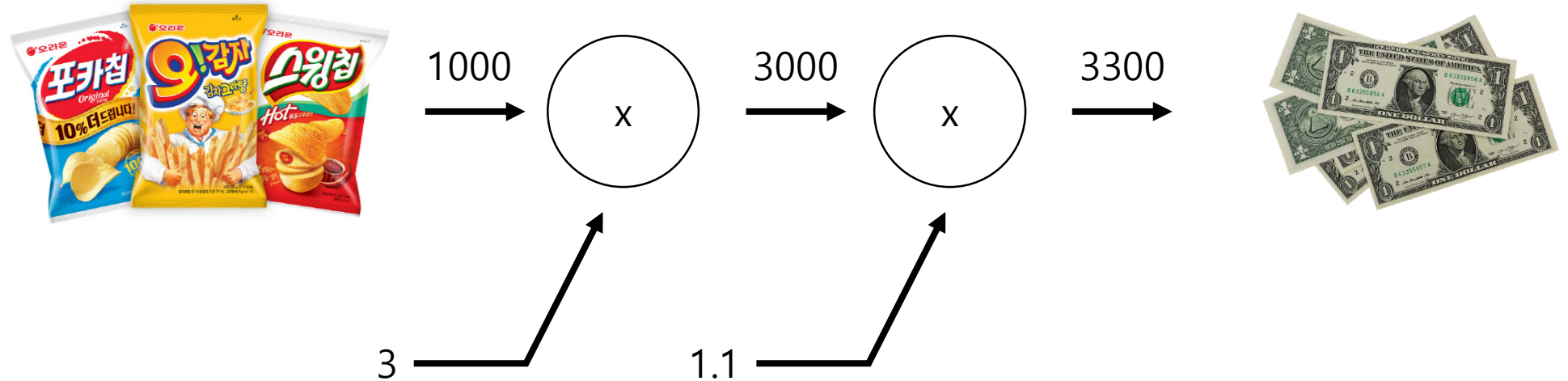
Linear
Regression

Neural Network

Logistic
Regression

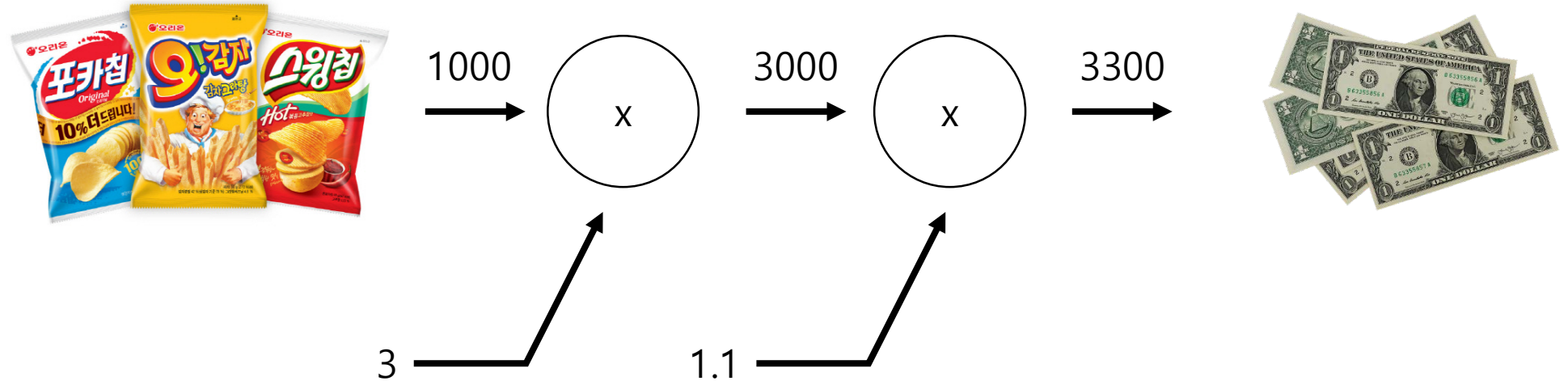
■ 신경망(Neural Network)에서의 학습

- 문제 1) 철수는 1000원짜리 과자를 3개 샀다. 이 때 철수의 지불액은? (단, 소비세 10% 부과)
- 계산그래프(Computational Graph) : 노드(Node)와 에지(Edge)로 구성



■ 신경망(Neural Network)에서의 학습

- 문제 1) 철수는 1000원짜리 과자를 3개 샀다. 이 때 철수의 지불액은? (단, 소비세 10% 부과)
- 계산그래프(Computational Graph) : 노드(Node)와 에지(Edge)로 구성



- 순전파(Forward Propagation) : 계산을 왼쪽에서 오른쪽으로 진행하는 단계

Neural Network

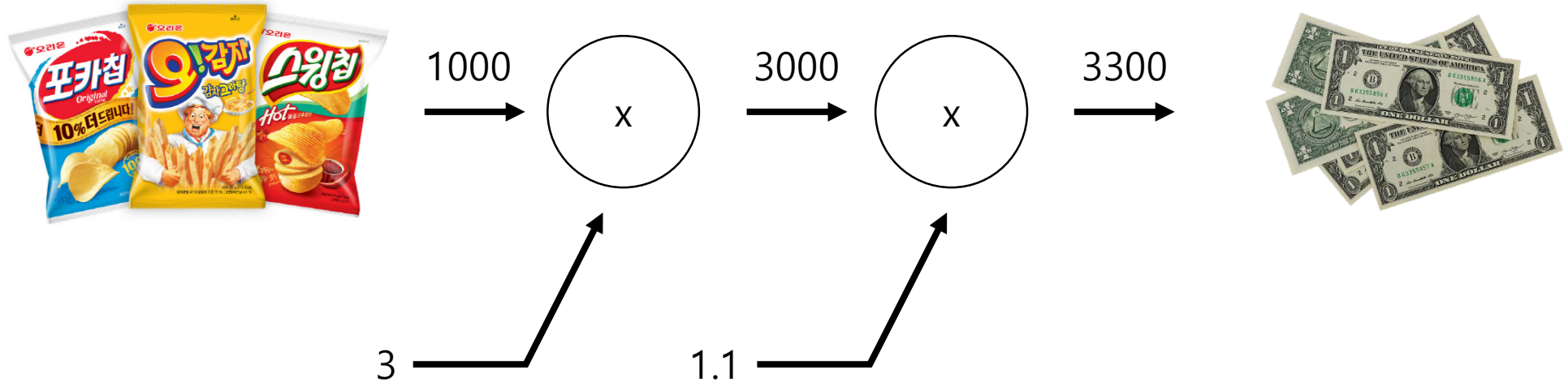
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 신경망(Neural Network)에서의 학습
 - 문제2) 과자를 하나 더 살 때, 지불 금액은 얼마나 증가하는가?



Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 신경망(Neural Network)에서의 학습
 - 문제2) 과자를 하나 더 살 때, 지불 금액은 얼마나 증가하는가?
 - = 과자 수량(x)의 변화 당 지불 금액(y)의 변화량
 - = 1100
 - $= \frac{\partial y}{\partial x}$

Neural Network

Model Training

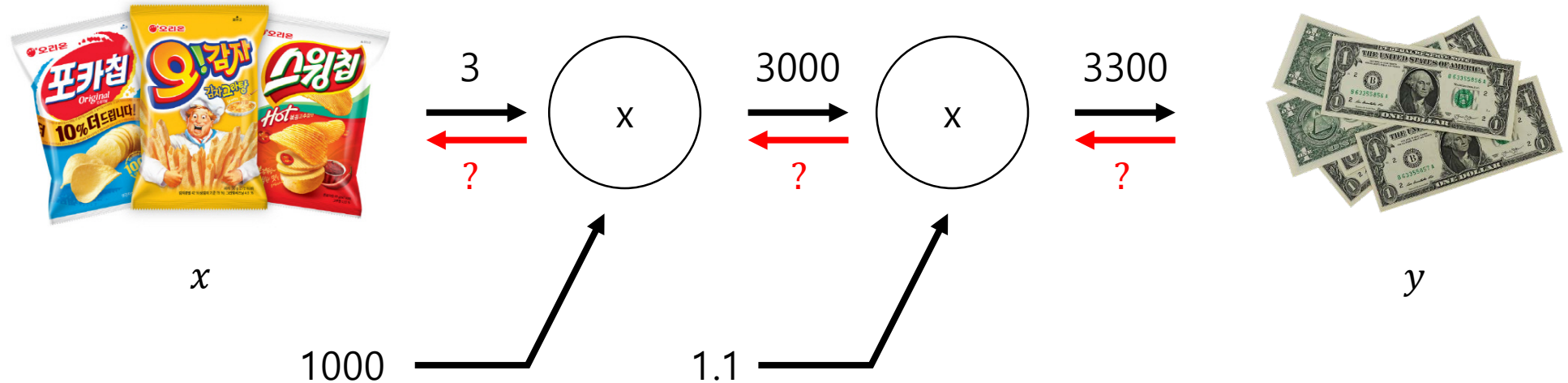
Linear
Regression

Neural Network

Logistic
Regression

■ 신경망(Neural Network)에서의 학습

- 문제2) 과자를 하나 더 살 때, 지불 금액은 얼마나 증가하는가? $\frac{\partial y}{\partial x}$



Neural Network

Model Training

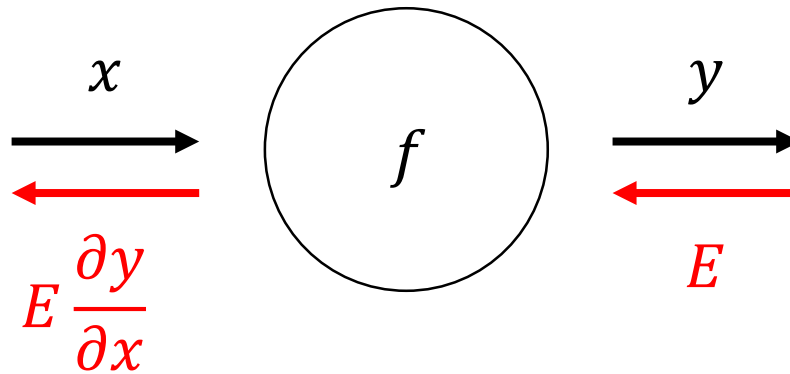
Linear Regression

Neural Network

Logistic Regression

■ 역전파(Backpropagation)

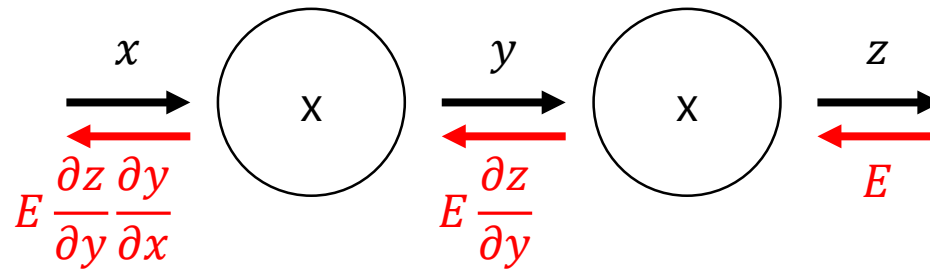
- 국소적 미분을 오른쪽에서 왼쪽으로 전달하는 것
- 이러한 과정을 반복적으로 거치면 원하는 값을 쉽게 구할 수 있음



■ 역전파(Backpropagation)

- 역전파가 가능한 이유 : 연쇄법칙(Chain Rule)

$$\frac{\partial z}{\partial y} \frac{\partial y}{\partial x} = \frac{\partial z}{\cancel{\partial y}} \frac{\cancel{\partial y}}{\partial x} = \frac{\partial z}{\partial x}$$



$$\text{■ } E \frac{\partial z}{\partial y} \frac{\partial y}{\partial x} = E \frac{\partial z}{\partial x}$$

Neural Network

Model Training

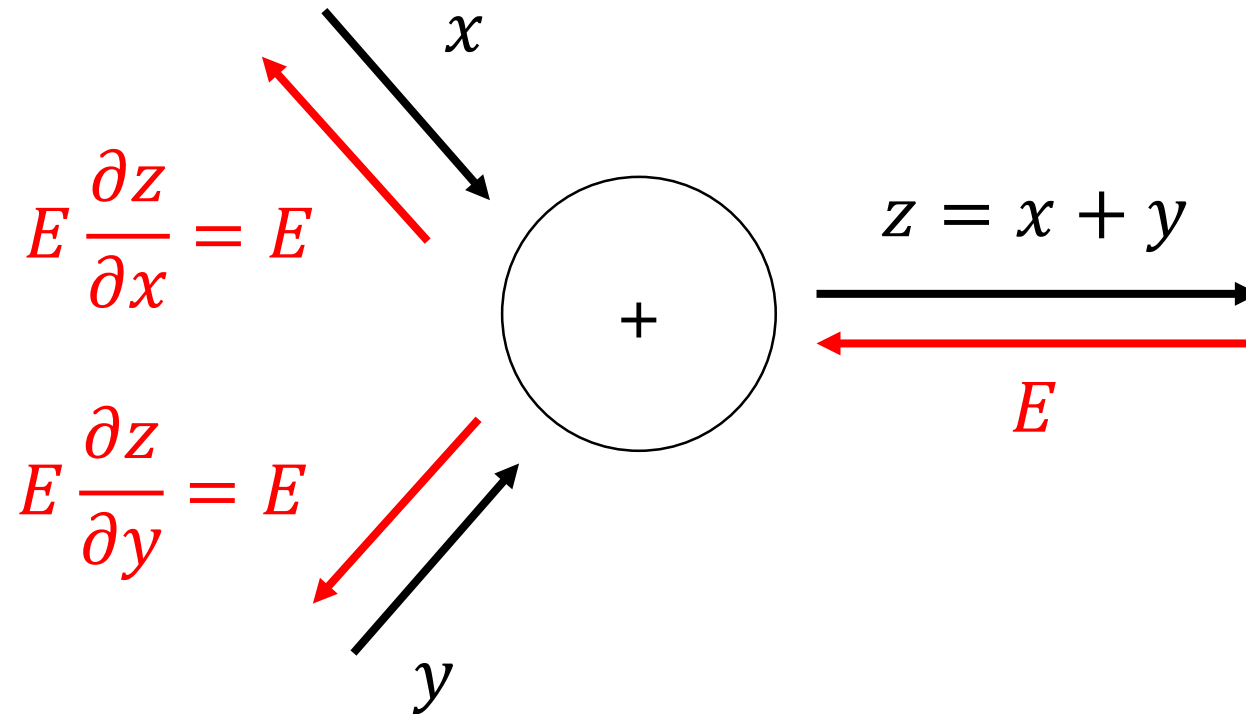
Linear Regression

Neural Network

Logistic Regression

■ 역전파(Backpropagation)

- 덧셈 노드의 역전파
- 앞서 받은 값이 그대로 전달됨



Neural Network

Model Training

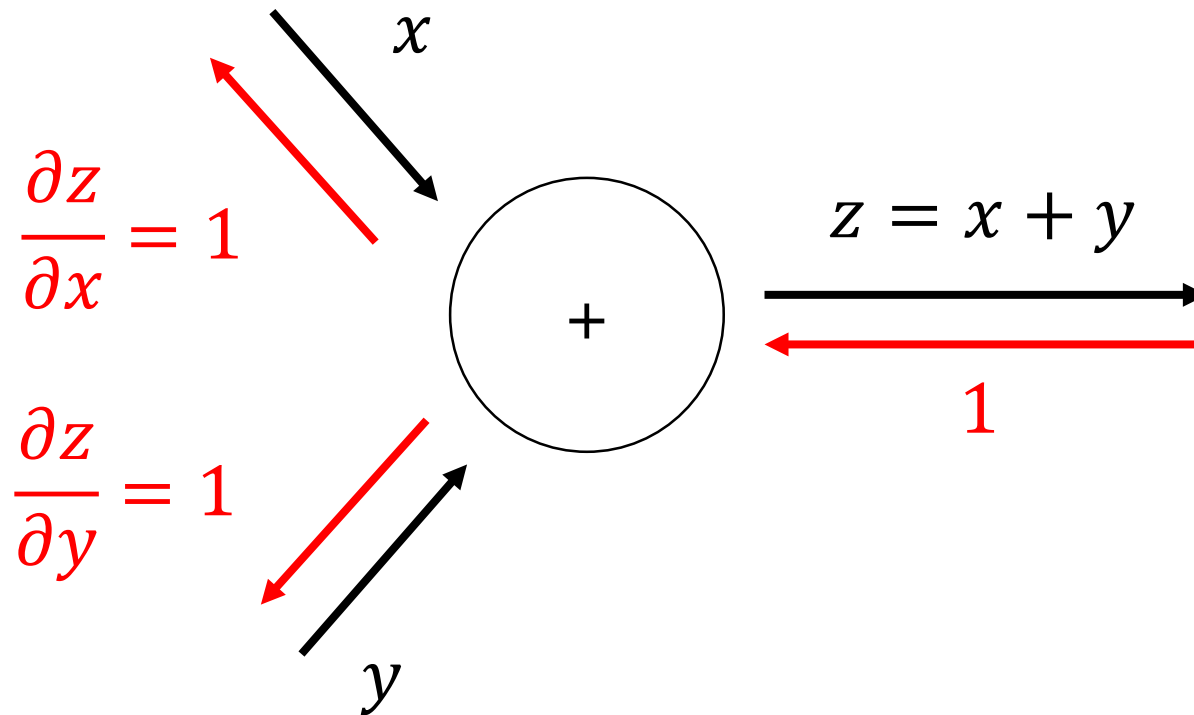
Linear Regression

Neural Network

Logistic Regression

■ 역전파(Backpropagation)

- 덧셈 노드의 역전파
- 앞서 받은 값이 그대로 전달됨



Neural Network

Model Training

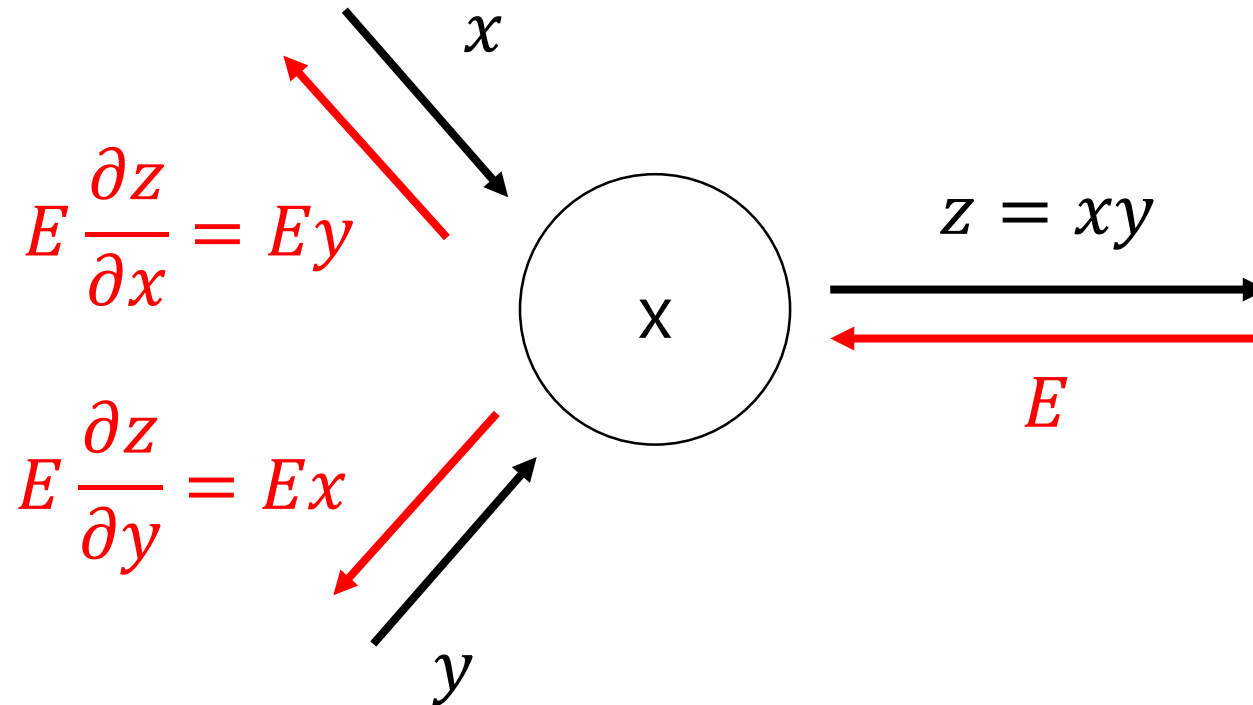
Linear Regression

Neural Network

Logistic Regression

■ 역전파(Backpropagation)

- 곱셈 노드의 역전파
- 앞서 받은 값에 다른 곱해진 값이 전달됨



Neural Network

Model Training

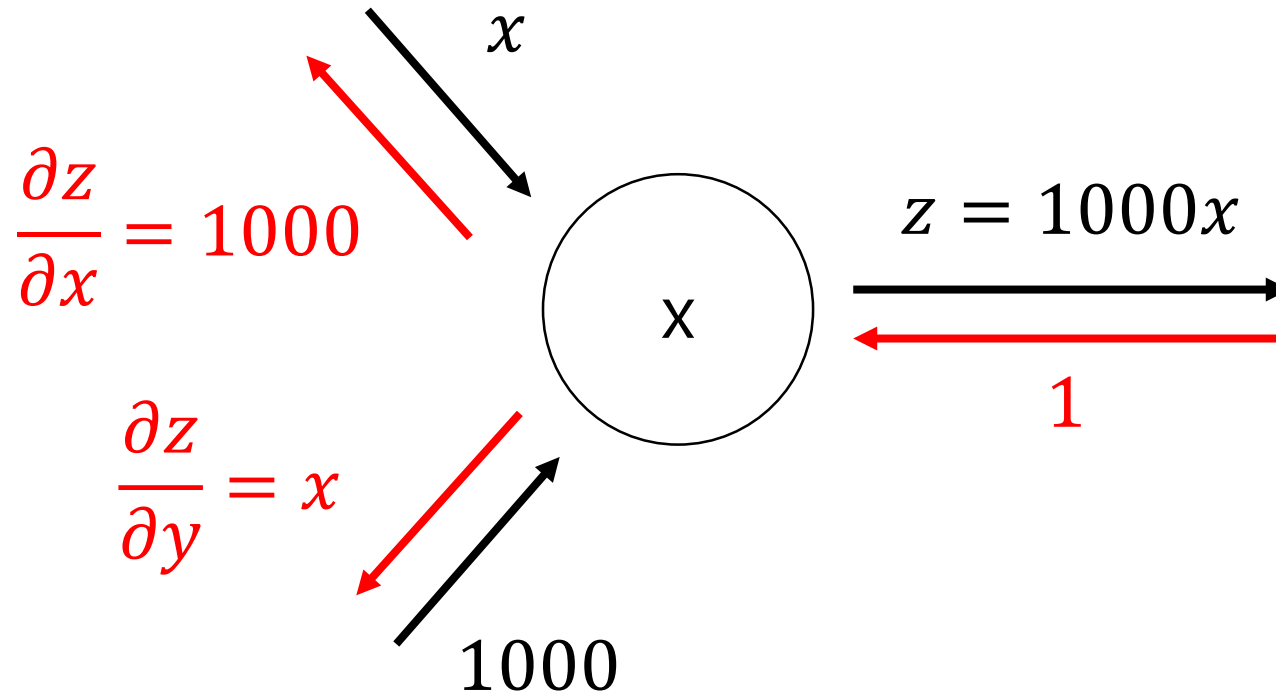
Linear Regression

Neural Network

Logistic Regression

■ 역전파(Backpropagation)

- 곱셈 노드의 역전파
- 앞서 받은 값에 다른 곱해진 값이 전달됨



Neural Network

Model Training

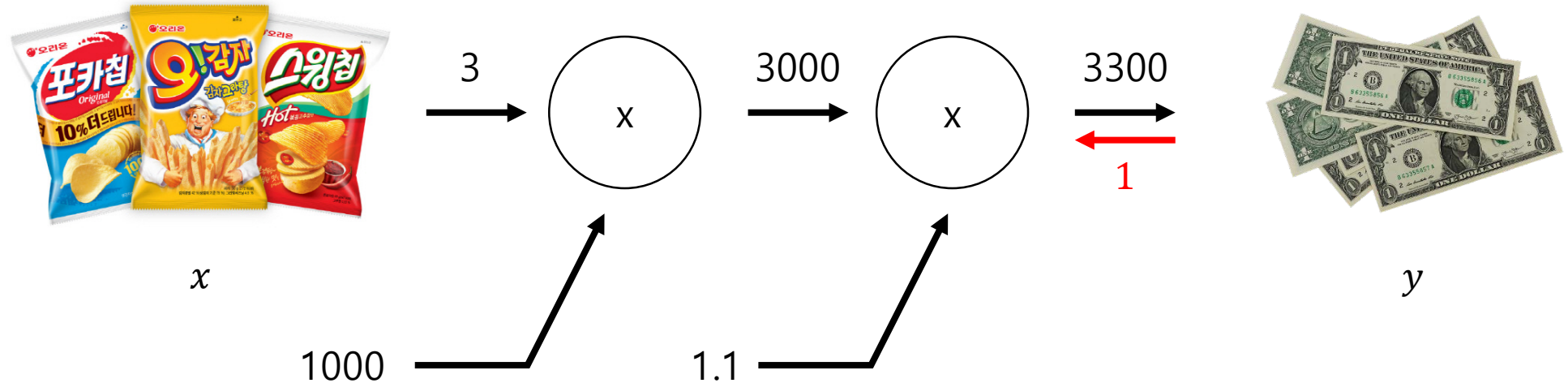
Linear
Regression

Neural Network

Logistic
Regression

■ 신경망(Neural Network)에서의 학습

- 과자를 하나 더 살 때, 지불 금액은 얼마나 증가하는가?



Neural Network

Model Training

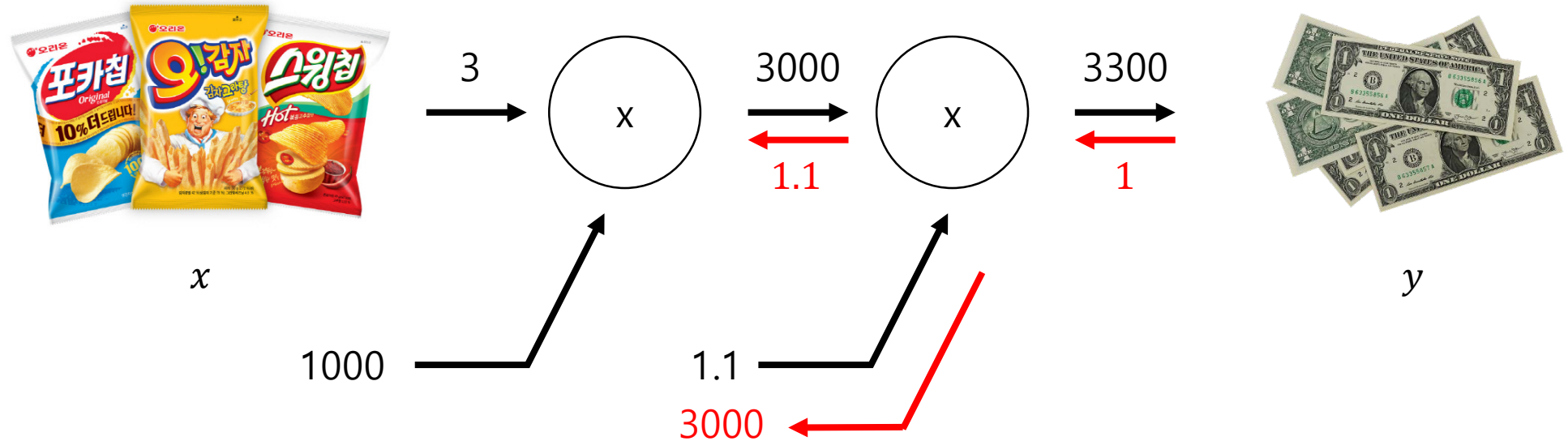
Linear
Regression

Neural Network

Logistic
Regression

■ 신경망(Neural Network)에서의 학습

- 과자를 하나 더 살 때, 지불 금액은 얼마나 증가하는가?



Neural Network

Model Training

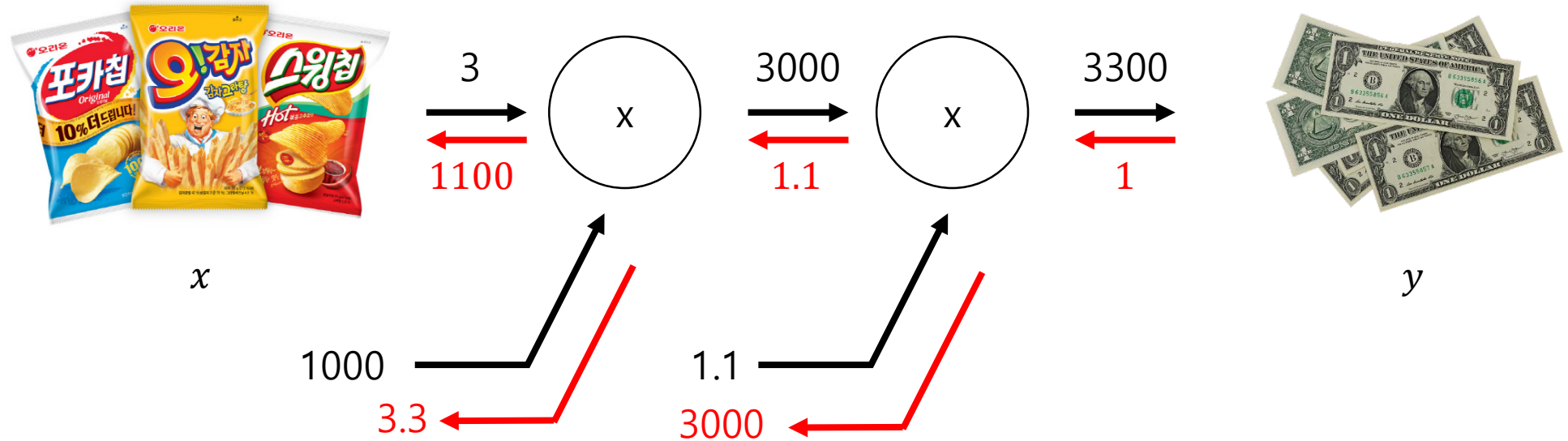
Linear
Regression

Neural Network

Logistic
Regression

■ 신경망(Neural Network)에서의 학습

- 과자를 하나 더 살 때, 지불 금액은 얼마나 증가하는가?



Neural Network

Model Training

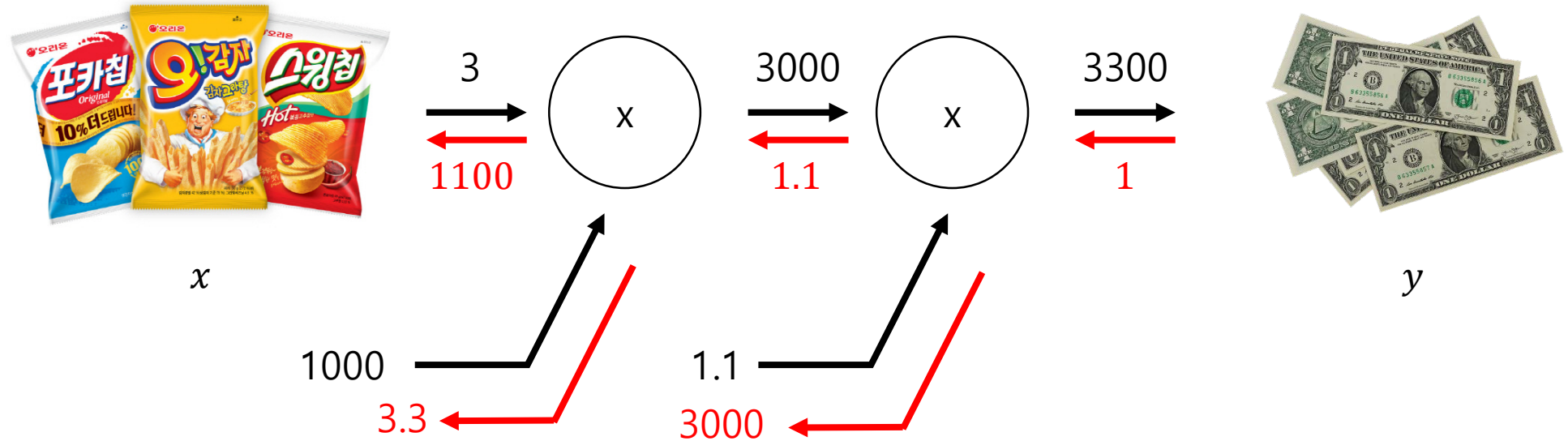
Linear
Regression

Neural Network

Logistic
Regression

■ 신경망(Neural Network)에서의 학습

- 과자를 하나 더 살 때, 지불 금액은 얼마나 증가하는가?



- = 과자 수량(x)의 변화 당 지불 금액(y)의 변화량 = $1100 = \frac{\partial y}{\partial x}$

Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 신경망에서 역전파의 의미
 - 신경망에서 역전파는 어떻게 사용되는가?

Neural Network

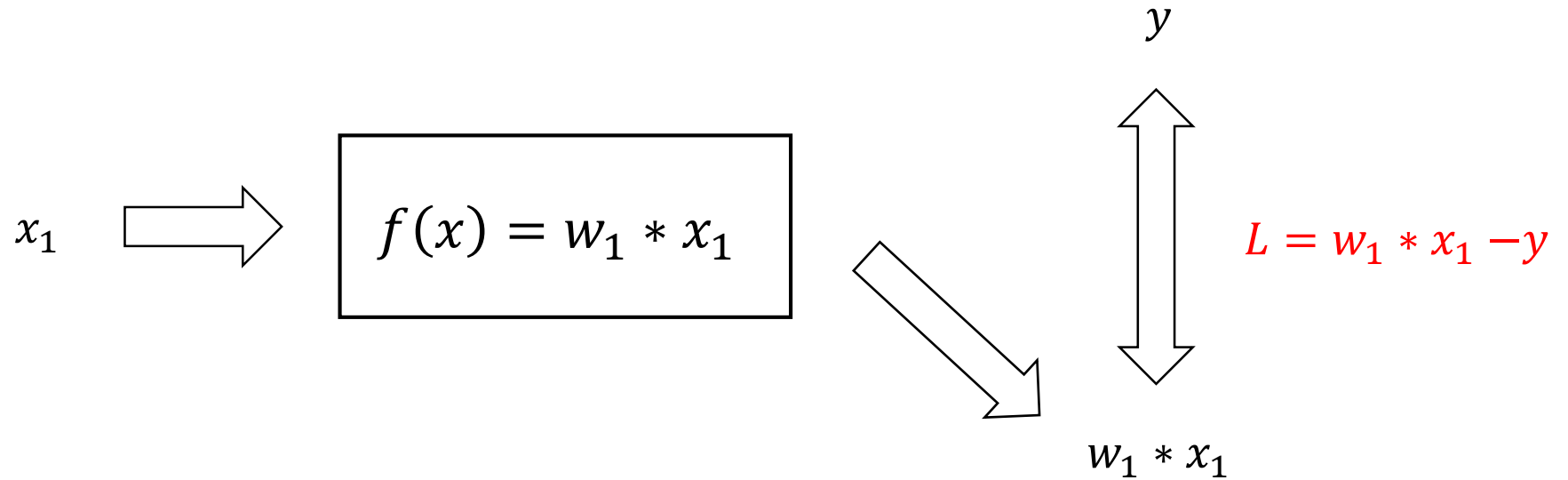
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



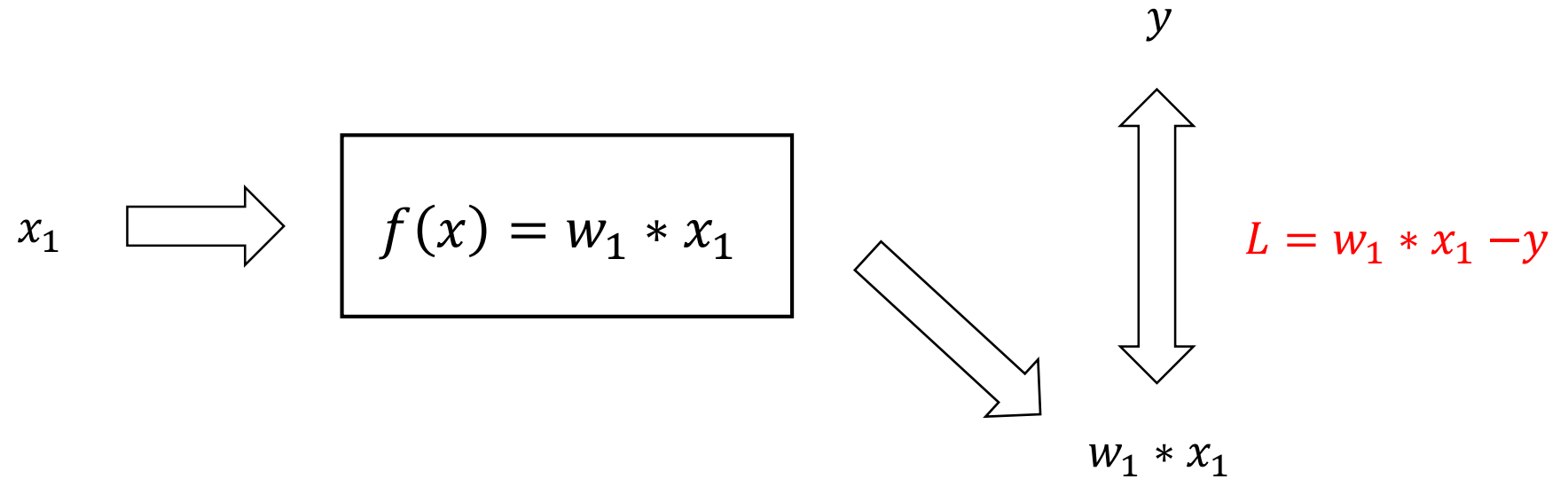
Model Training

Linear Regression

Neural Network

Logistic Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



- 우리가 찾아야하는 것은?

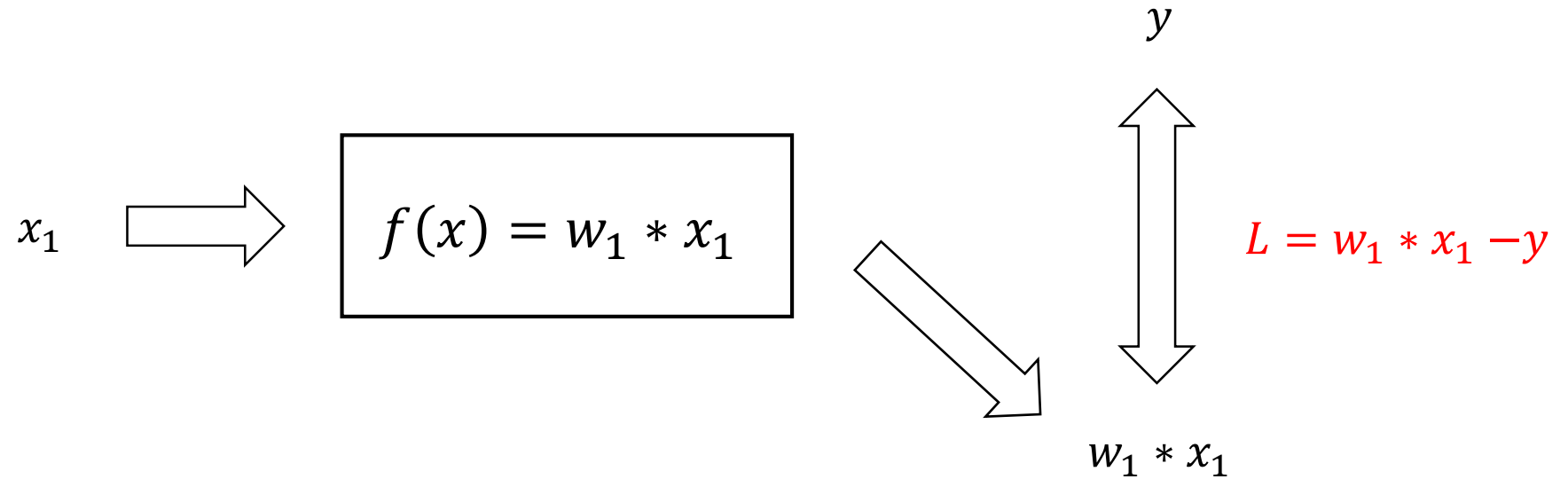
Model Training

Linear Regression

Neural Network

Logistic Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



- 우리가 찾아야하는 것은? **w_1**

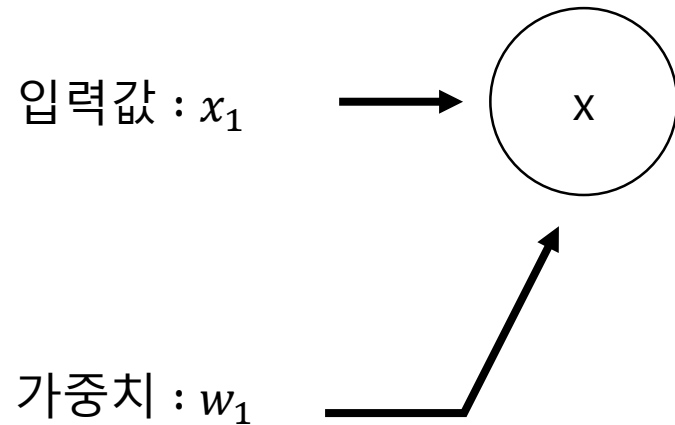
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



Neural Network

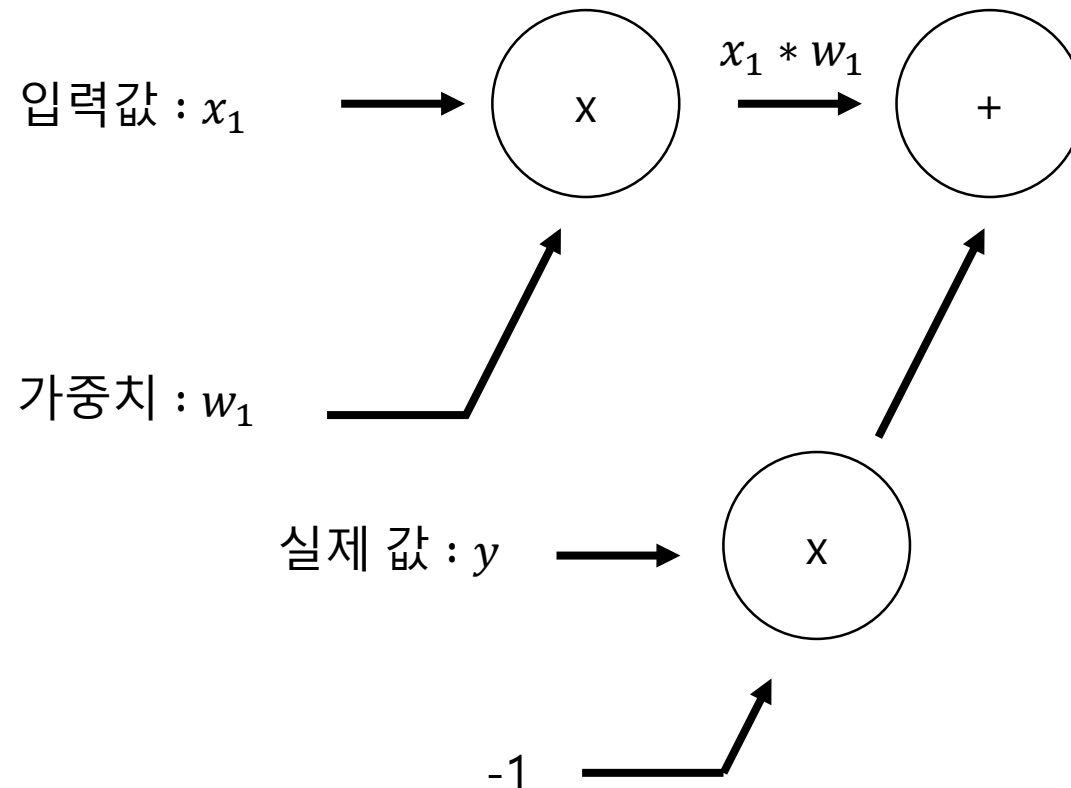
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



Neural Network

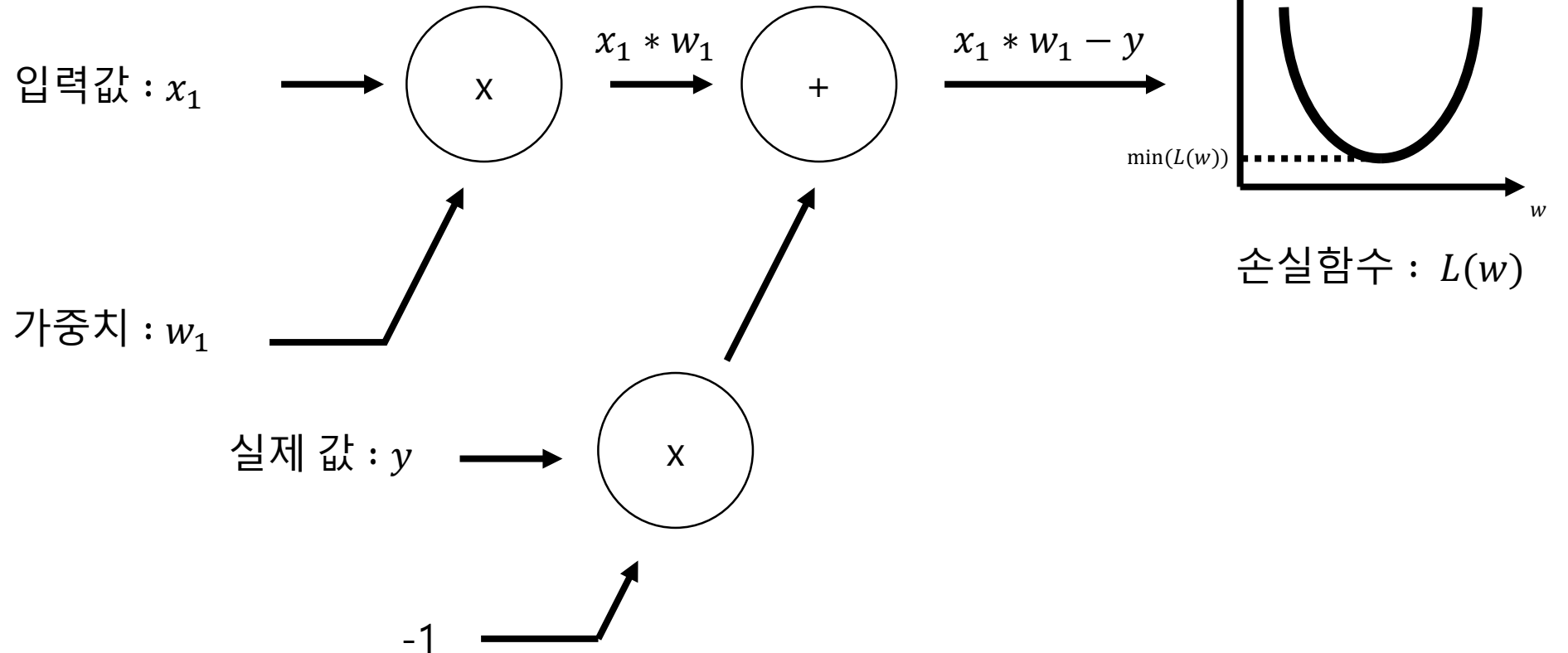
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



Neural Network

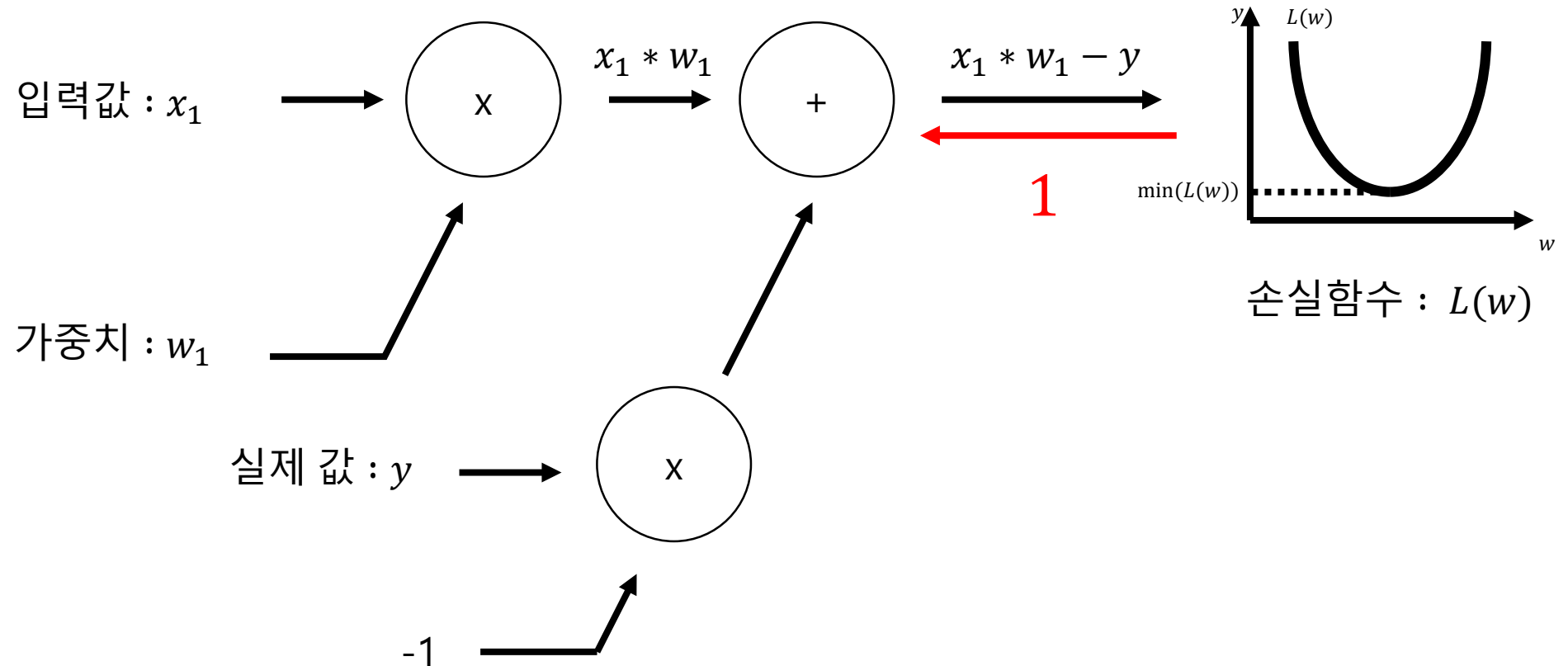
Model Training

Linear Regression

Neural Network

Logistic Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



Neural Network

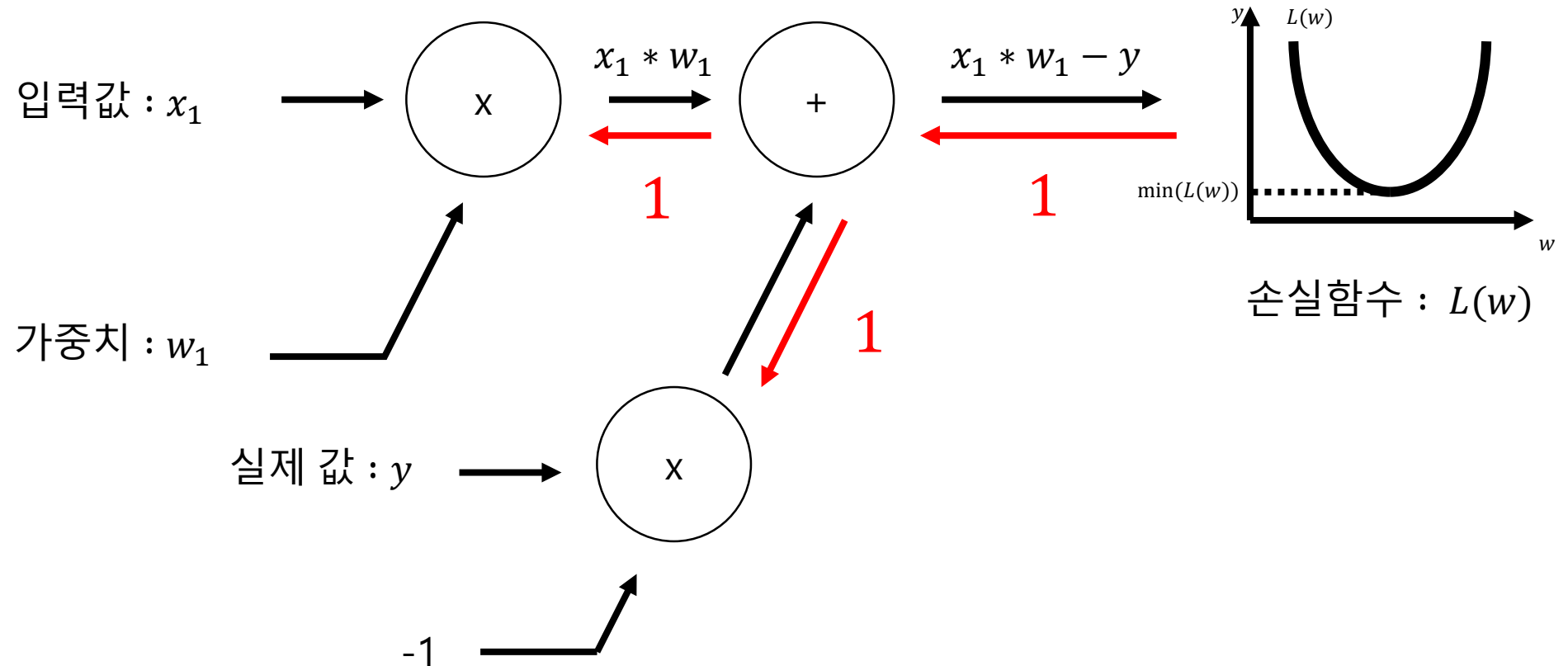
Model Training

Linear Regression

Neural Network

Logistic Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



Neural Network

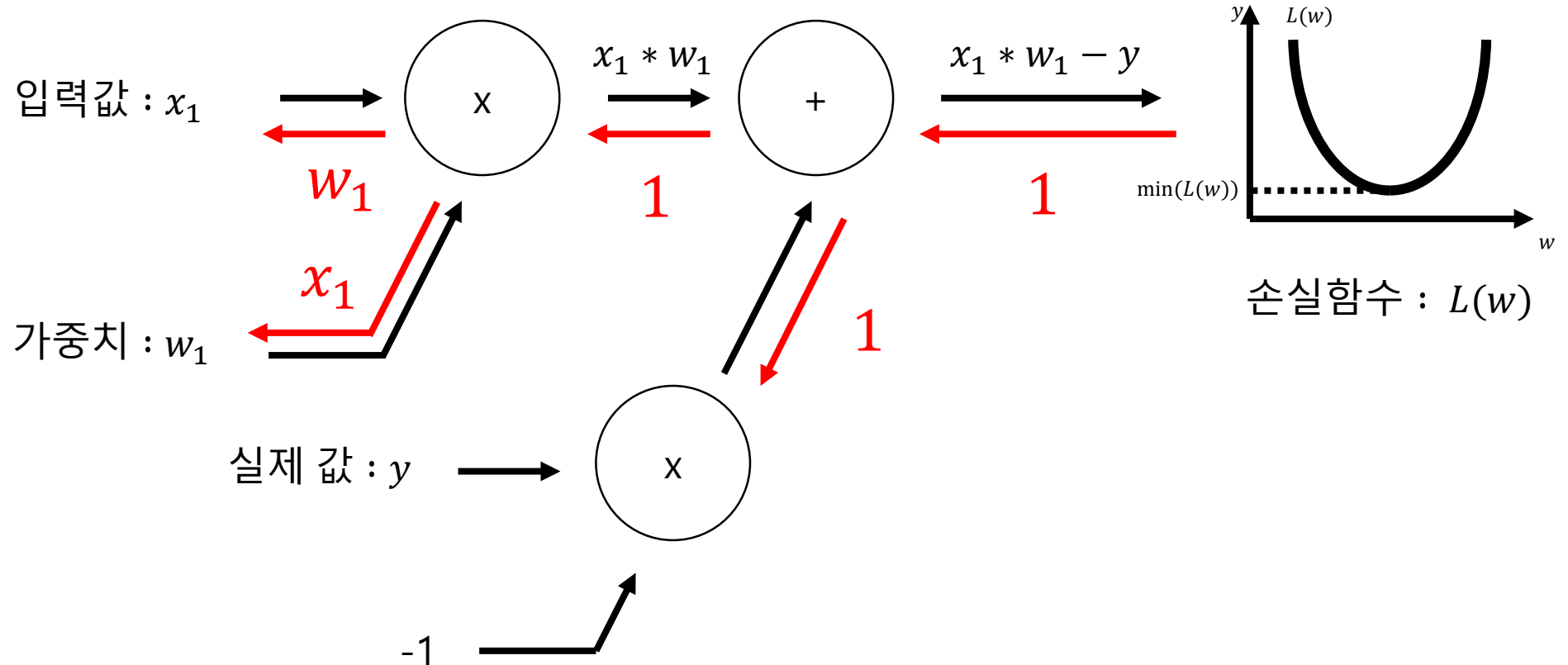
Model Training

Linear Regression

Neural Network

Logistic Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



Neural Network

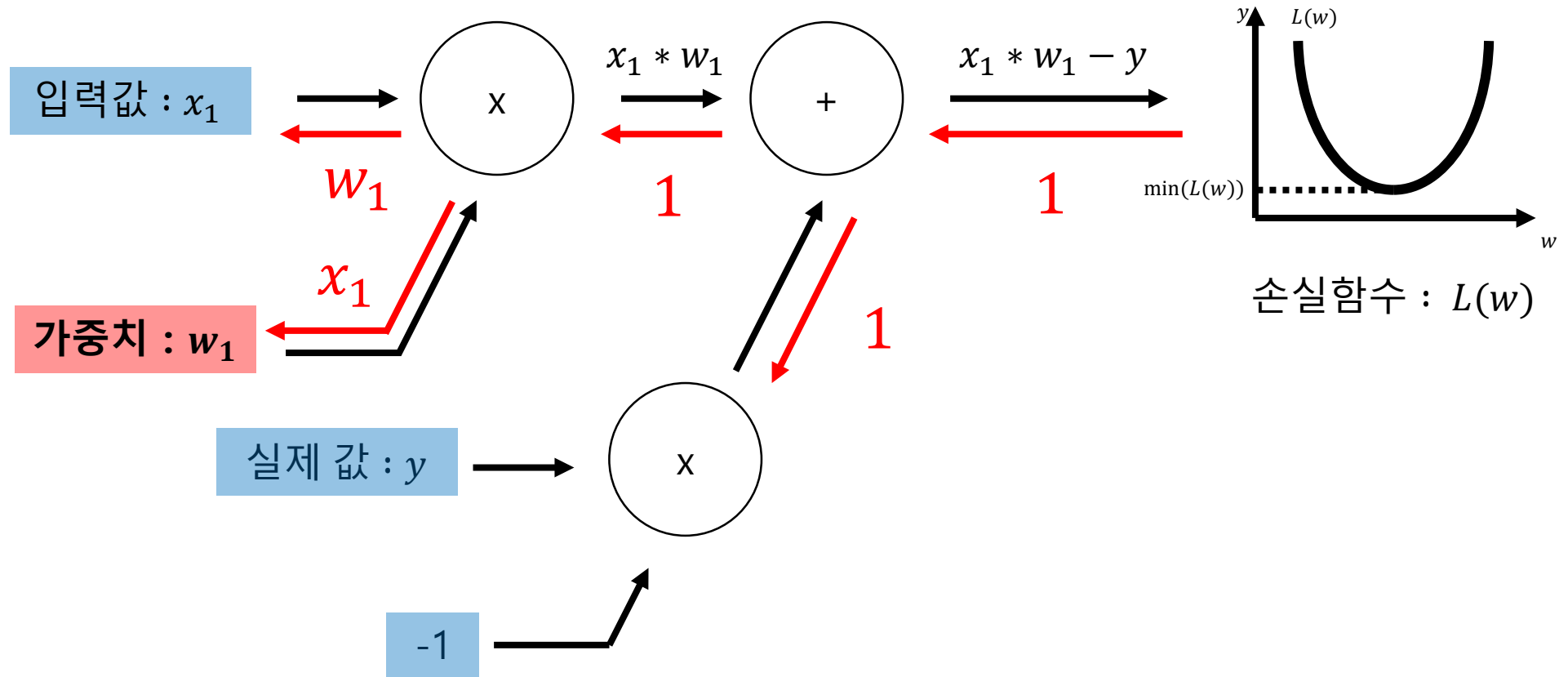
Model Training

Linear Regression

Neural Network

Logistic Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



Neural Network

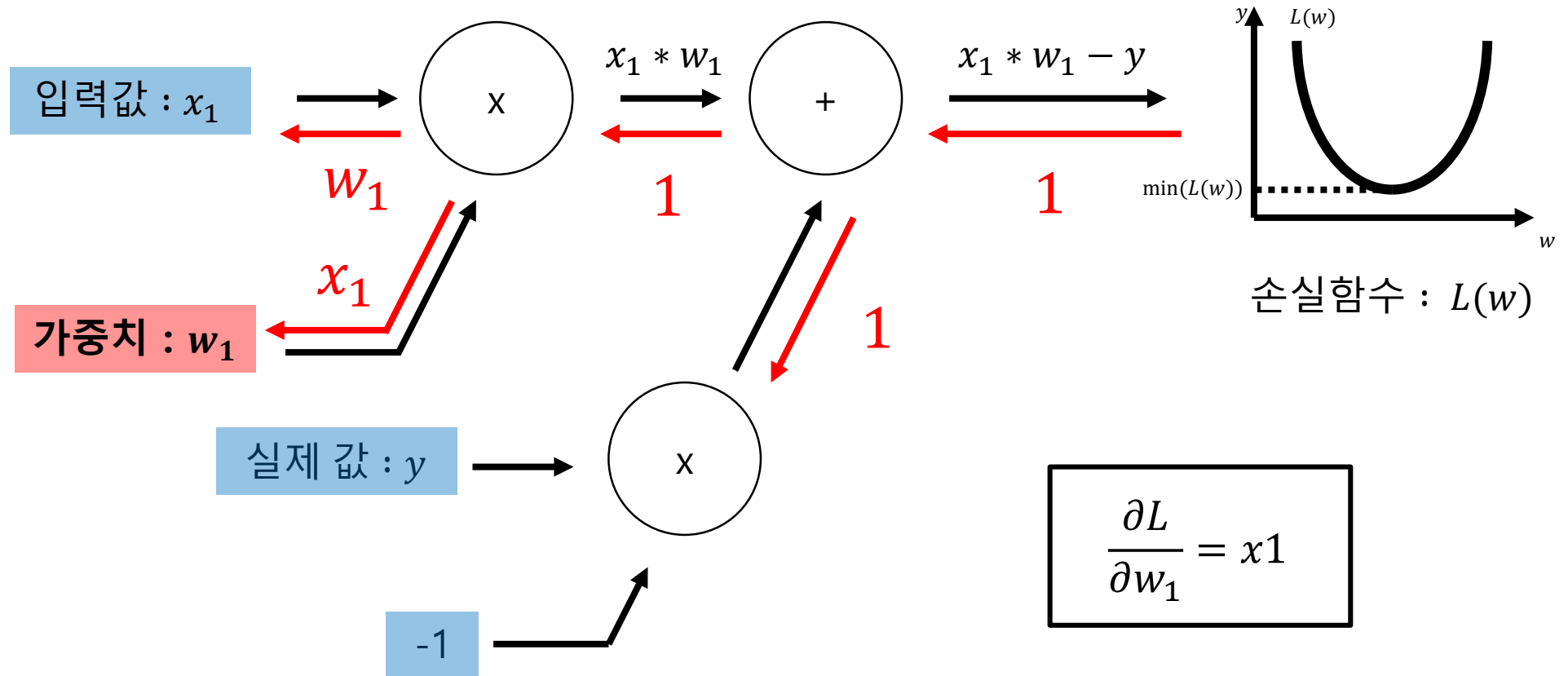
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 신경망에서 역전파의 의미
 - 퍼셉트론의 계산그래프화



■ 신경망에서 역전파의 의미

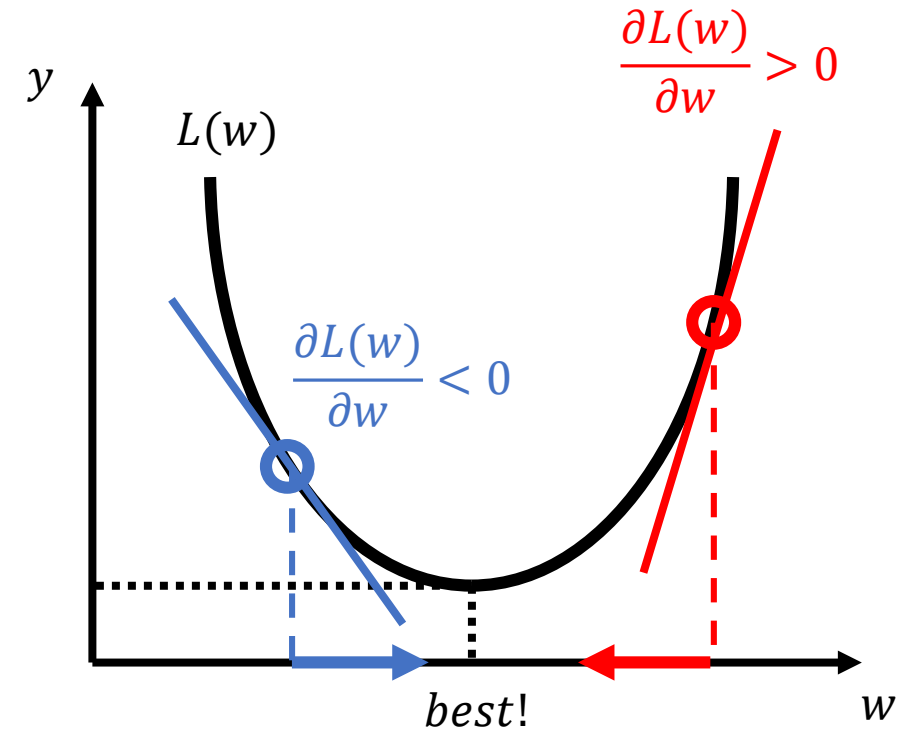
- 퍼셉트론의 계산그래프화

$$\frac{\partial L}{\partial w_1} = x_1 > 0$$

x_1 이 0보다 크면?

$$\frac{\partial L}{\partial w_1} = x_1 < 0$$

x_1 이 0보다 작으면?



Neural Network

Model Training

Linear
Regression

Neural Network

Logistic
Regression

■ 신경망에서 역전파의 의미

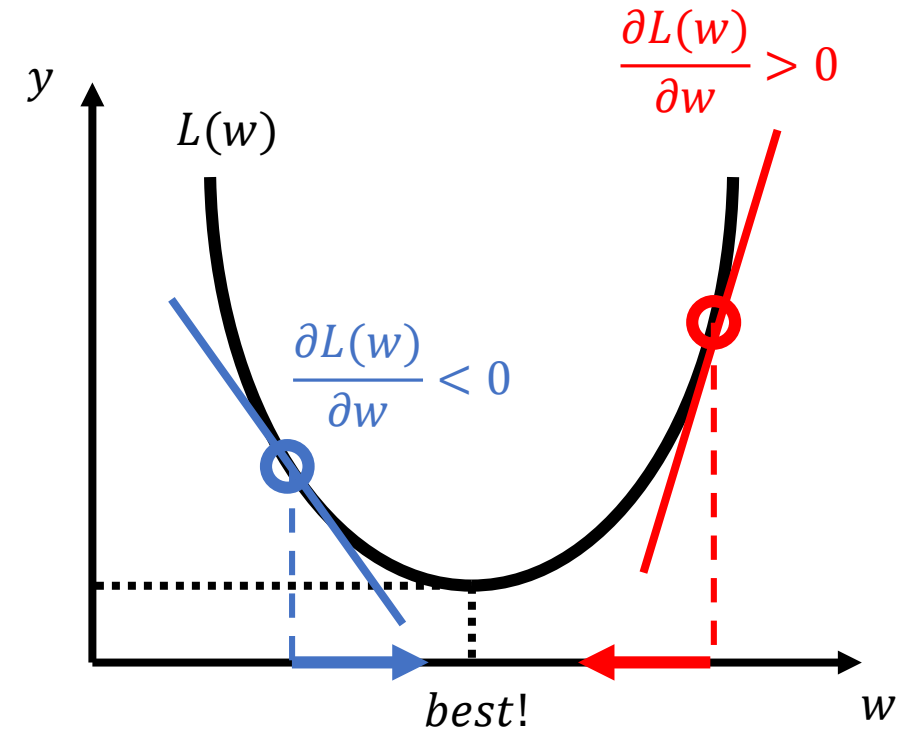
- 퍼셉트론의 계산그래프화

$$\frac{\partial L}{\partial w_1} = x_1 > 0$$

x_1 이 0보다 크면?
= w_1 이 증가하면 L 이 증가

$$\frac{\partial L}{\partial w_1} = x_1 < 0$$

x_1 이 0보다 작으면?



Neural Network

Model Training

Linear
Regression

Neural Network

Logistic
Regression

■ 신경망에서 역전파의 의미

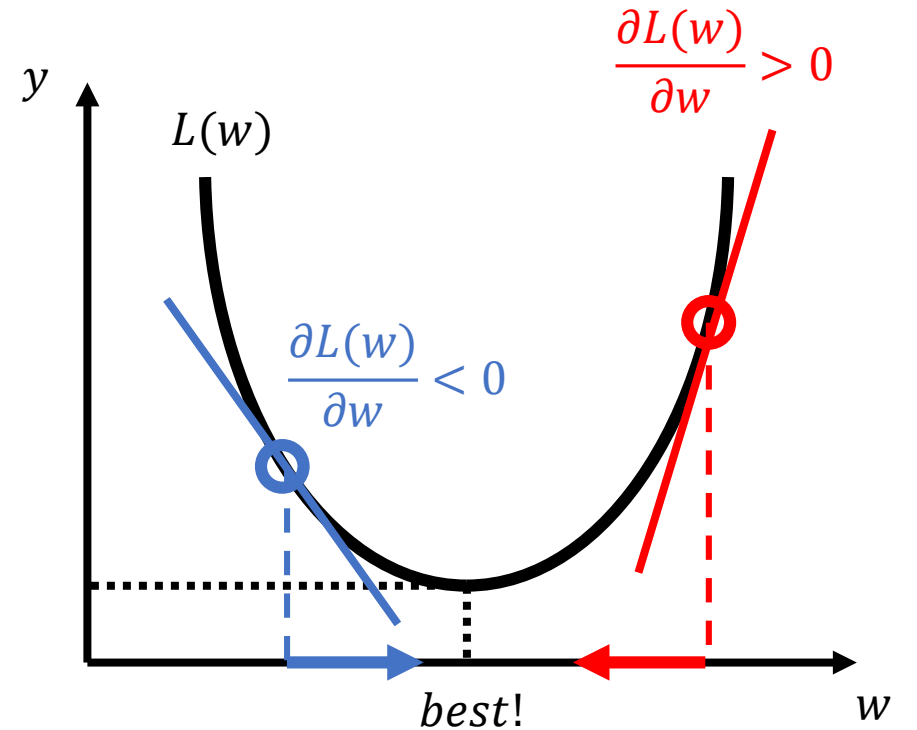
- 퍼셉트론의 계산그래프화

$$\frac{\partial L}{\partial w_1} = x_1 > 0$$

x_1 이 0보다 크면?
= w_1 이 증가하면 L 이 증가

$$\frac{\partial L}{\partial w_1} = x_1 < 0$$

x_1 이 0보다 작으면?
= w_1 이 증가하면 L 이 감소



Neural Network

Model Training

Linear
Regression

Neural Network

Logistic
Regression

■ 신경망에서 역전파의 의미

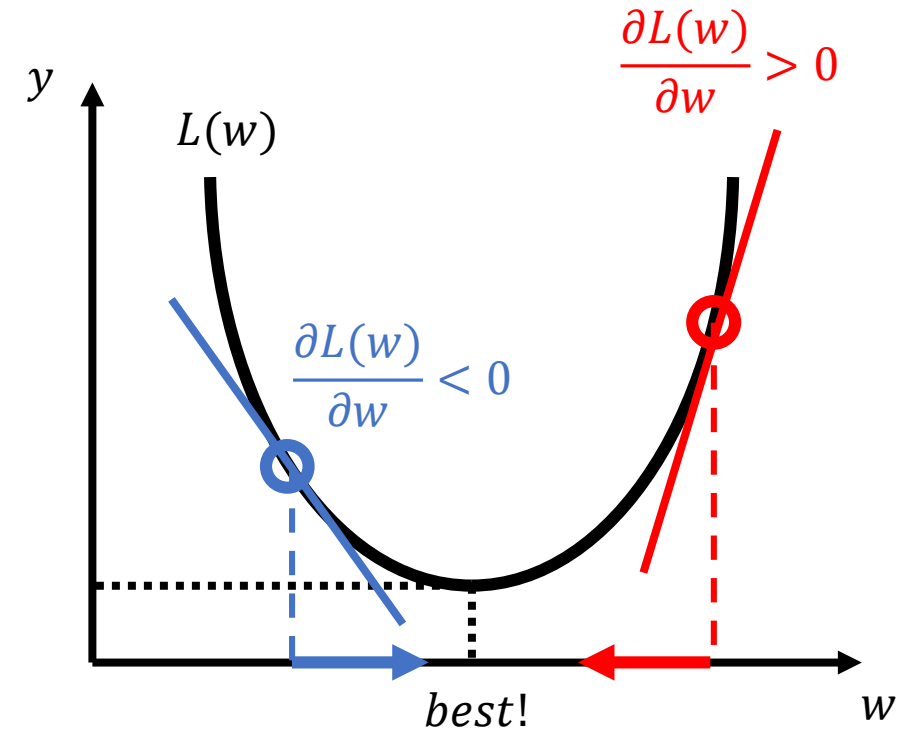
- 퍼셉트론의 계산그래프화

$$\frac{\partial L}{\partial w_1} = x_1 > 0$$

x_1 이 0보다 크면?
= w_1 이 증가하면 L 이 증가

$$\frac{\partial L}{\partial w_1} = x_1 < 0$$

x_1 이 0보다 작으면?
= w_1 이 증가하면 L 이 감소



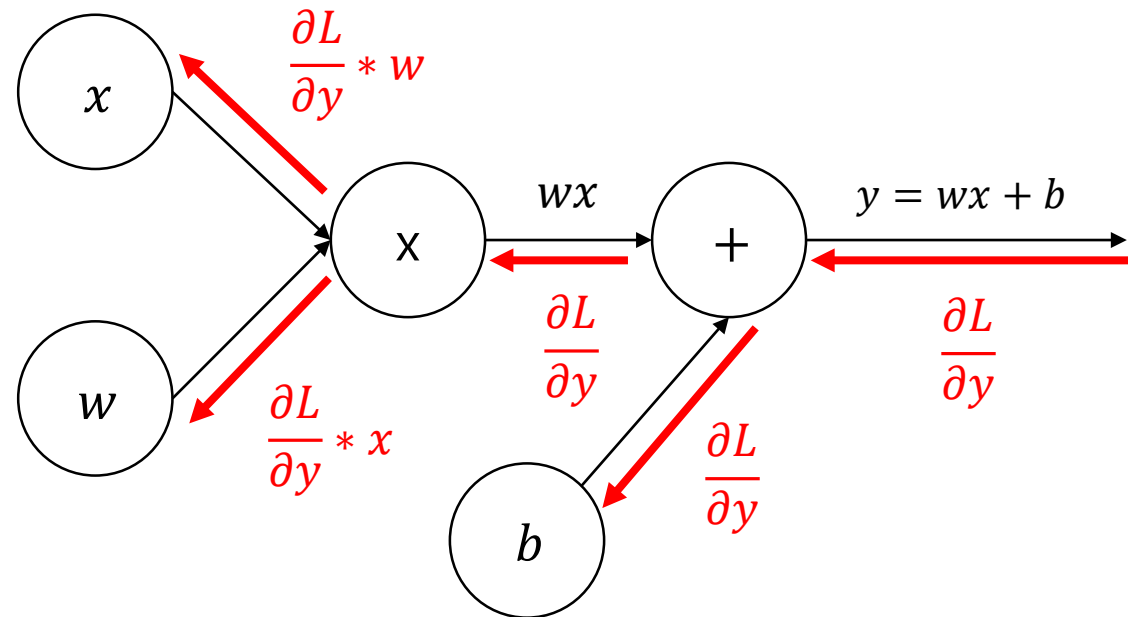
- 위를 활용하여 최적의 가중치 값을 찾아내는 것이 목표

■ 선형 회귀에서의 역전파

- $MSE = \frac{\sum err^2}{n} = \frac{\sum_x (w^T x - y)^2}{n}$

- $y = wx + b$

- $grad = \frac{\partial}{\partial w} MSE = \frac{1}{n} \frac{\partial}{\partial w} \sum_x (wx - y)^2 = \frac{\sum_x 2(wx - y)x}{n}$



$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial y} * x$$

Neural Network

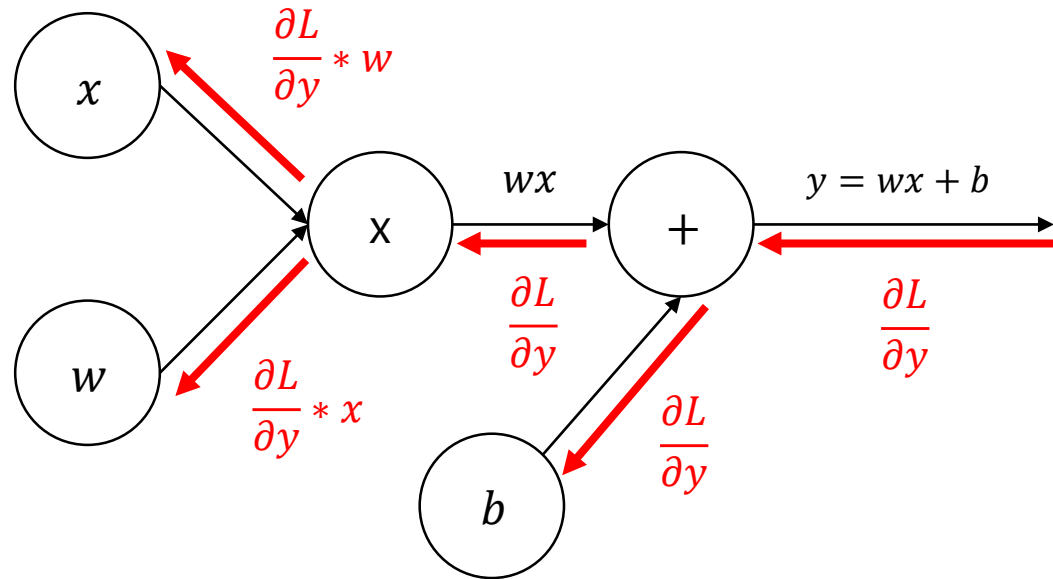
Model Training

Linear
Regression

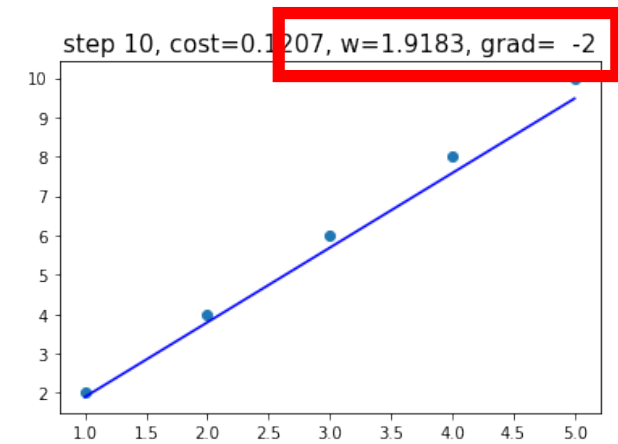
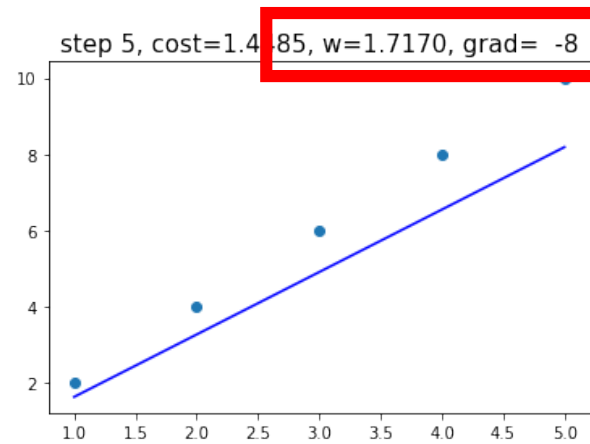
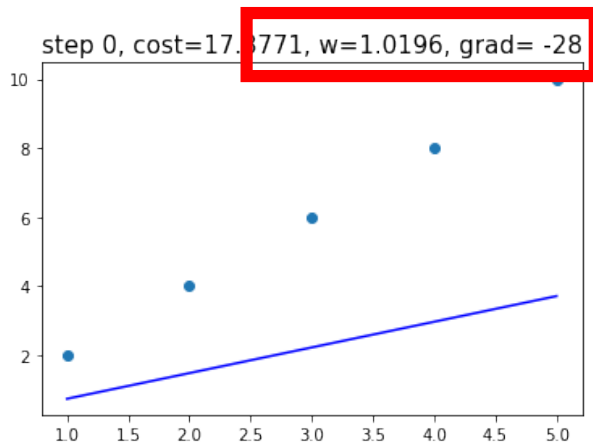
Neural Network

Logistic
Regression

■ 선형 회귀에서의 역전파



$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial y} * x$$



Neural Network

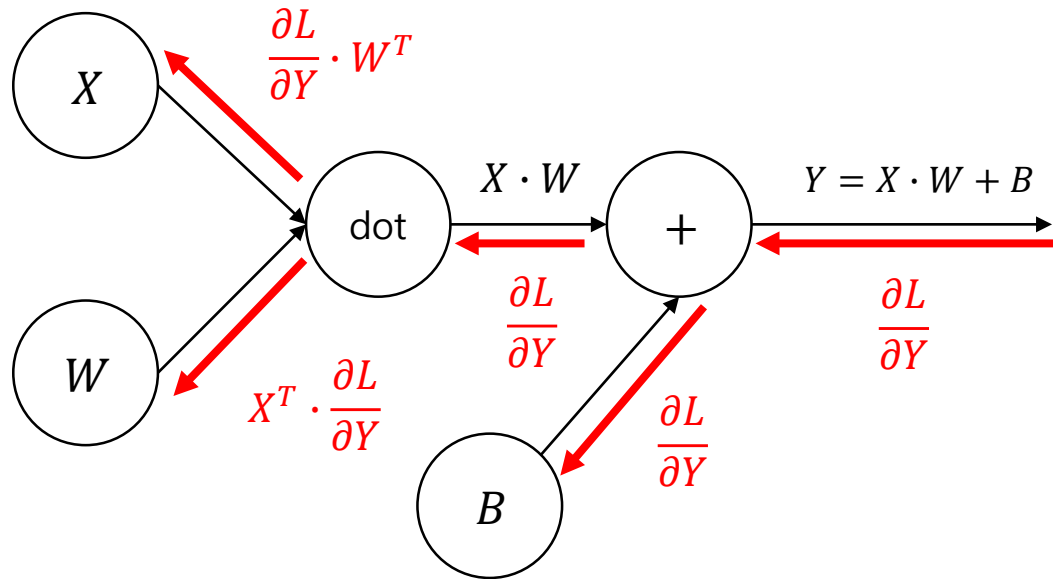
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 다변량 선형 회귀에서의 역전파



$$\frac{\partial L}{\partial W} = X^T \cdot \frac{\partial L}{\partial Y}$$

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}$$

$$W = \begin{pmatrix} w_{11} & \cdots & w_{1l} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{ml} \end{pmatrix}$$

$$Y = \begin{pmatrix} y_{11} & \cdots & y_{1l} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nl} \end{pmatrix}$$

Model Training

Linear
Regression

Neural Network

Logistic
Regression

4. Logistic Regression

Logistic Regression

Model Training

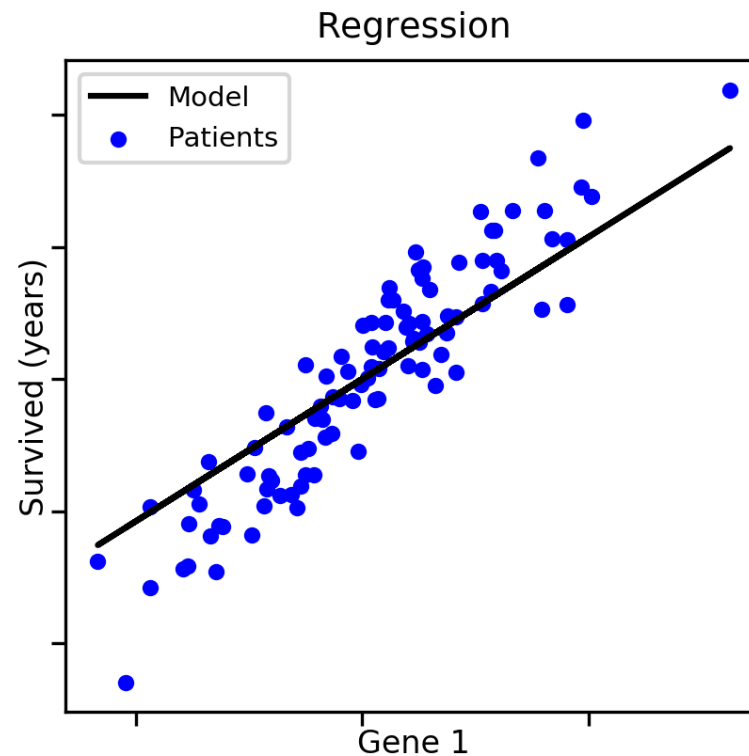
Linear
Regression

Neural Network

Logistic
Regression

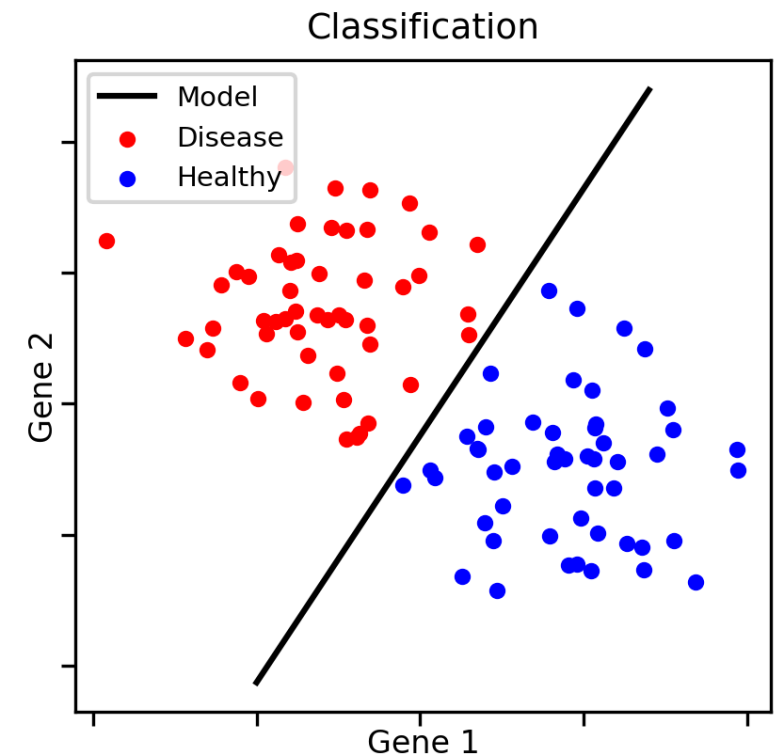
교사 학습의 분류

- 회귀(Regression) : 연속적인 값을 예측
 - EX) 집의 가격 예측



<https://aldro61.github.io/microbiome-summer-school-2017/sections/basics/>

- 분류(Classification) : 이산적인 값을 예측
 - EX) 암 발병 예측



<https://aldro61.github.io/microbiome-summer-school-2017/sections/basics/>

Logistic Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

- **교사 학습의 분류**
 - 연체 횟수, 연체 금액 → 신용불량자 여부
 - 정답이 존재하므로 **교사 학습(Supervised Learning)**에 해당
 - Input(X) : 연체 횟수, 연체 금액
 - Output(Y) : 신용불량자 여부
 - 목표 변수가 이산(0,1)이므로 **분류(Classification)**에 해당
 - 분류 모델 :
 - **Logistic Regression**
 - Decision Tree
 - SVM

Logistic Regression

Model Training

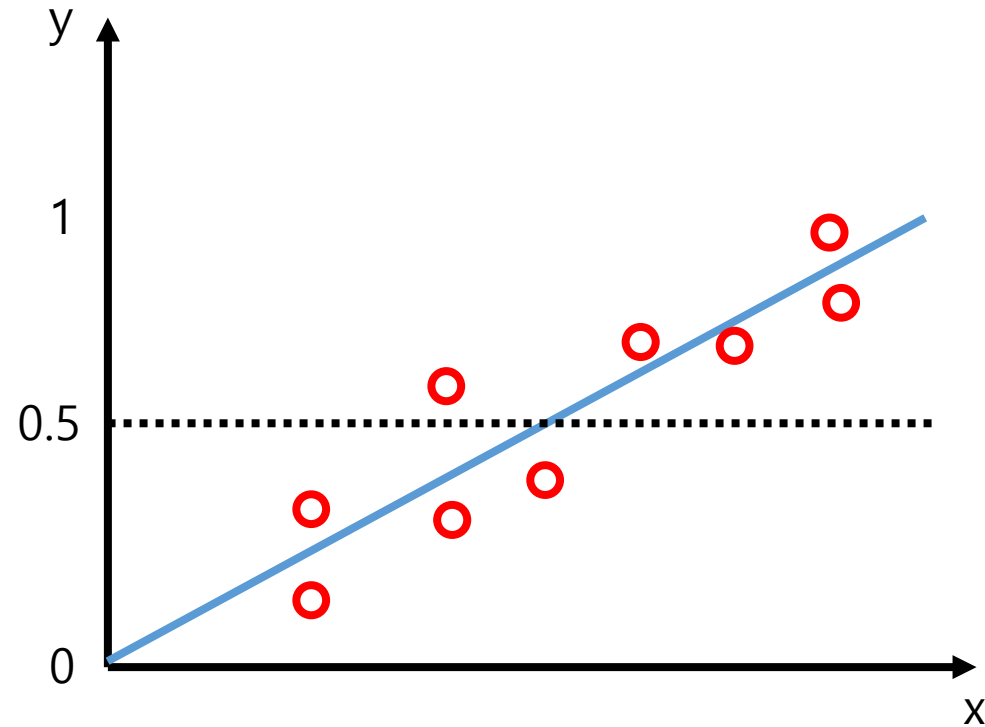
Linear
Regression

Neural Network

Logistic
Regression

로지스틱 회귀 분석(Logistic Regression)

- 선형 회귀 분석(Linear Regression)에서의 결과값을 분류에 활용하면 어떨까?
- 결과값을 0, 1로 놓고 회귀
- EX) 예측값이 0.5보다 크면 신용불량자



Logistic Regression

Model Training

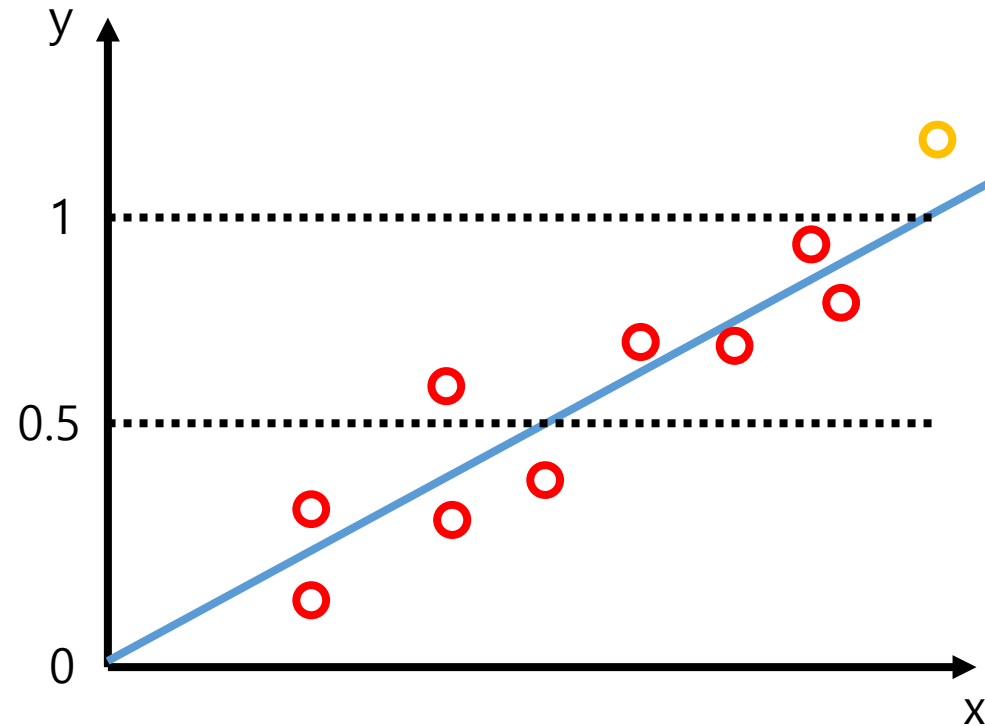
Linear
Regression

Neural Network

Logistic
Regression

로지스틱 회귀 분석(Logistic Regression)

- 그러나 이 경우 아래와 같은 단점이 있다.
 - 선형 회귀 모델이 1보다 크거나 0보다 작은 값을 반환할 수 있다.
 - 반환된 값이 확률과는 다른 의미이다.



Logistic Regression

Model Training

Linear
Regression

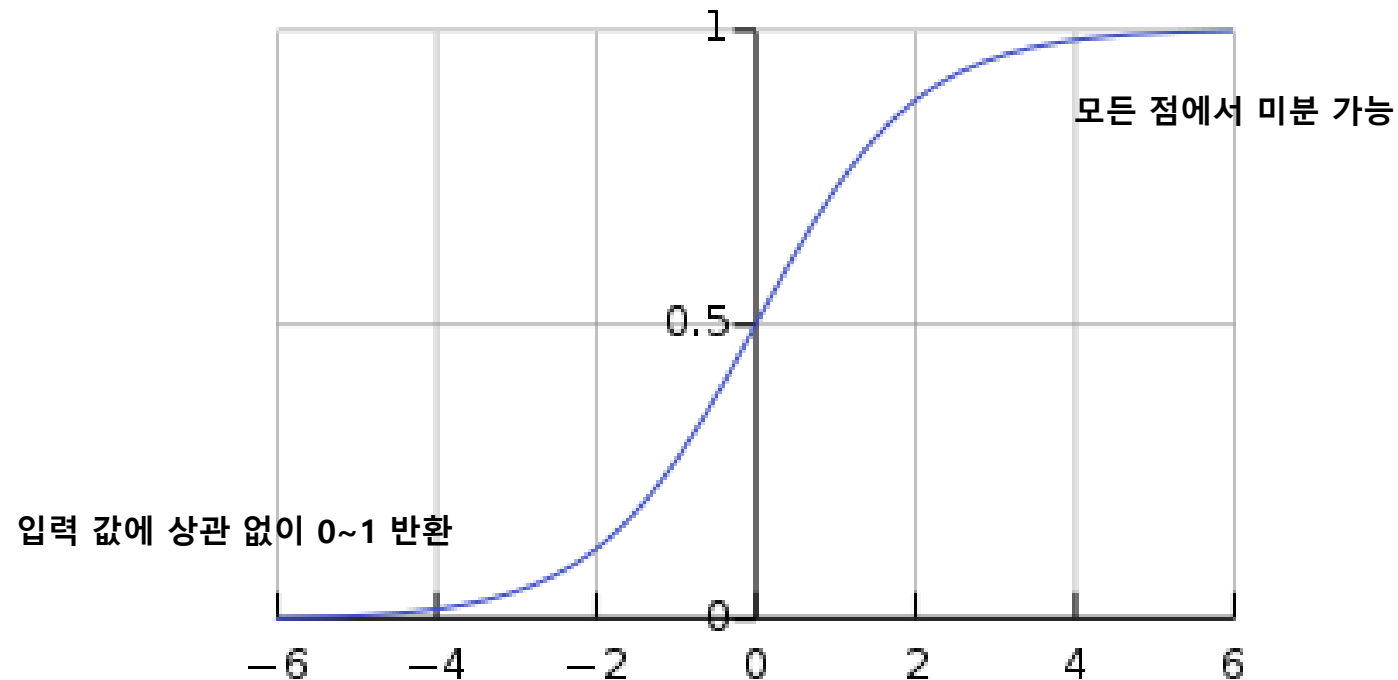
Neural Network

Logistic
Regression

로지스틱 회귀 분석(Logistic Regression)

- 로지스틱 함수(Logistic Function) or 시그모이드 함수(Sigmoid Function)

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



Logistic Regression

Model Training

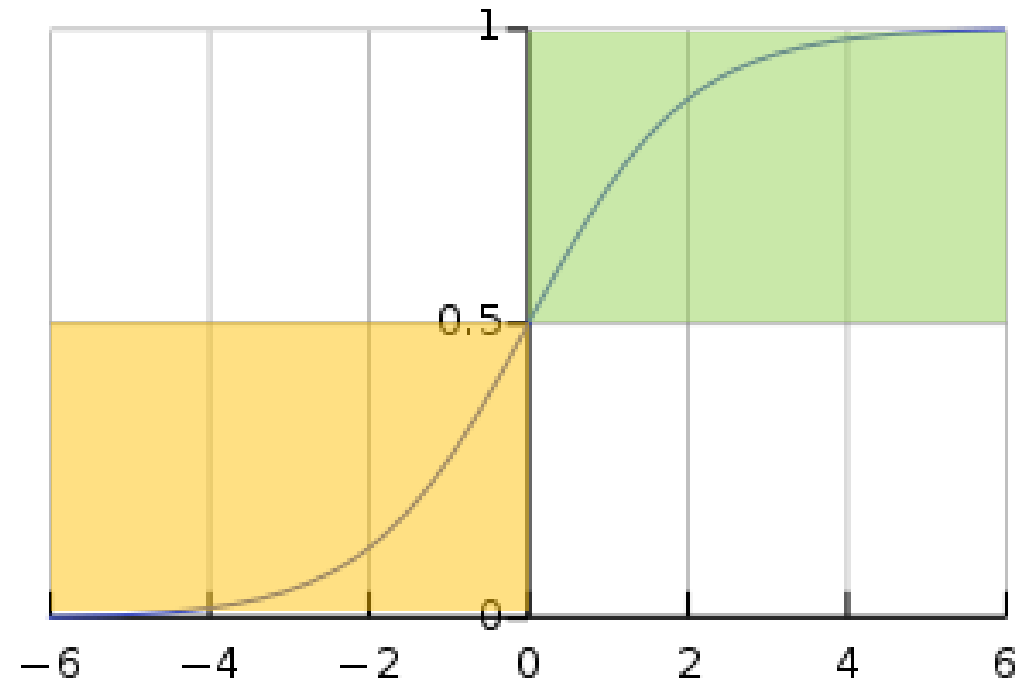
Linear
Regression

Neural Network

Logistic
Regression

로지스틱 회귀 분석(Logistic Regression)

- $f_w(x) = \sigma(wx) = \frac{1}{1+\exp(-wx)}$: 예측값
- $f_w(x) = P(y = 1|x, w)$: 입력 값을 x 로 갖는 값이 1로 분류될 확률
- $P(y = 1|x, w) + P(y = 0|x, w) = 1$
- $y = \begin{cases} 1 & (\text{if } wx > 0) \\ 0 & (\text{if } wx < 0) \end{cases}$



Logistic Regression

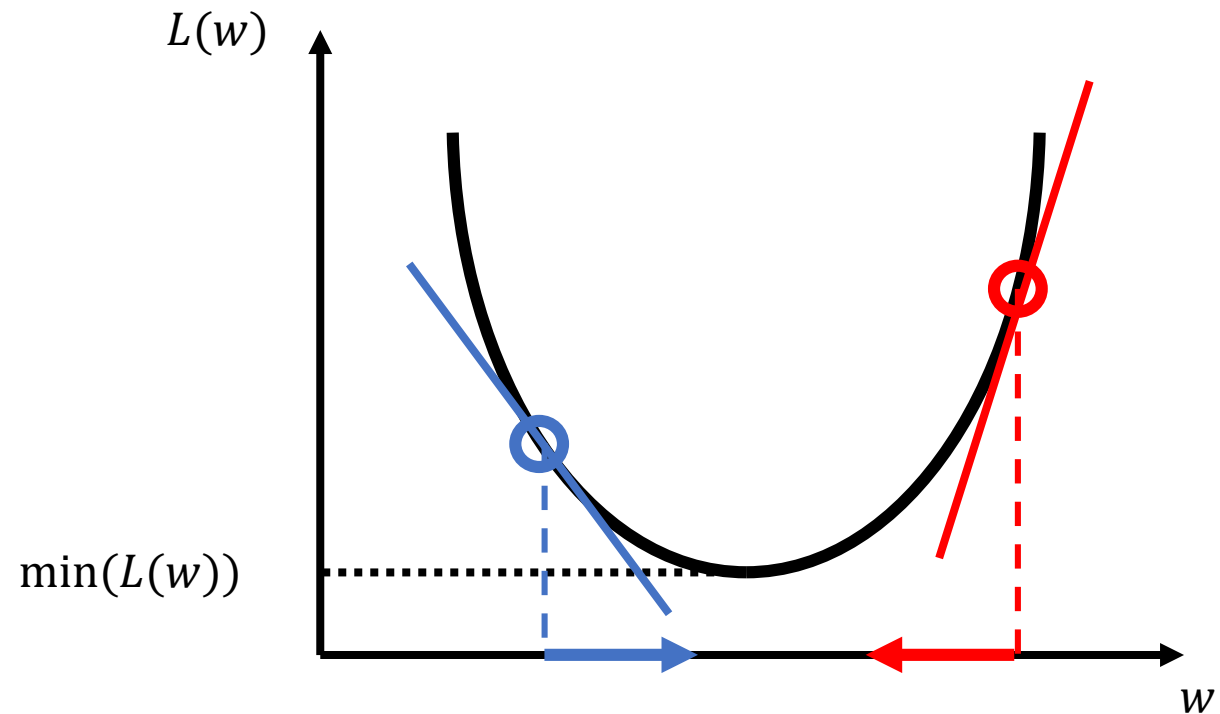
Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 로지스틱 회귀 분석의 손실 함수
 - MSE를 손실 함수로 사용하면?
 - w 로 나타낸 손실 함수가 볼록(Convex)하지 않음



Logistic Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

로지스틱 회귀 분석의 손실 함수

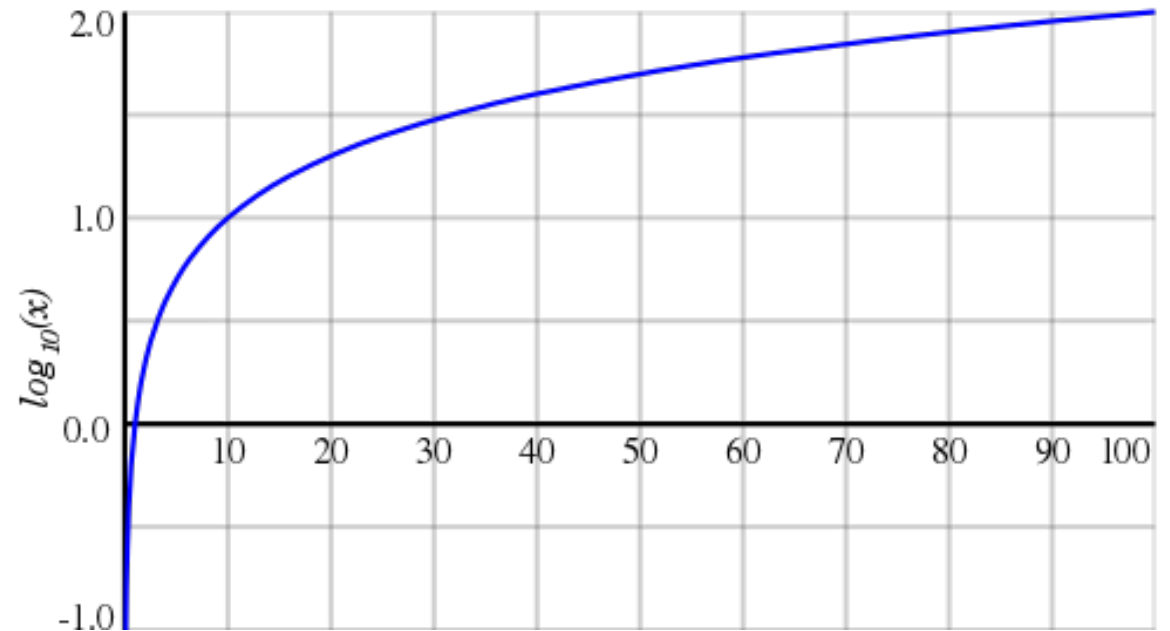
$$\text{Loss}(f_w(x), y) = \begin{cases} -\log(f_w(x)) & (\text{if } y = 1) \\ -\log(1 - f_w(x)) & (\text{if } y = 0) \end{cases}$$

$y = 1$ 일 때 :

- $f_w(x) = 1$ 이면 $\text{Loss} = 0$
- $f_w(x) = 0$ 이면 $\text{Loss} = \infty$

$y = 0$ 일 때 :

- $f_w(x) = 0$ 이면 $\text{Loss} = 0$
- $f_w(x) = 1$ 이면 $\text{Loss} = \infty$



Logistic Regression

Model Training

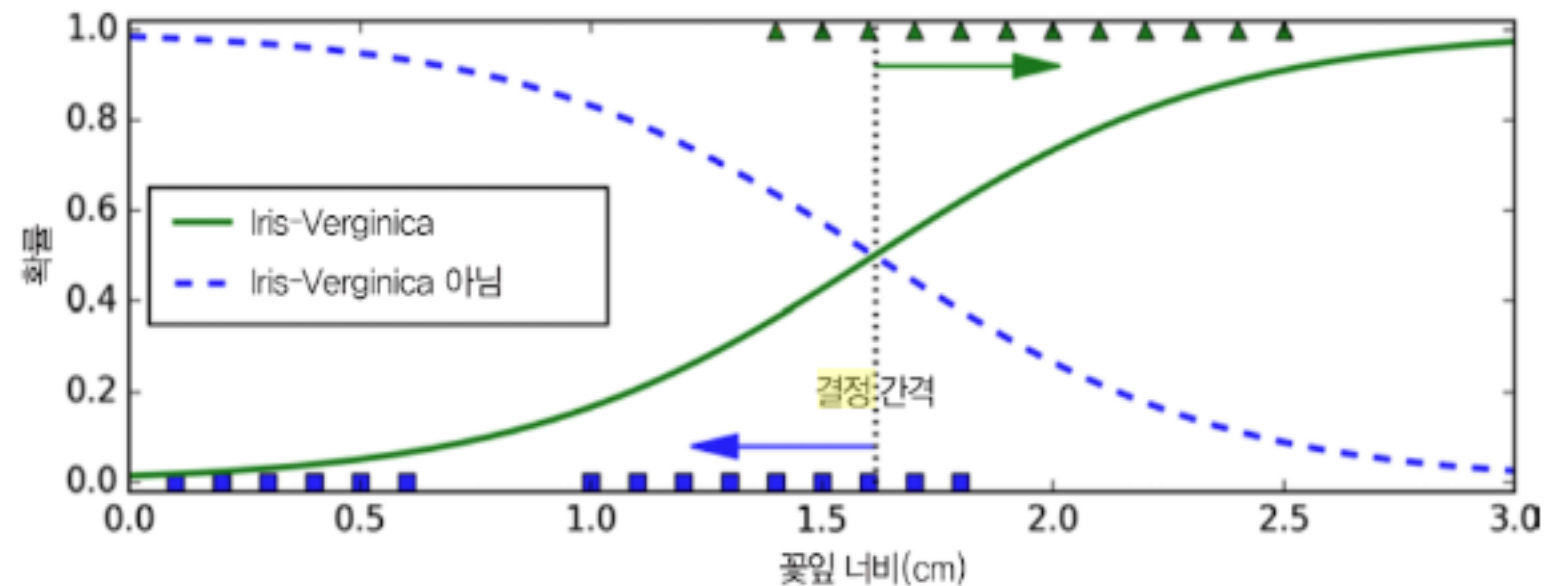
Linear
Regression

Neural Network

Logistic
Regression

로지스틱 회귀 분석의 손실 함수

- $Loss(f_w(x), y) = -y \log(f_w(x)) - (1 - y) \log(1 - f_w(x))$
 - 볼록함수(Convex)
 - 경사하강법(Gradient Descent) 사용 가능



Logistic Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 로지스틱 회귀 분석의 손실 함수
 - 맞춰야할 분류가 3개 이상일 경우?
 - 동물 = 0
 - 식물 = 1
 - 사람 = 2

Logistic Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 로지스틱 회귀 분석의 손실 함수
 - 맞춰야할 분류가 3개 이상일 경우?
 - 동물 = 0
 - 식물 = 1
 - 사람 = 2
 - 식물 + 식물 = 사람?
 - 동물 + 사람 = 사람?

Logistic Regression

Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 로지스틱 회귀 분석의 손실 함수
 - 맞춰야할 분류가 3개 이상일 경우?
 - One-Hot Encoding
- 동물 : [1, 0, 0]
- 식물 : [0, 1, 0]
- 사람 : [0, 0, 1]

Model Training

Linear
Regression

Neural Network

Logistic
Regression

- 로지스틱 회귀 분석의 손실 함수
 - 교차 엔트로피 오차(Cross Entropy Error, CEE)
 - 분류에서 자주 사용되는 손실 함수
 - 정답과 오차의 엔트로피의 합
 - One-Hot Encoding의 경우 아래와 같이 기술 가능
 - $CEE = -\sum y_k \log(f_w(x)_k)$
 - 앞의 *Loss*의 경우 Binary Cross Entropy Loss라고 정의
 - $BCE = -y \log(f_w(x)) - (1 - y) \log(1 - f_w(x))$

Model Training

Linear
Regression

Neural Network

Logistic
Regression

실습