회귀

Regression

머신러닝/딥러닝

임 경 태

Week	Chapter	Contents
1	1, 2장	강의 소개, 파이썬 복습
2	1, 3장	파이썬 복습, Numpy, Pandas
3	1, 4장	딥러닝을 위한 미분
4	5장	회귀
5	5장	분류
6	6장	XOR문제
7	7장	딥러닝
8	1~7장	중간고사
9	8장	MNIST 필기체 구현 (팀 프로젝트)
10	9장	오차역전파
11	11장	합성곱 신경망(CNN)
12	12장	순환 신경망(RNN)
13	10장	자율주행 (Collision Avoidanve, Transfer Learning)
14	11장	자율주행 (Load Following)
15	8~12장	기말고사 (or 프로젝트 발표)

CONTENTS

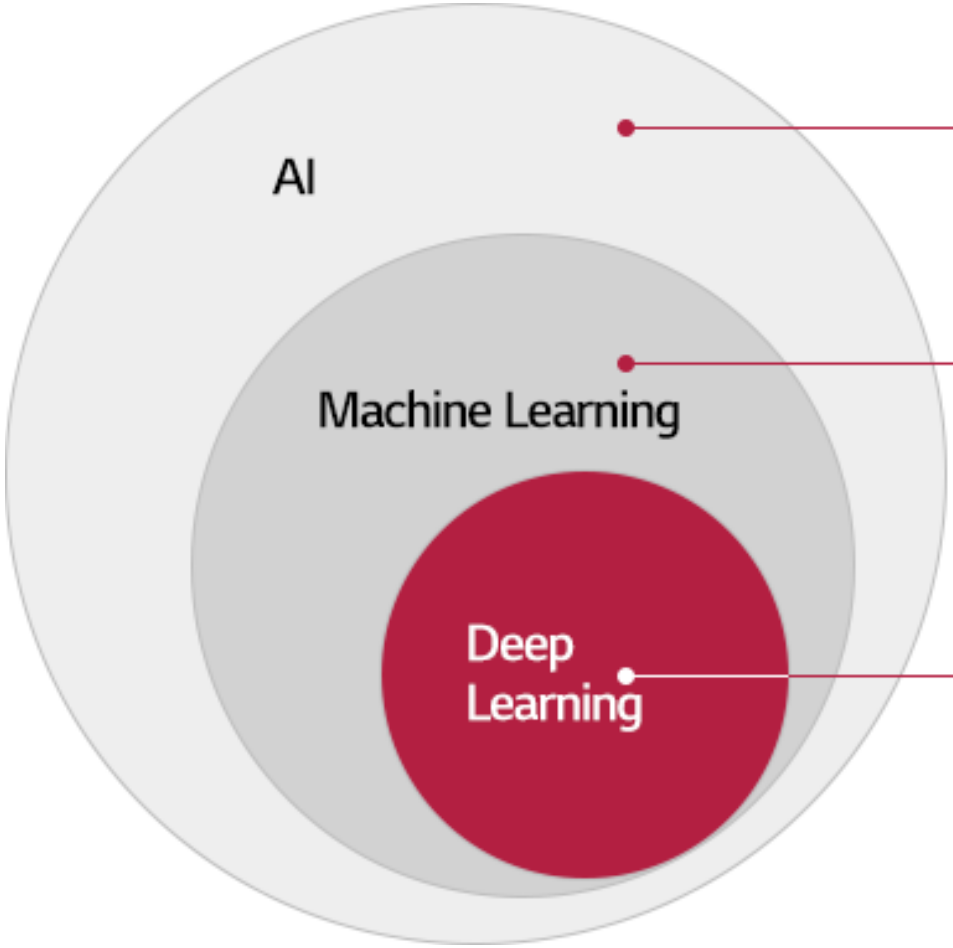
목적 : 지도학습의 대표적인 알고리즘인 선형회귀에 대한 학습

목표 : 선형 회귀와 경사 하강법에 대한 이해

내용 : 머신 러닝 개요, 선형 회귀, 경사 하강법

CONTENTS

머신 러닝 개요



인공지능(Artificial Intelligence)

사람이 해야 할 일을 기계가
대신할 수 있는 모든 자동화에 해당

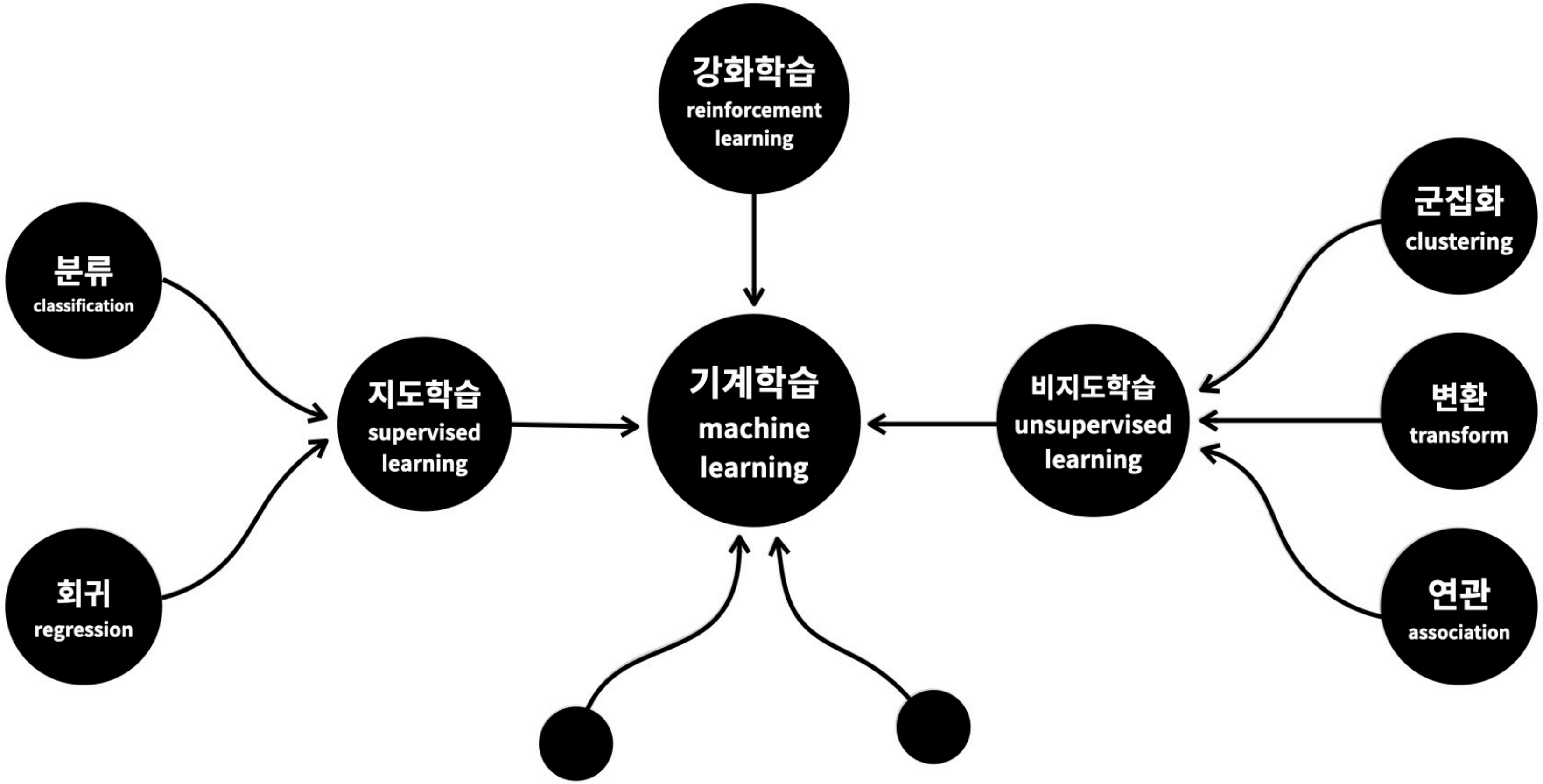
머신러닝(Machine Learning)

명시적으로 규칙을 프로그래밍하지 않고
데이터로부터 의사결정을 위한 패턴을
기계가 스스로 학습

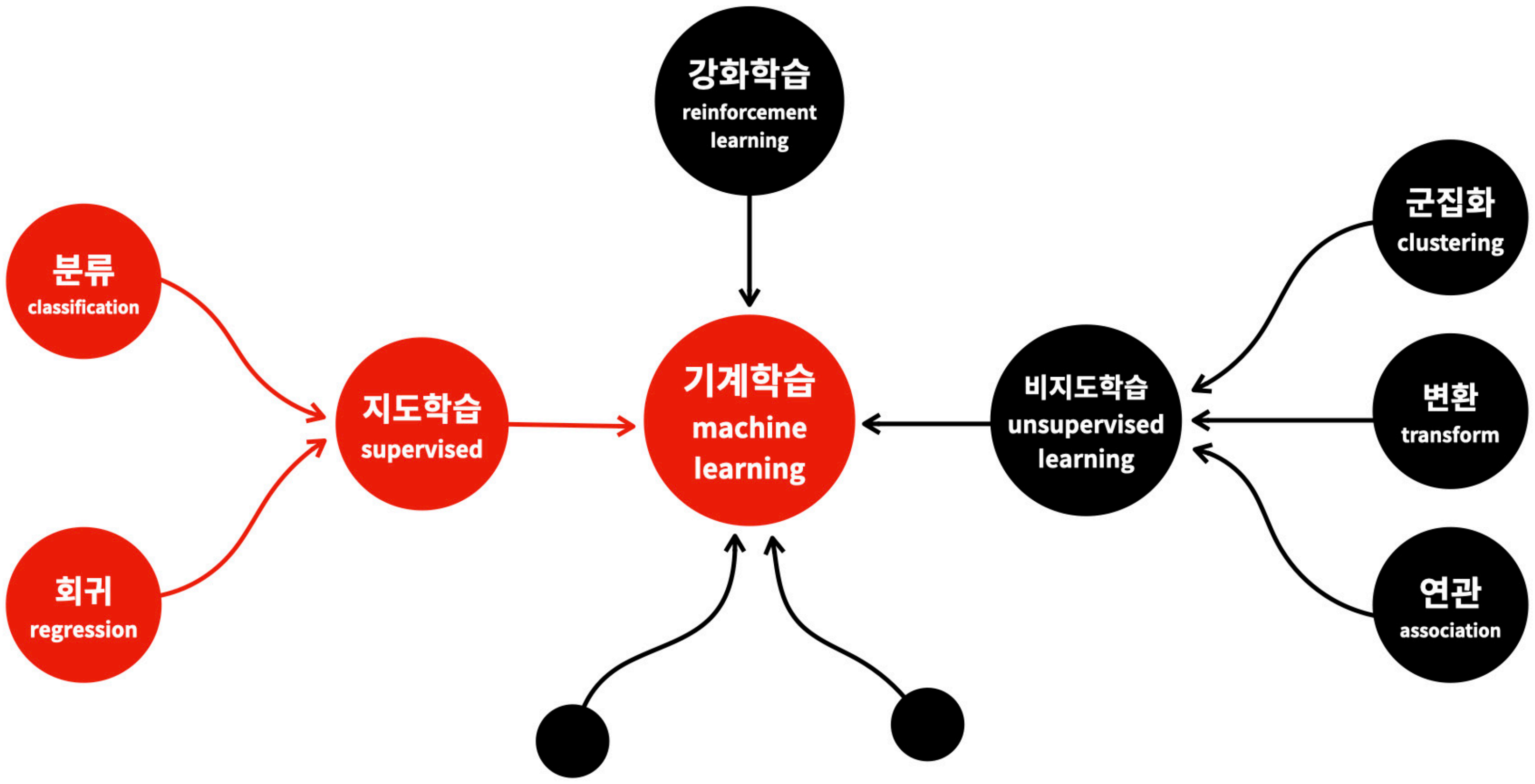
딥러닝(Deep Learning)

인공신경망 기반의 모델로, 비정형
데이터로부터 특징 추출 및 판단까지
기계가 한 번에 수행

머신 러닝 개요


강화학습
reinforcement learning
분류
classification
지도학습
supervised learning
기계학습
machine learning
비지도학습
unsupervised learning
군집화
clustering
변환
transform
연관
association
회귀
regression

머신 러닝 개요



강화학습
reinforcement learning
분류
classification
지도학습
supervised
기계학습
machine learning
비지도학습
unsupervised learning
군집화
clustering
변환
transform
회귀
regression
연관
association

머신 러닝 개요 - 정리

- 머신 러닝은 크게 지도학습, 비지도학습, 강화학습의 세가지로 정답 label 존재 여부, 학습 방법 등을 기준으로 나뉜다.

- 회귀와 분류는 정답 label이 존재하는 데이터를 학습하는 알고리즘으로 지도학습에 속한다.

CONTENTS

선형 회귀

- 회귀는 주어진 데이터를 통해 미지의 데이터를 연속적인 값으로 예측하는 알고리즘으로 일반적으로 여러 개의 입력을 갖는다.

간단한 회귀 예제 (입력이 한 개)

- 공부시간 → 시험 성적
- 1인당 GDP → 삶의 만족도 등

선형 회귀

- 선형회귀모델 : 출력변수 Y를 입력변수 X들의 선형결합으로 표현한 모델

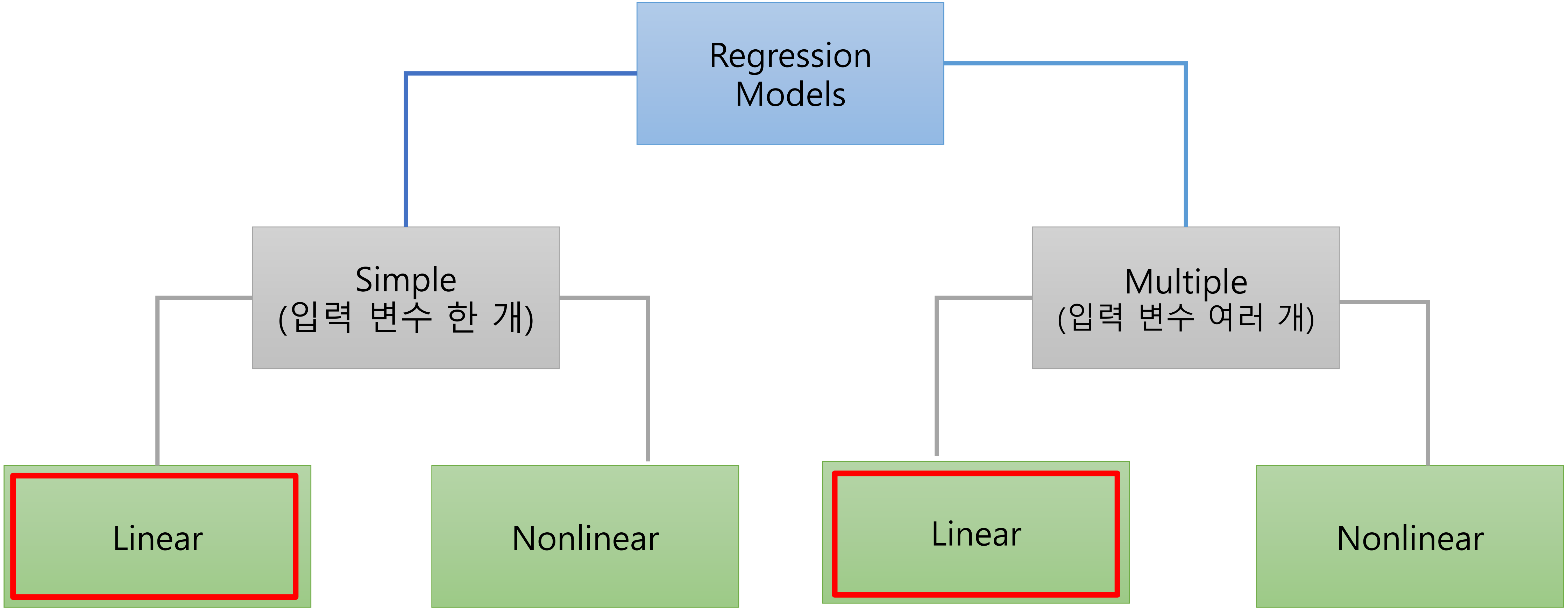
- 선형 결합 : 변수들을 상수 배, 덧셈을 통해 결합

ex) $w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3 + \cdots + w_n x_n$

w_i : 상수, x_i : 변수

특히 w_0는 편향이라고 하며 b(bias)로 표현함.

선형 회귀


Regression Models
Simple
(입력 변수 한 개)
Multiple
(입력 변수 여러 개)
Linear
Nonlinear
Linear
Nonlinear

선형 회귀

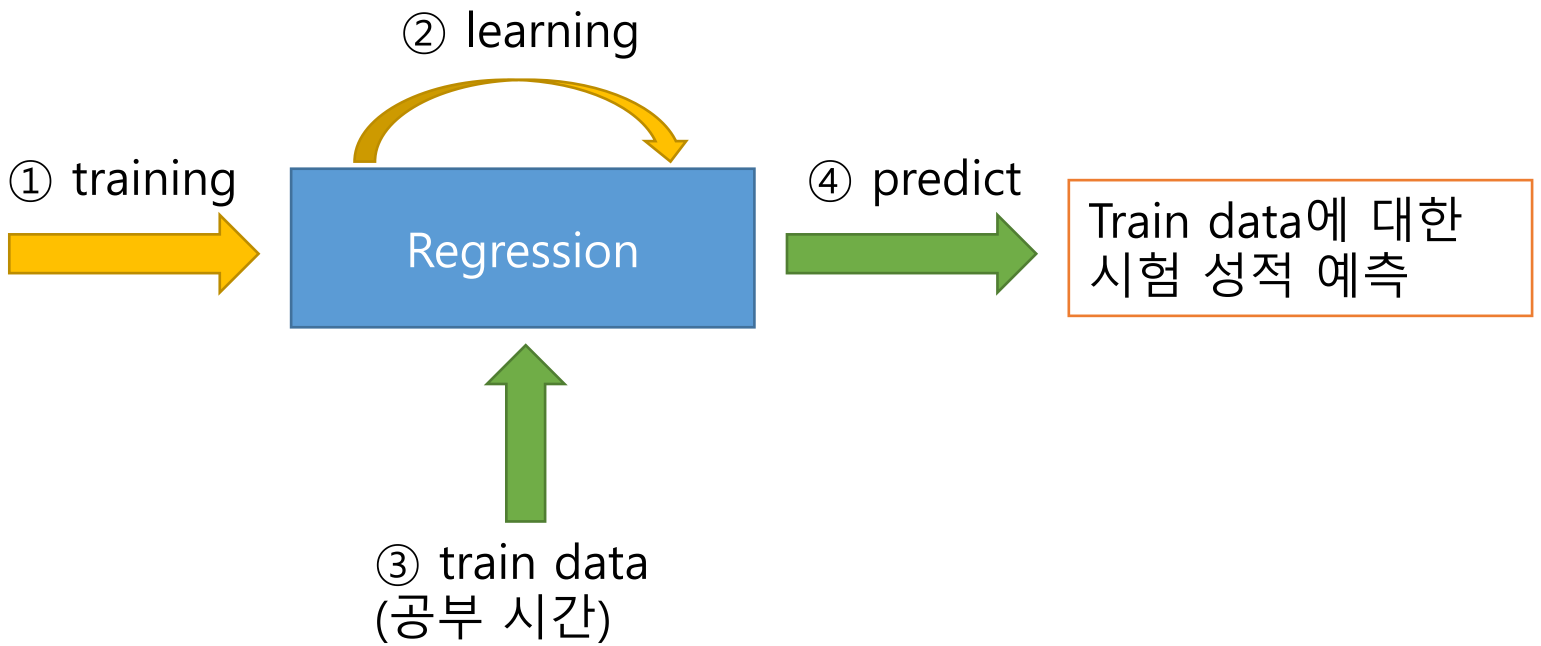
- $\hat{y} = \boldsymbol{wx} + \boldsymbol{b}$ - y 예측 값 (교재에서는t로 표현)

- $\boldsymbol{w} = [w_1\ w_2\ w_3 \cdots]^T$ - 가중치 벡터, 학습을 통해 최적값을 구함
- $\boldsymbol{x} = [x_1\ x_2\ x_3 \cdots]$ - 특성 벡터
- $\boldsymbol{b}$ - 편향

선형 회귀 Simple, Linear 예제

공부 시간 → 시험 성적 예제

이름	수학 공부시간	점수
야쓰오	2	3
오공	4	4
블리츠	6	5
페이커	8	6



선형 회귀 Simple, Linear 예제

DATA

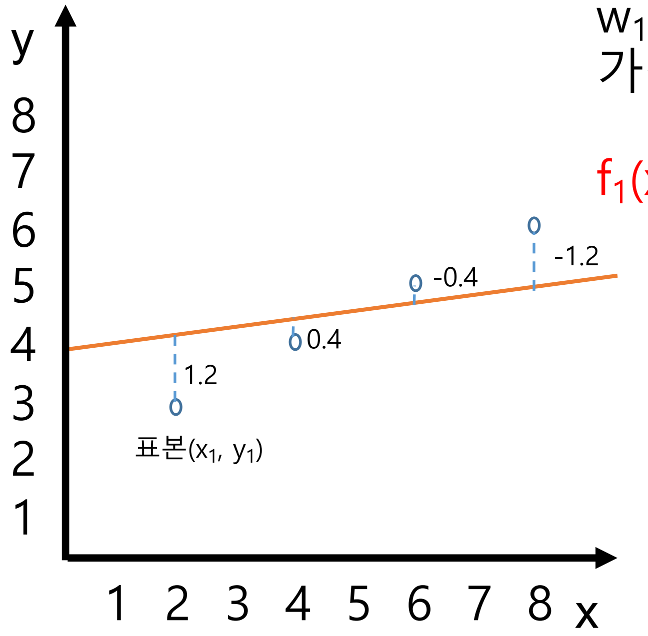
이름	수학 공부시간	점수
야쓰오	2	3
오공	4	4
블리츠	6	5
페이커	8	6

Find w and b → Predict $\hat{y}$

x : 공부 시간, y : 시험 성적
$\hat{y}$: 미지의 공부시간에 대해 예측한 시험 성적

❑ 선형회귀 평가

- 관측된 표본데이터와 우리가 선택한 w, b값으로 생성한 직선을 비교해 차이가 적은 w, b값을 찾자



$w_1 = 0.1$, $b_1 = 0.4$로 가정해보자

$f_1(x) = 0.1*x + 0.4$

어떤 평가척도를 사용해서 w, b를 평가할까?

- 관측된 표본(x_1, y_1)과 f1과의 **편차**$= (f_1(x_1) - y_1) = 1.2$
- 전체 편차의 합$= \Sigma_{i=1}(f_1(x_i) - y_i) = 0$??
- 전체 편차의 평균$= 1/n*\Sigma_{i=1}(f_1(x_i) - y_i) = 0$??
- 편차 제곱의 평균$= 1/n*\Sigma_{i=1}(f_1(x_i) - y_i)^2 = 0.8$

loss function (MSE, Mean Square Error)

- $MSE = \frac{1}{n}\sum_{i=1}^{n}\left((wx^{(i)} + b) - y^{(i)}\right)^2$

  $= \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{y^{(i)}} - y^{(i)}\right)^2$
- $\widehat{y^{(i)}}$: 예측한 결과
- $y^{(i)}$: 정답

- MSE를 최소로 하는 것이 목표 → 경사 하강법

선형 회귀 multiple, Linear 예제

- f(평수, 위치, 방의 개수,⋯) → 집값 예측
- f(나이, 흡연 기간, 가족력,⋯) → 폐암 예측
- f(성별, 나이, 수영 여부,⋯) → 생존 가능성 예측
- f(강우량, 일조량, 지형,⋯) → 농작물 가격 예측

CONTENTS

경사 하강법 (Gradient Descent)

- 최적의 매개변수(w)를 찾기 위한 알고리즘

최적의 매개변수란 손실함수를 최소로 만드는 매개변수

보통 매개변수 공간이 광대하여 어디에서 손실함수가 최소가 되는지 짐작 불가능

→ 기울기(gradient)를 이용하여 손실함수의 최솟값을 찾음

복습 – Gradient

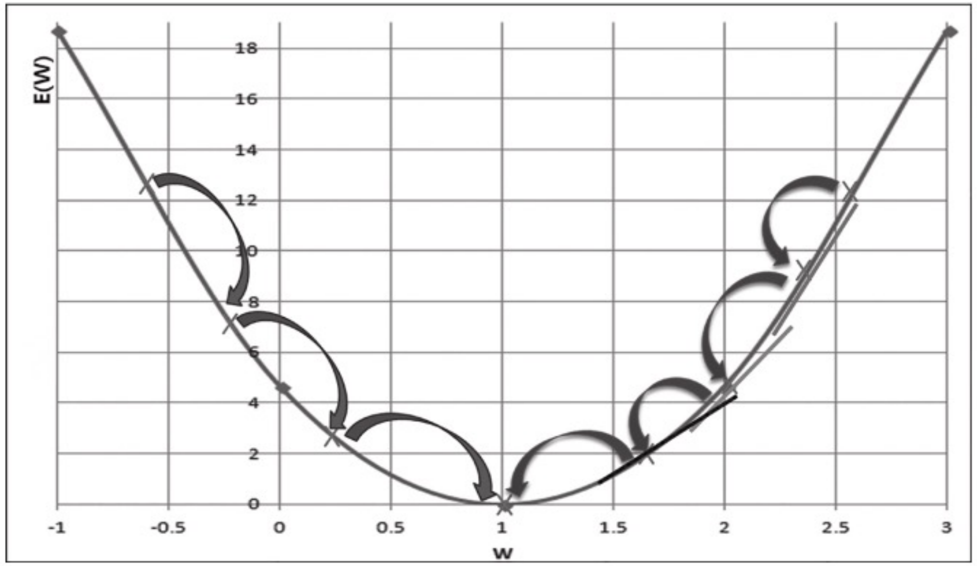
- Gradient란 최대의 증가율을 나타내는 벡터.

- Gradient가 최대의 증가율을 나타내기 때문에 Gradient의 반대방향을 따라가면 Local minimum에 가장 빠르게 도달한다.

- 즉 벡터 $\nabla f(x, y, z) = \left\langle \frac{\partial f(x,y,z)}{\partial x}, \frac{\partial f(x,y,z)}{\partial y}, \frac{\partial f(x,y,z)}{\partial z} \right\rangle$ 는 $f(x, y, z)$에서 가장 가파른 방향을 나타냄.

경사 하강법 (Gradient Descent)

Simple, Linear 예제의 w, $E(w)$그래프



[그림 5.12] 경사하강법(Gradient Descent Algorithm) 원리

- 경사 하강법

1. 임의의 가중치 w를 선택 (random)

2. 선택된 w에서의 기울기 $\frac{\partial E}{\partial w}$를 구함

(여기서 E는 loss function, $\frac{\partial E}{\partial w}$는 w가 변할 때 E의 변화량을 의미)

3. 미분 값 $\frac{\partial E}{\partial w}$가 작아지는 방향으로 w를 업데이트하면 local minimum에 도달

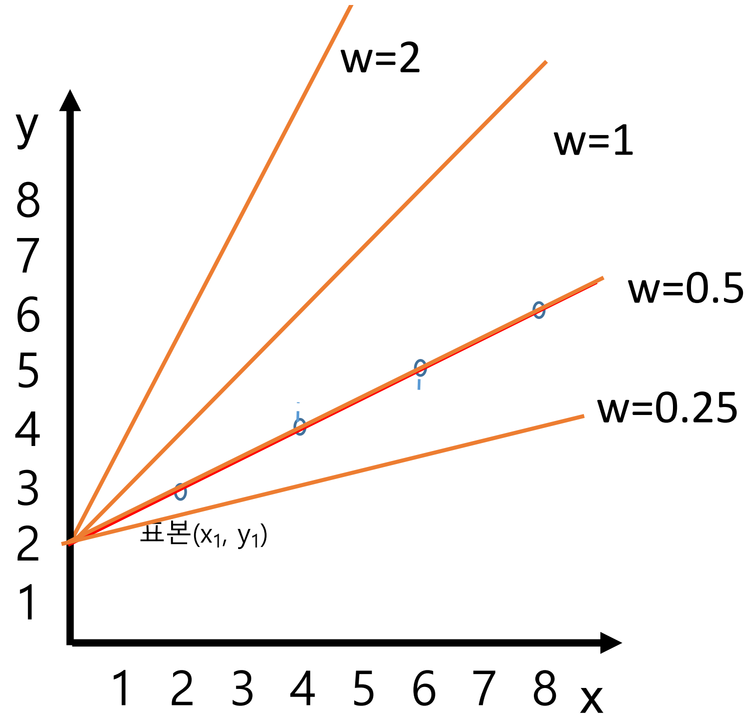
선형회귀 오류 피드백

$$\underset{W,b}{\text{minimize}}\, cost(W, b)$$

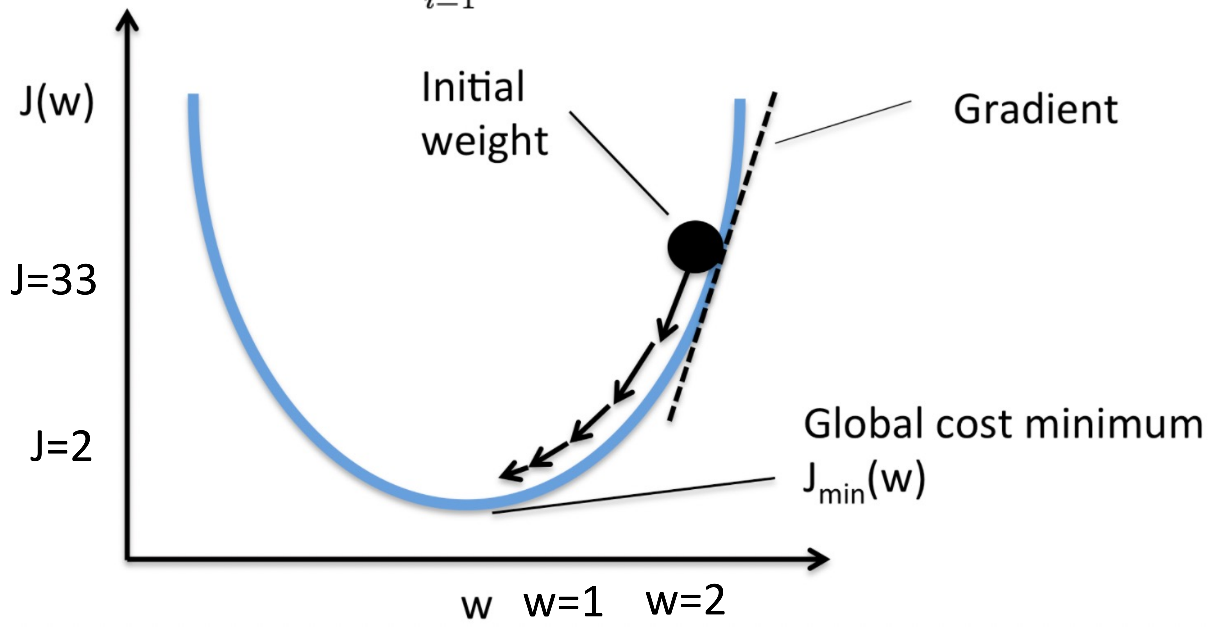
❑ Gradient Descent Algorithm

- 미분을 이용해 cost값이 작아지는 방향으로 조금씩 w값을 업데이트 하자!!

$$W := W - \alpha \frac{\partial}{\partial W} cost(W)$$



$$cost(W) = \frac{1}{m} \sum_{i=1}^{m} (Wx^{(i)} - y^{(i)})^2$$



이미지출처: https://sebastianraschka.com/faq/docs/closed-form-vs-gd.html

경사 하강법 (Gradient Descent)

- 경사 하강법 공식

$$w \leftarrow w - \alpha \frac{\partial E}{\partial w}$$

α : 학습률 (learning rate), 한 번의 학습으로 얼마만큼 매개변수 값을 갱신할지 결정

$-\alpha \frac{\partial E}{\partial w}$: gradient의 반대방향으로 학습

감사합니다.