Machine Learning

Learn Machine Learning with Microsoft Azure ML Studio

Microsoft Student Partner
고희원 김다영 강민주 구노영

Machine Learning

✓Machine Learning 의 등장

- Limitations of Explicit Programming
- Explicit Programming : 개발자가 입력 조건과 프로그램 상태 조건에 따라 소프트웨어가 동작하는 방식을 직접 구현하는 프로그래밍 방식
- Complex Problem : Spam mail filtering, Autonomous vehicle?

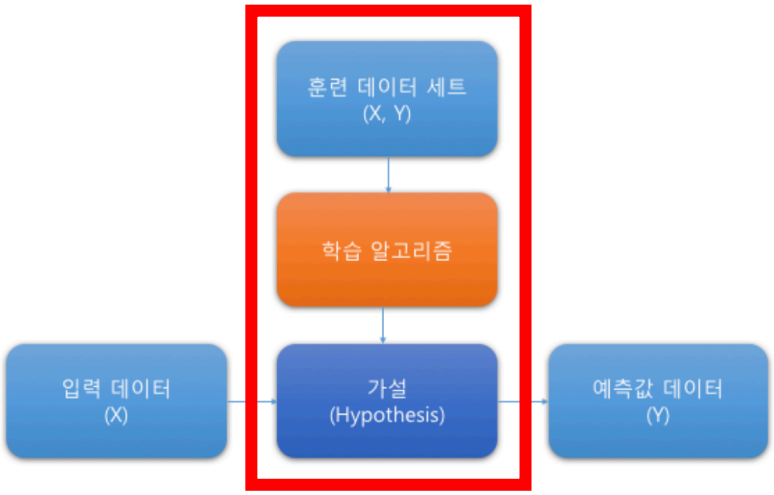
✓Machine Learning 이란?

- “ Machine Learning is a field of computer science that gives computers the ability to learn without being explicitly programmed.” By. Arthur Samuel (1959)

Machine Learning

✓용어

- **Training Data Set**
  (X,Y) = (메일 내용, 스팸 여부)
- **Input Data**
  X = 새로운 메일 내용
- **Predicted Value**
  Y = 스팸 여부
- **Hypothesis (Model)**
  f(X) = Y
- **Target function**



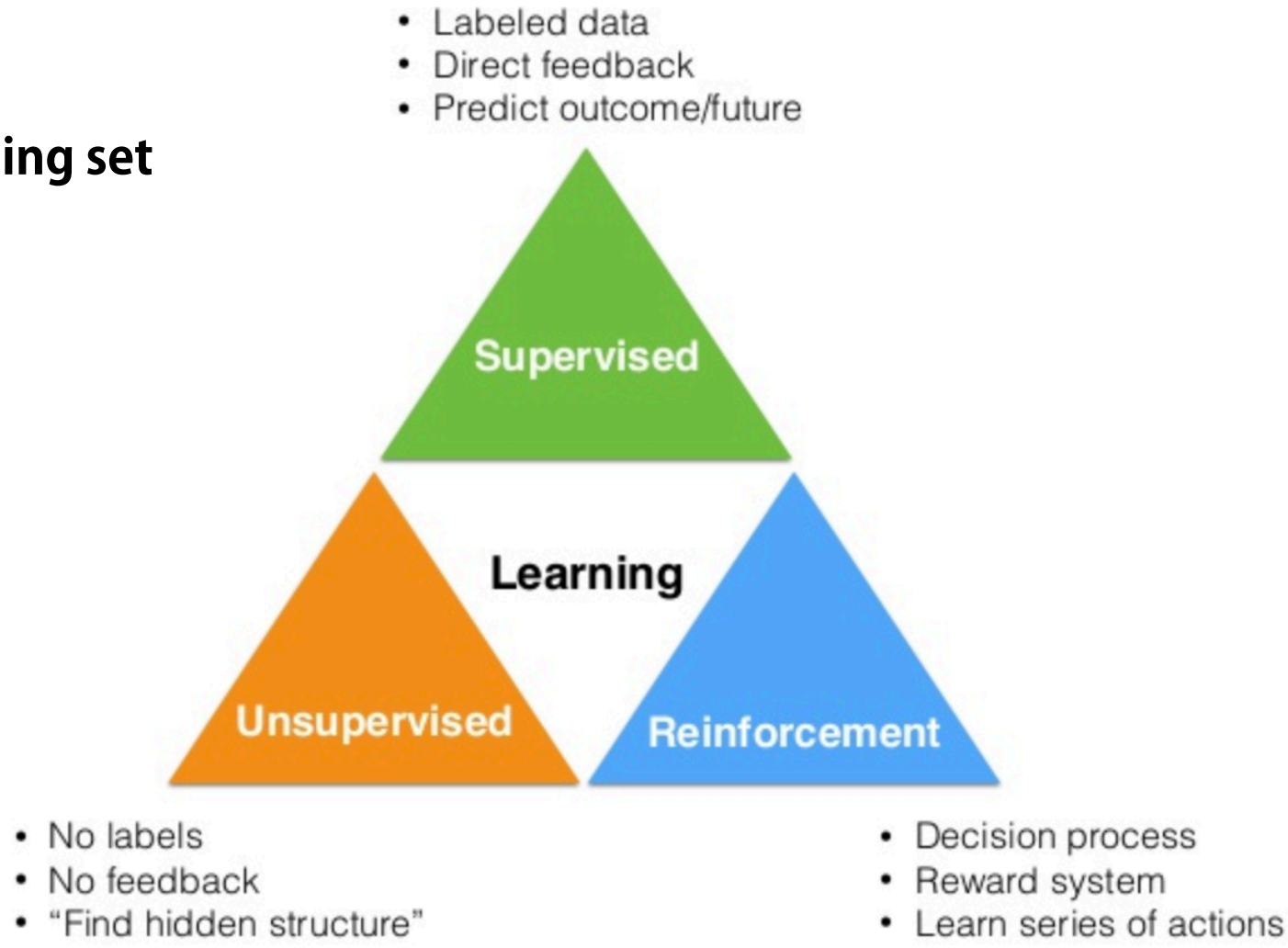
✓ 오차(Error) : 예측값과 목표값의 차이

✓ 오차를 수정하는 방식

- 최소제곱법(method of least squares — LSM or LMS)
- 최우추정법(maximum likelihood method — MLE)
- 최대사후확률추정(maximum a posterior estimation — MAP)
- 기울기 하강법(gradient descent — GD)
- 역전파(backpropagation — BP)

Machine Learning

- ✓ **Supervised Learning**
  - learn with labeled examples - training set
- ✓ **Unsupervised Learning**
  - learn with unlabeled data
- ✓ **Semi-supervised Learning**
  - Supervised + Unsupervised
- ✓ **Reinforcement Learning**
  - reward feedback required by each behavior



Machine Learning

✓Supervised Learning

✓Classification (For Discrete Value)

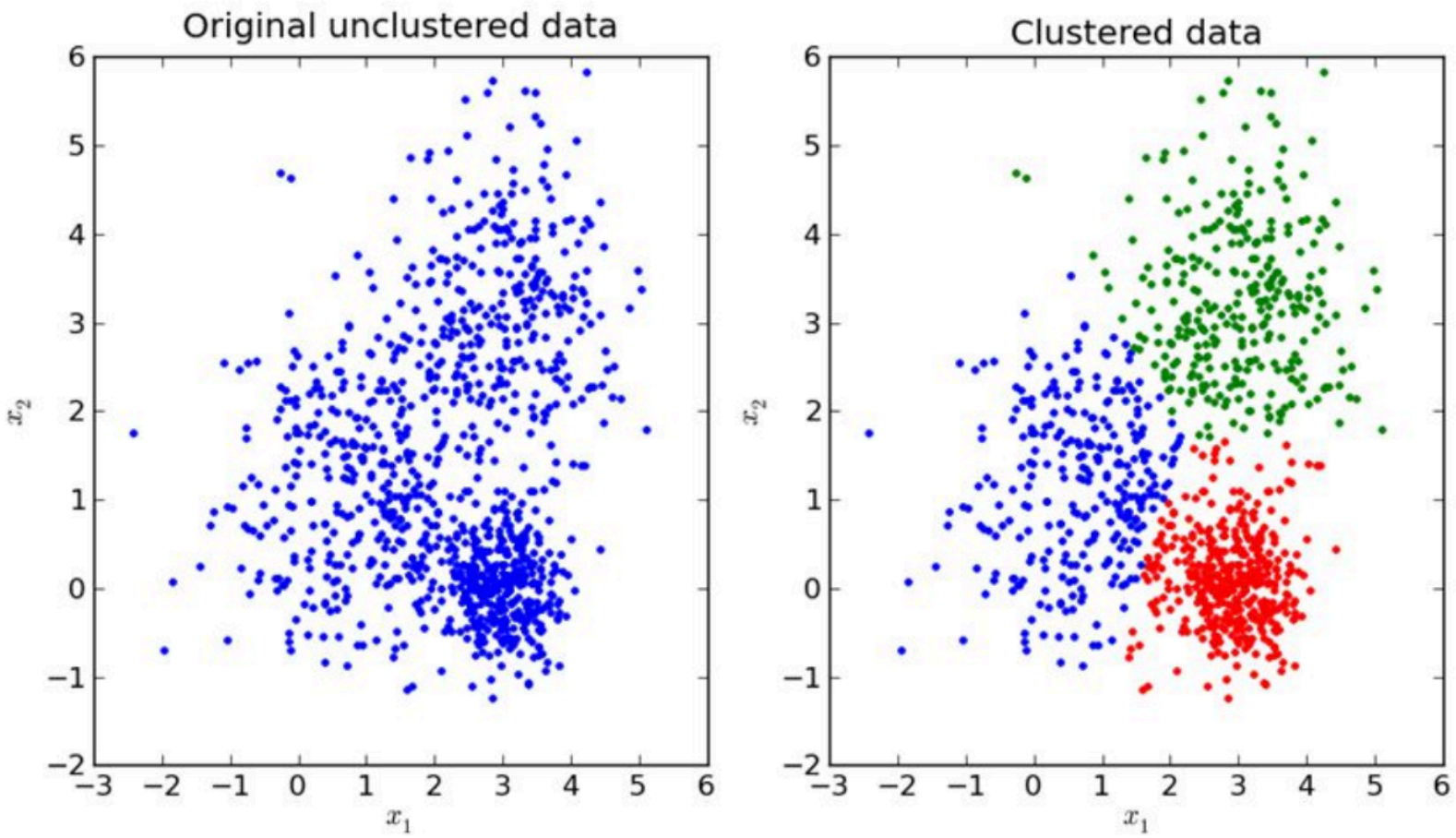
✓Regression (For Continuous Value)

✓Deep Learning Algorithms
- Convolutional Neural Network(CNN)
- Recurrent Neural Network(RNN)

✓Examples
- Number Recognition
- Spam mail filtering

Machine Learning

✓Unsupervised Learning



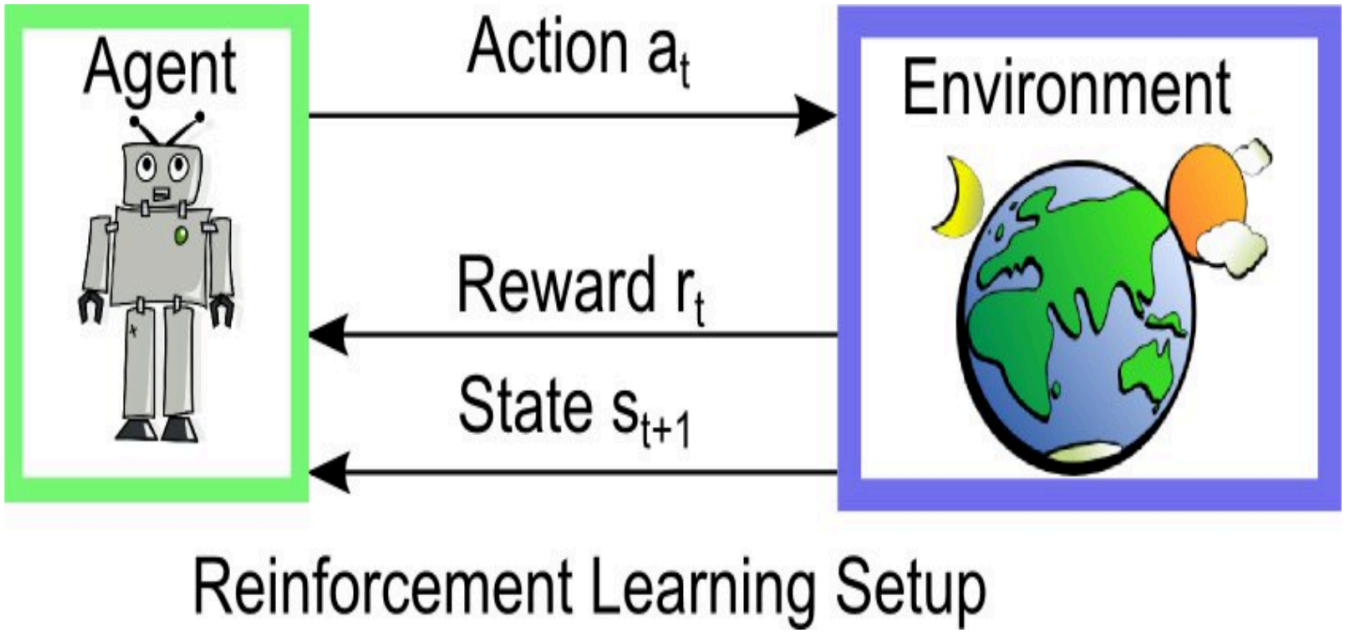
✓Clustering

✓Deep Learning Algorithms
- AutoEncoders

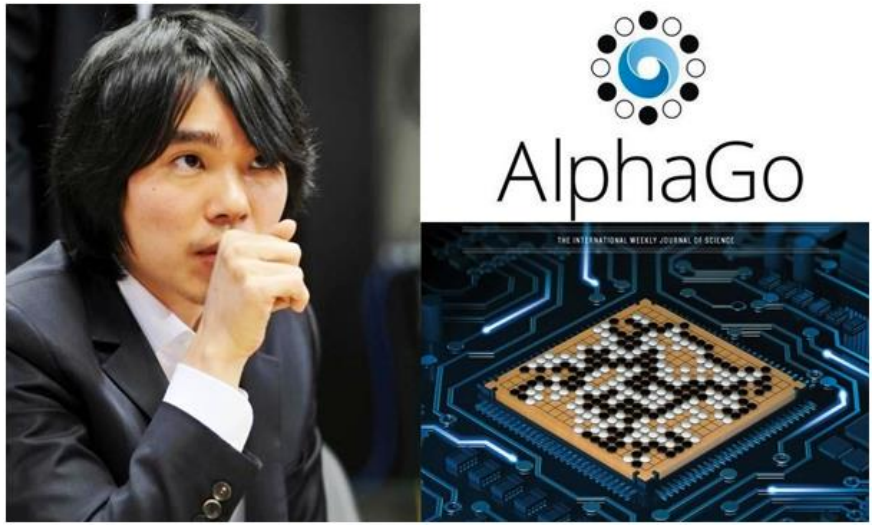
✓Examples
- Google News Clustering
- Word Clustering

Machine Learning

✓Reinforcement Learning



Reinforcement Learning Setup



✓Deep Learning Algorithms

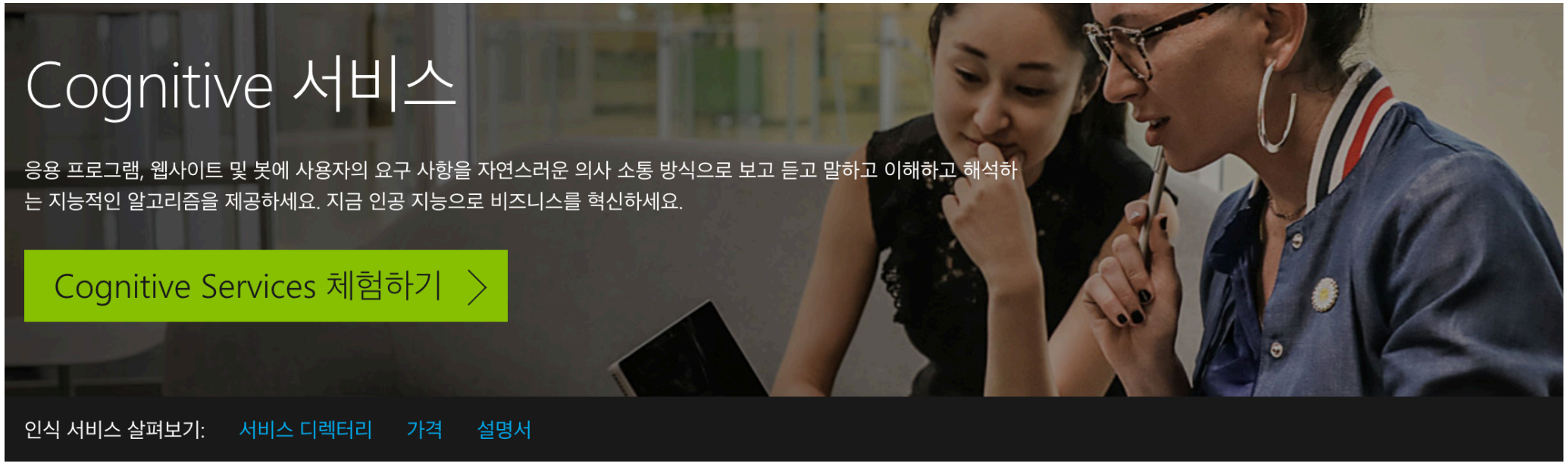
- Q-learning
- Deep-Q-Nework(DQN)

✓Examples

- “AlphaGo”
- Autonomous Car driving

Machine Learning

✓Microsoft Cognitive API



인공지능을 사용하여 비즈니스 문제 해결

시각

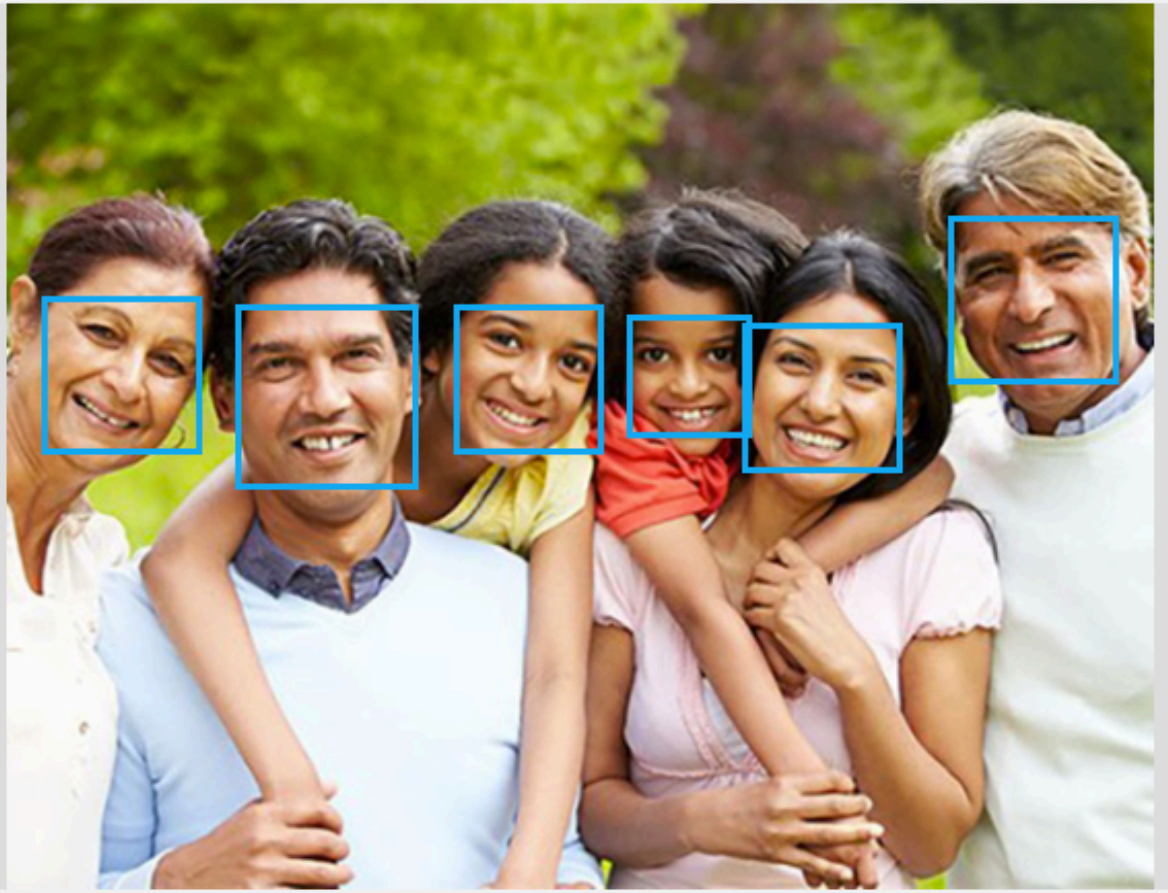
이미지 처리 알고리즘을 통해 사진을 스마트하게 인식, 캡션 작성, 조정합니다.

음성

음성을 텍스트로 변환하고 신원 확인에 음성을 사용하거나 앱에 화자 인식 기능을 추가합니다.

Machine Learning

✓Microsoft Cognitive Face API

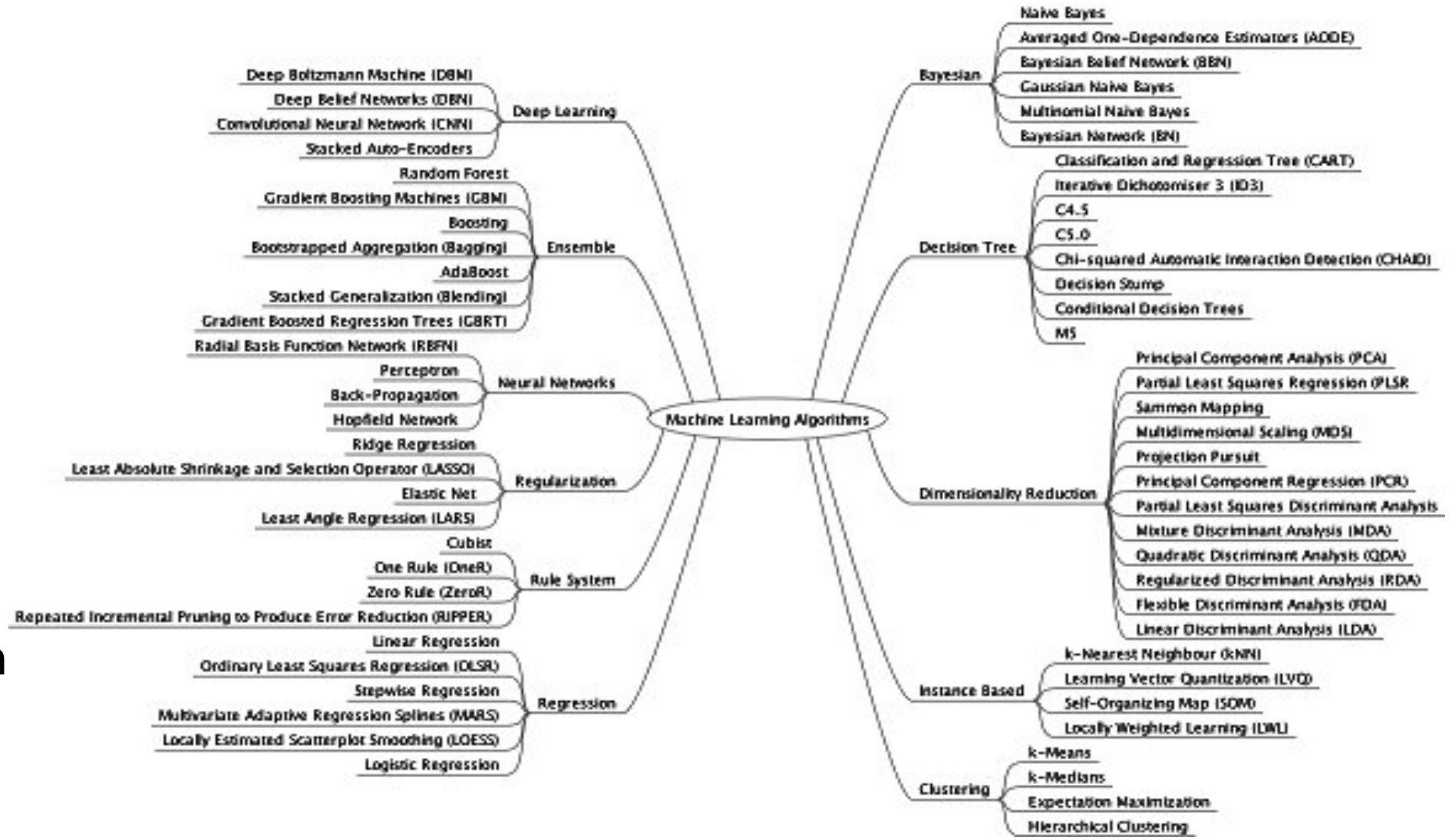


```
Detection result:
6 faces detected

JSON:
[
  {
    "faceRectangle": {
      "top": 159,
      "left": 117,
      "width": 95,
      "height": 95
    },
    "scores": {
      "anger": 2.80264025E-08,
      "contempt": 6.347948E-08,
      "disgust": 9.236748E-06,
      "fear": 2.52110336E-15,
      "happiness": 0.999983549,
      "neutral": 7.132026E-06,
      "sadness": 2.71582626E-12,
      "surprise": 5.42829448E-10
    }
  },
```

Machine Learning

✓Algorithms



✓Linear Regression
✓Logistic Regression
✓Neural Network
✓ ...

Linear Regression

Linear Regression : What is Linear Regression?

✓Linear Regression 이란?

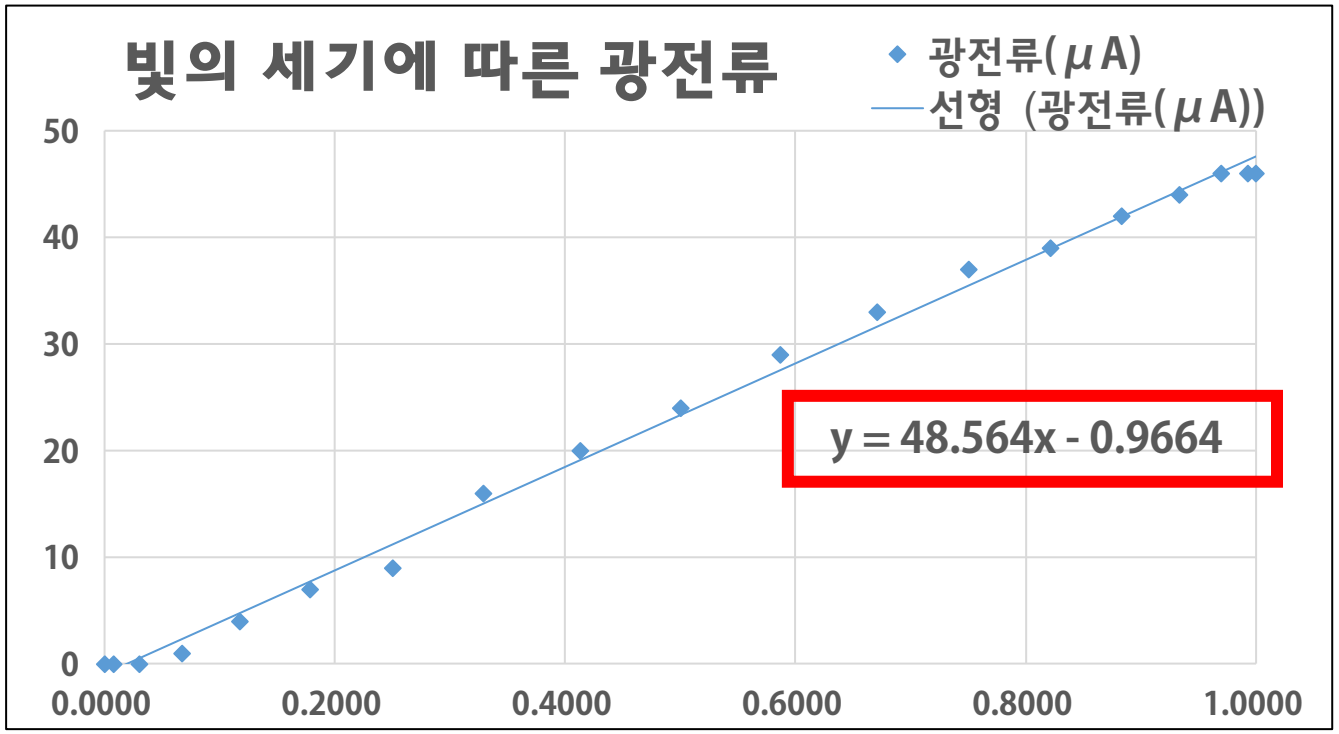
- 종속 변수(y)와 독립 변수(x)간의 관계가 선형적이라는 가정 하에, 이 둘의 상관 관계를 선형적으로 모델링
- 선형적(linear) : 그래프로 그렸을 때 직선 형태로 나타나는 관계 (Ex> 힘의 크기와 가속도의 크기)

✓용어

- Independent variable : input value(x)
- Dependent variable : predicted output(y) for the set of input values
- Parameter(or Weight) : scale factor(θ) for input value

Linear Regression : What is Linear Regression?

빛의 세기(상대값)	광전류(μA)
1.0000	46
0.9924	46
0.9698	46
0.9330	44
0.8830	42
0.8214	39
0.7500	37
0.6710	33
0.5868	29
0.5000	24
0.4132	20
0.3290	16
0.2500	9
0.1786	7
0.1170	4
0.0670	1
0.0302	0
0.0076	0
0.0000	0



- **Dependent variable : 빛의 세기(상대값)**
- **Independent variable : 광전류(μA)**
- **Weight:** $\theta_1 = 48.564, \theta_0 = -0.9664$
- **model : y = 48.564x - 0.9664**

Linear Regression : Simple Linear Regression

✓Simple Linear Regression : 하나의 독립 변수

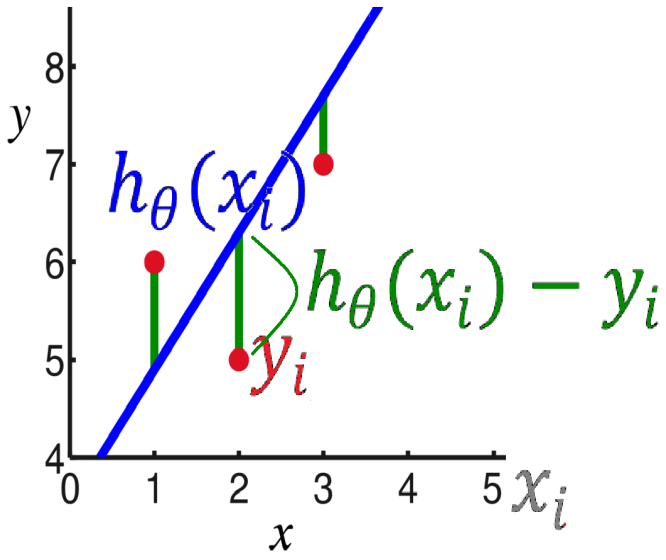
- **n개의 training set이 있다고 할 때,** $h_\theta(x_i) = \hat{y}_i = \theta_0 + \theta_1 x_i (\mathrm{i} = 1,\ 2,\ \ldots\ ,\ \mathrm{n})$ **꼴의 관계임을 가정할 수 있다.**

✓Cost function (loss function) : 회귀식의 효율성 검증

- $\boldsymbol{cost} = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2$

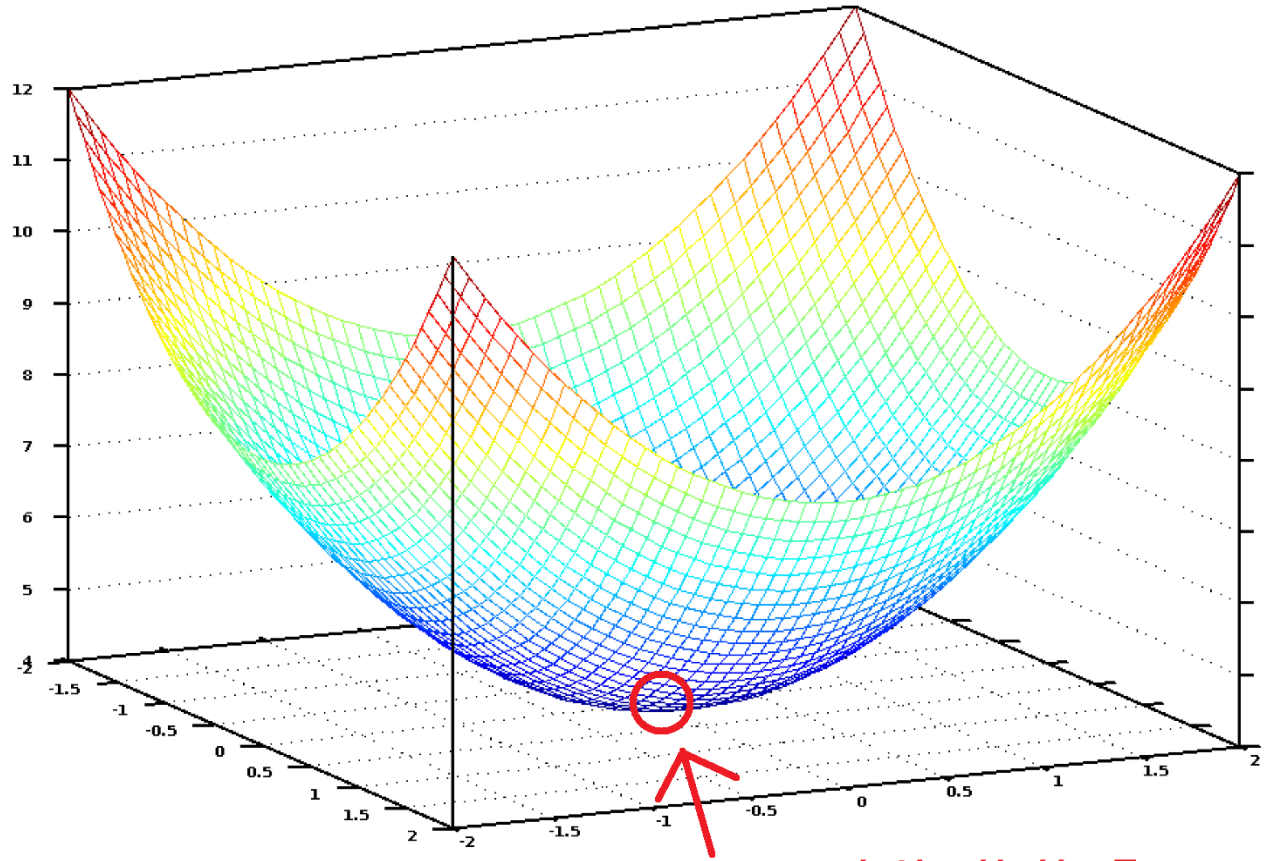
$$= \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x_i) - y_i)^2$$

- ***Goal : minimize the cost function***



Linear Regression : How to minimize the Cost Function

✓Graph of Cost Function : $z = J(\theta_0, \theta_1)$



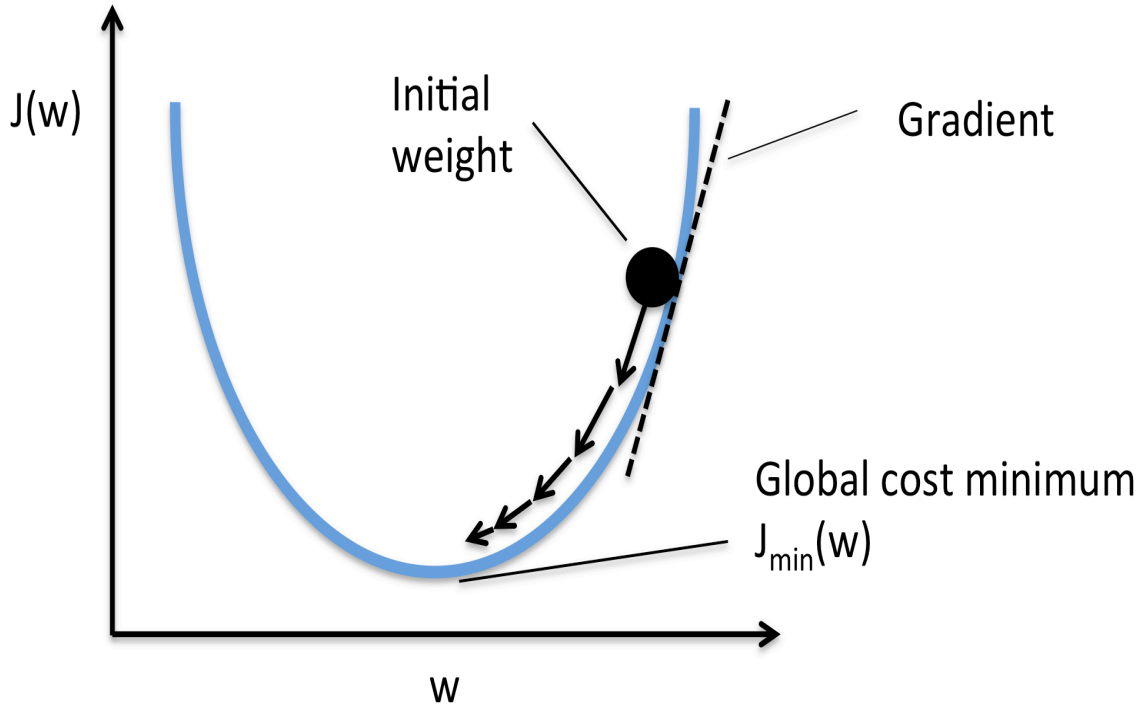
cost function이 최소화되는 곳

= 가장 효율적인 회귀식

Linear Regression : How to minimize the Cost Function

✓Gradient Descent

$$\theta_i := \theta_i - \alpha \frac{\partial}{\partial \theta_i} J(\theta_0, \theta_1) \; for \; i = 0, i = 1$$



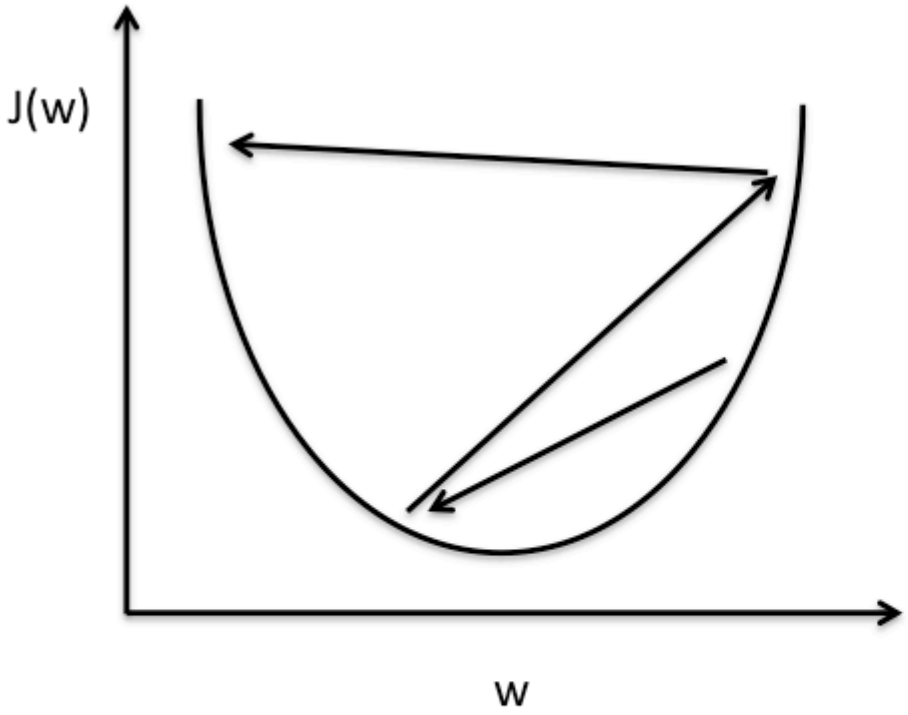
- $\alpha(> 0)$ **"learning rate"**
  : **local minimum**을 향해 가는 속도

- $\frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1) = \frac{\partial}{\partial \theta_0} \left[\frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2 \right] = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)$
- $\frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) = \frac{\partial}{\partial \theta_1} \left[\frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2 \right] = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i) x_i$

: 이동할 방향과 크기

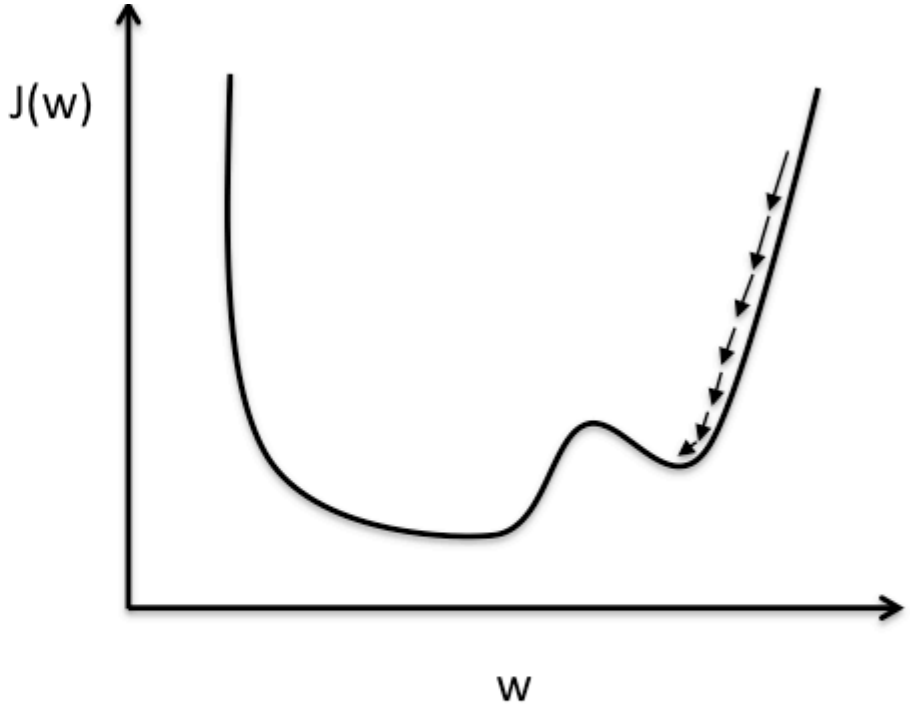
Linear Regression : How to minimize the Cost Function

When Learning Rate is Too Big



Large learning rate: Overshooting.

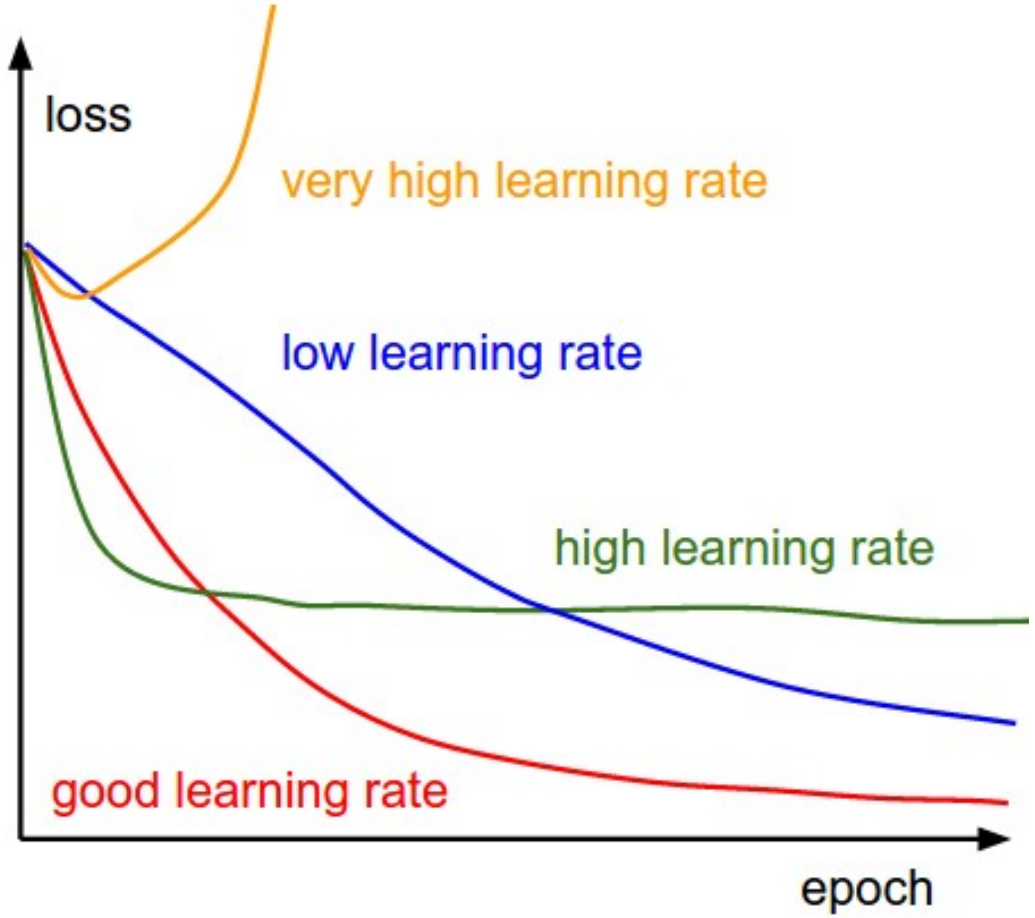
When Learning Rate is Too Small



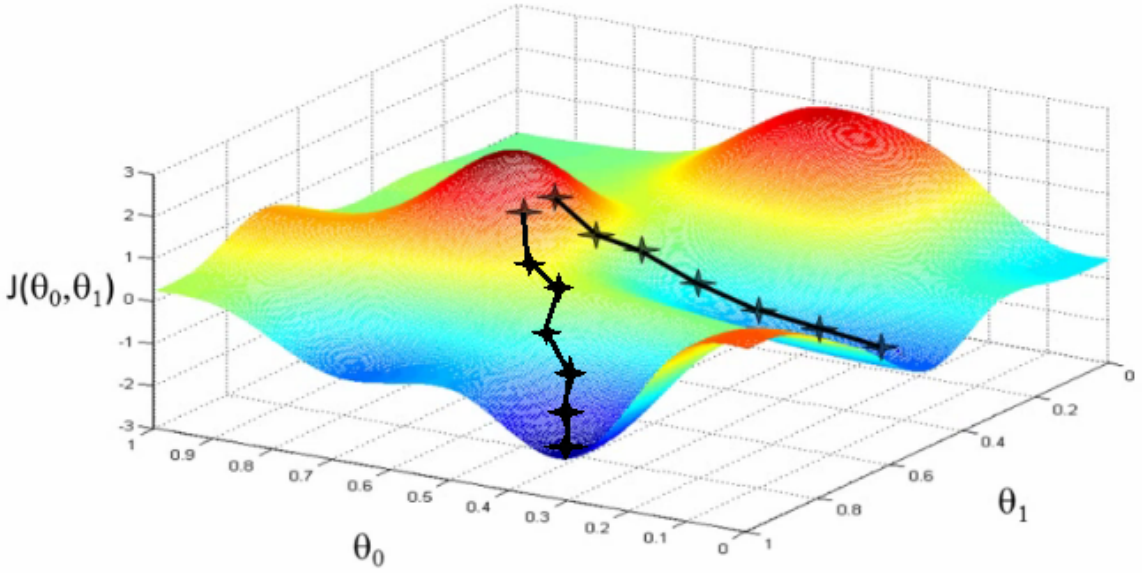
Small learning rate: Many iterations until convergence and trapping in local minima.

Linear Regression : How to minimize the Cost Function

Learning Rate



Local minimum vs Global minimum



✓Linear regression의 cost function은 Convex function!

➢Local min. = Global min. = optimal value

Linear Regression : Multiple Linear Regression

✓**Multiple Linear Regression : 여러 개의 독립 변수**

✓**Model**

$$h_\theta(x) = \hat{y}_i = \theta_0 x_{i0} + \theta_1 x_{i1} + \cdots + \theta_p x_{ip}\ , \forall x_{i0} = 1\ (i = 1,2, \ldots, n)$$

✓**Training set**

$$Let\ \mathbb{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \hat{\mathbb{Y}} = \begin{bmatrix} h_\theta(x_1) \\ h_\theta(x_2) \\ \vdots \\ h_\theta(x_n) \end{bmatrix}, \mathbb{X} = \begin{bmatrix} {x_1}^T \\ {x_2}^T \\ \vdots \\ {x_n}^T \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}, \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{bmatrix}$$

$$Then\ h_\theta(\mathbb{X}) = \hat{\mathbb{Y}} = \mathbb{X}\theta$$

✓**Cost Function :** $\mathrm{J}(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x_i) - y_i)^2$

✓**Gradient Descent :** $\theta := \theta - \alpha\frac{1}{m}\mathbb{X}^T(\mathbb{X}\theta - \mathbb{Y})$

Logistic Regression

Logistic Regression : What is logistic regression?

✓ Logistic Regression이란?

- 선형 회귀는 종속변수가 다양. 로지스틱 회귀는 종속 변수가 범주형 (이향형 문제)
- 데이터를 통해 Binary로 1 또는 0으로 분류 하는 학습 –〉 binary Classification

✓ 용어

- 독립변수: 입력값이나 원인 (예측변수, 복귀자, 통제변수)
- 종속변수: 결과물이나 효과 (설명변수)

Logistic Regression : What is logistic regression?

✓ Linear Regression

Output: multiple

✓ Logistic Regression

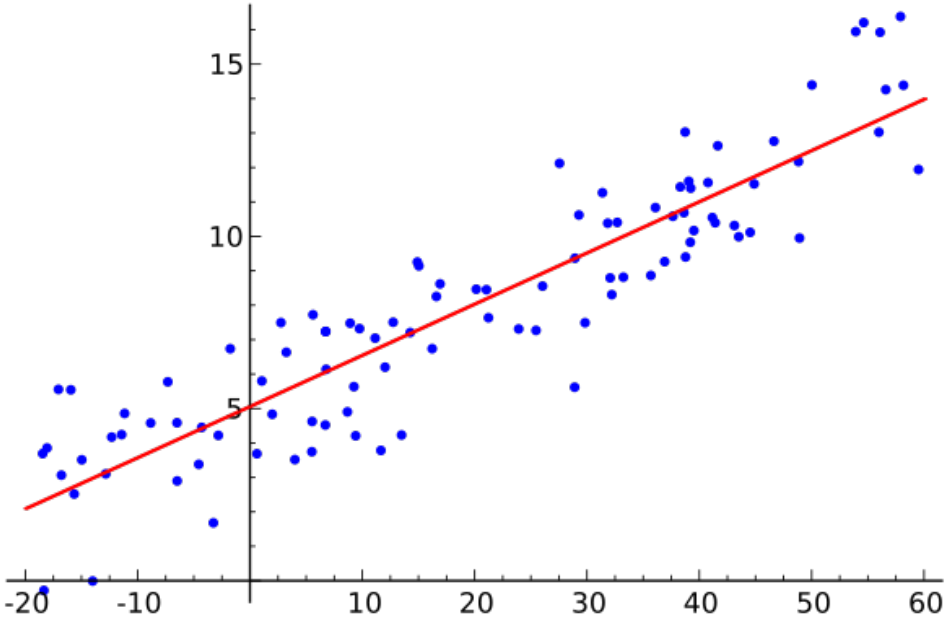
- Output: 0 or 1

(binary classification)

- Output: 0 or 1 or 2

(multi-class classification)

Linear Regression



Example of simple linear regression, which has one independent variable

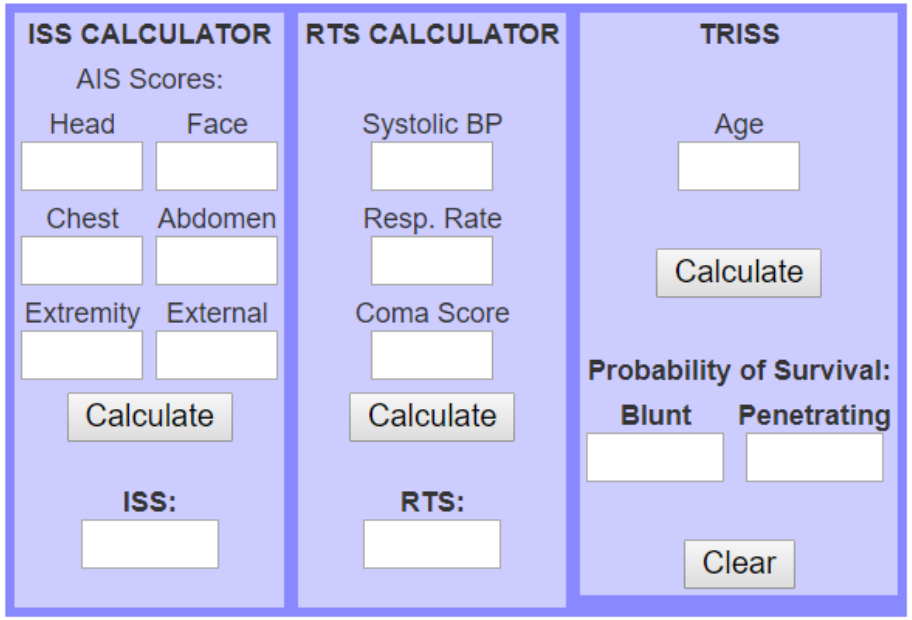
Logistic regression : What is logistic regression? (binary classification)

✓ 독립 변수가 [−∞,∞]의 어느 숫자이든 상관 없이
종속 변수 즉 결과 값이 항상 범위 [0,1] 사이에 존재

X is ∞
Y is 0 or 1

✓ **Example**

- Spam Detection: spam(1) or Ham(0)
- Facebook feed: show(1) or hide(0)
- 부상을 입은 환자들의 사망 예측을 위해 사용되는 Trauma and Injury Severity Score (TRISS)
- 이미지를 분석하여 lion dog cat 분류



(Injury Severity Score - Revised Trauma Score - Trauma Injury Severity Score)

Logistic regression : Sigmoid

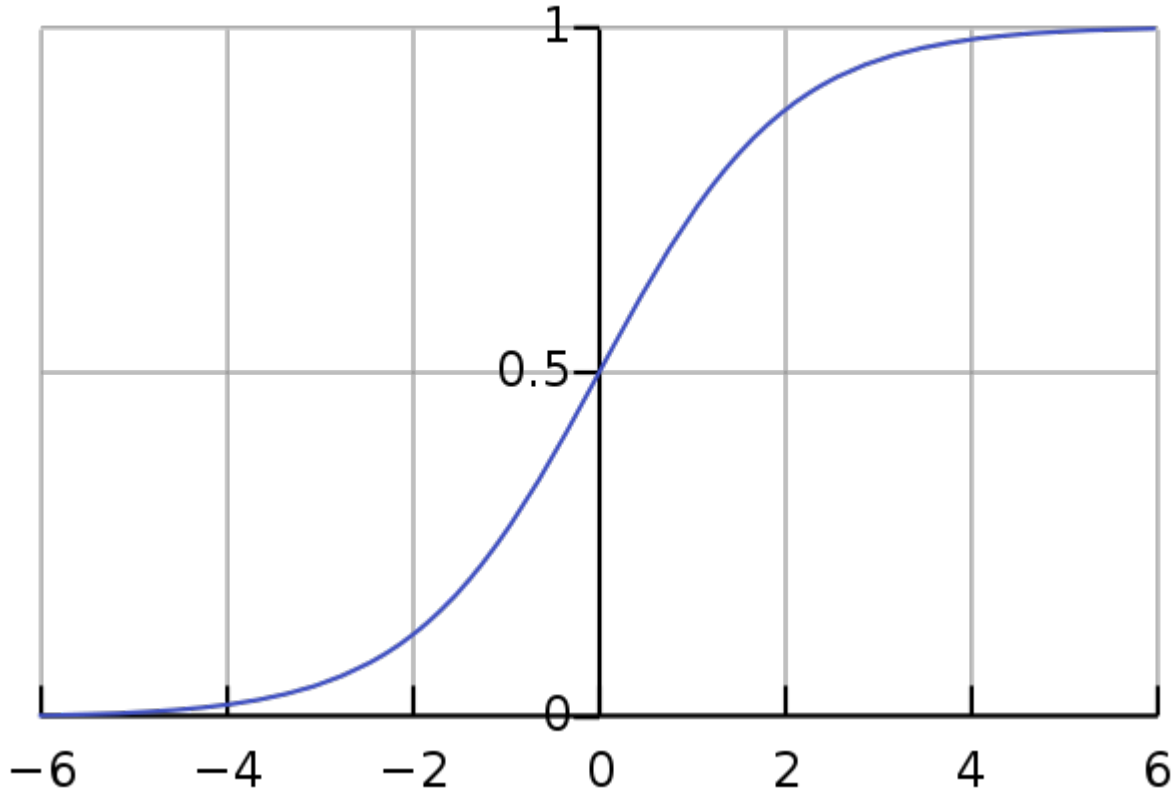
X is ∞
Y is 0 or 1

sigmoid:

Curved in two directions, like the letter "S", or the Greek ς (sigma).

-〉 모든 독립변수에 대해
종속변수는 0과 1사이의 값이 출력

Logistic Regression Hypothesis



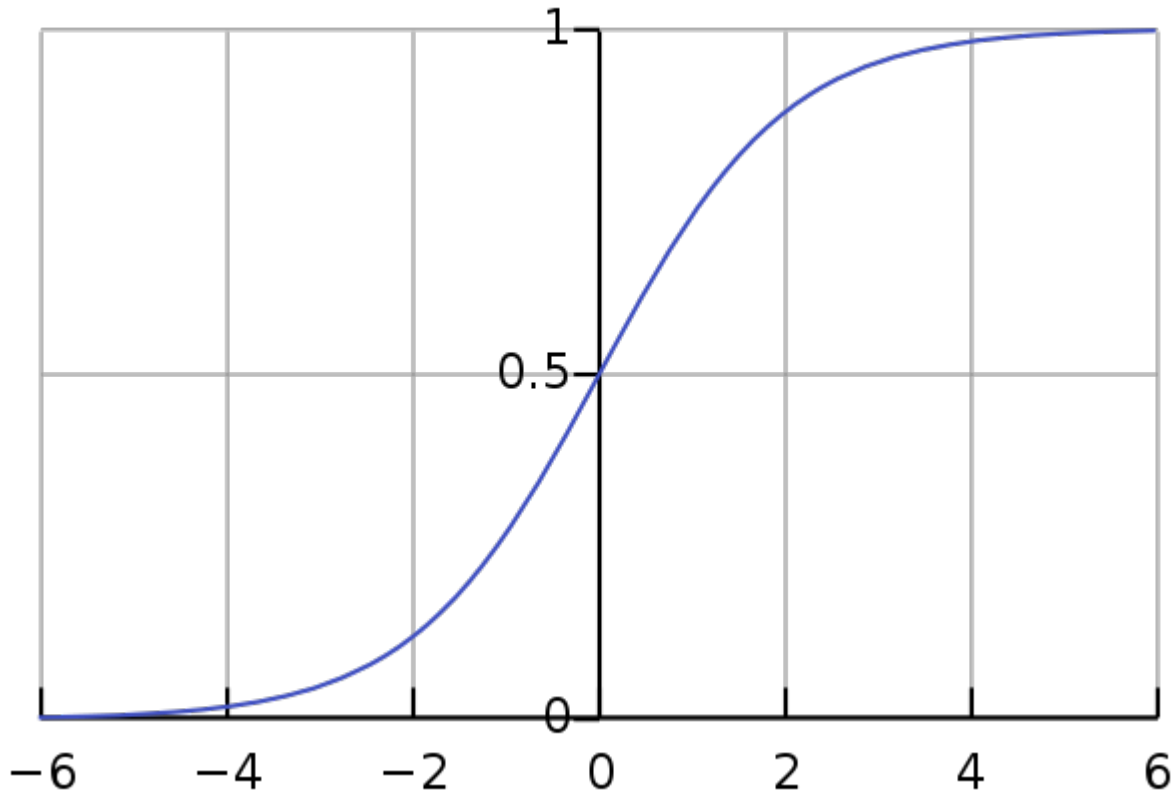
Logistic regression : Logistic Hypothesis

- ✓ **Logistic Hypothesis(Sigmoid)**

$$H(X) = \frac{1}{1 + e^{-W^T X}}$$

- ✓ **Suppose predict if H(x) >= 0.5 "y = 1"**
  predict if H(x) < 0.5 "y = 0"

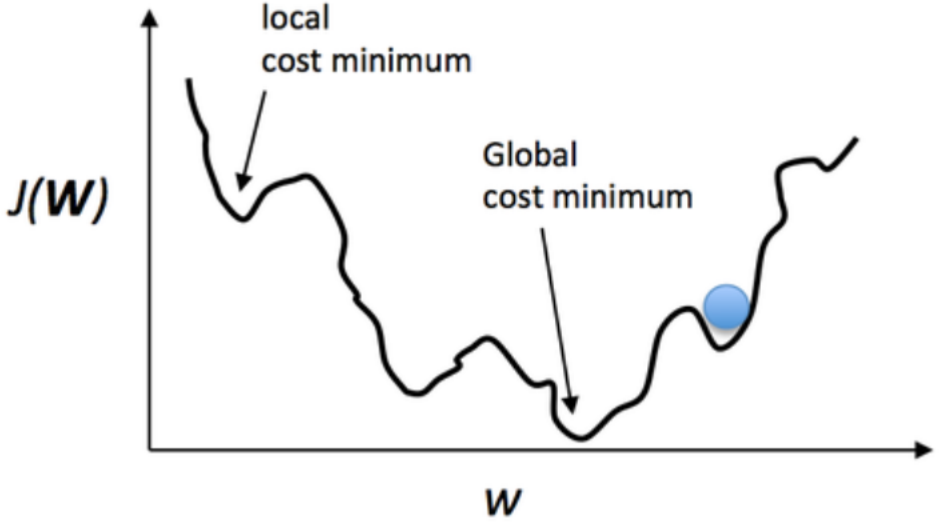
Logistic Regression Hypothesis



Logistic regression : Cost function

✓ **Cost Function**
(linear Regression)

$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$



–〉 Linear regression의 cost function을 이용하면, 볼록한 형태가 아니라 울퉁불퉁한 형태가 나오기 때문에 Gradient descent algorithm 을 통하여 최솟값을 구할 수 없게 된다.

Local cost minimum != Global cost minimum

Not Convex function!

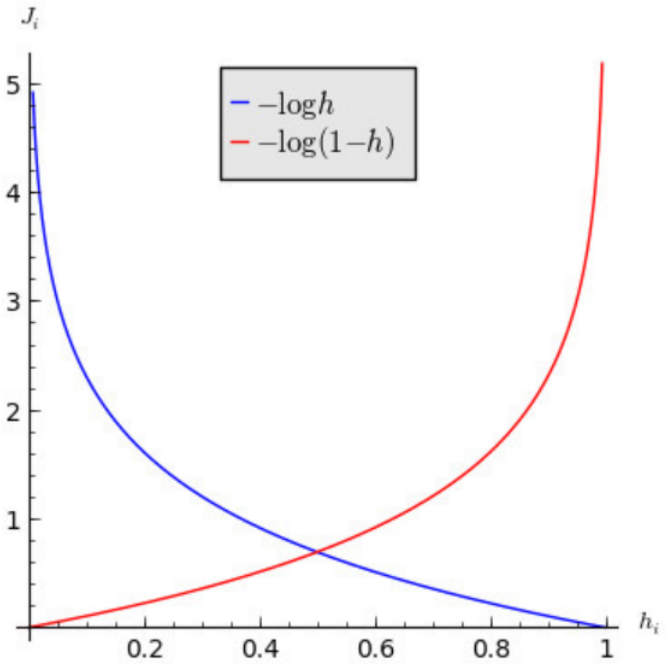
Logistic regression : Cost function

✓ **Cost Function**
(Logistic Regression)

$$cost(W) = \frac{1}{m}\sum \quad c(H(x), y)$$

$$c(H(x), y) = \begin{cases} -log(H(x)) & : y = 1 \\ -log(1 - H(x)) & : y = 0 \end{cases}$$

$$c(H(x), y) = ylog(H(x)) - (1 - y)log(1 - H(x))$$



-〉 log 값을 사용하게 되어 실제 데이터 결과 값인 y가 1일 때는

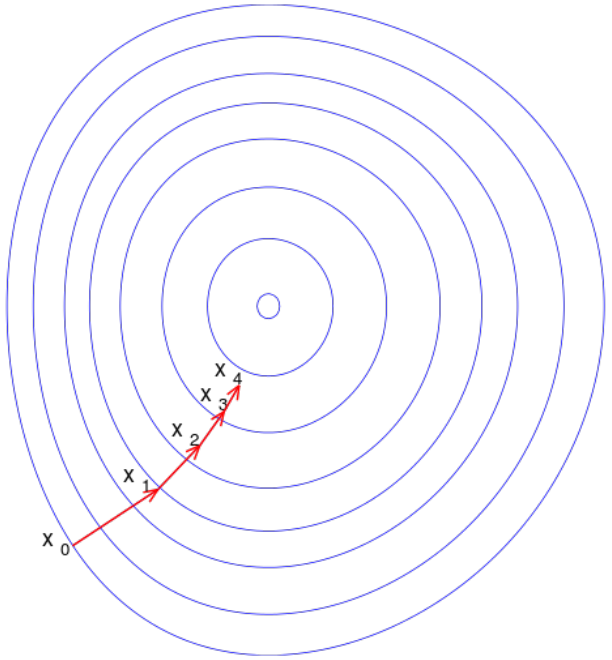
- 예측값(x)이 1에 가깝다면 (예측값 == 실제 값) cost값은 0으로 수렴(cost값이 최저),
- 예측값(x)이 0에 가깝다면 (예측값 != 실제 값) cost값은 cost값이 커짐.

Y가 0일때는 반대로 성립.

Logistic regression : Minimize cost - Gradient descent

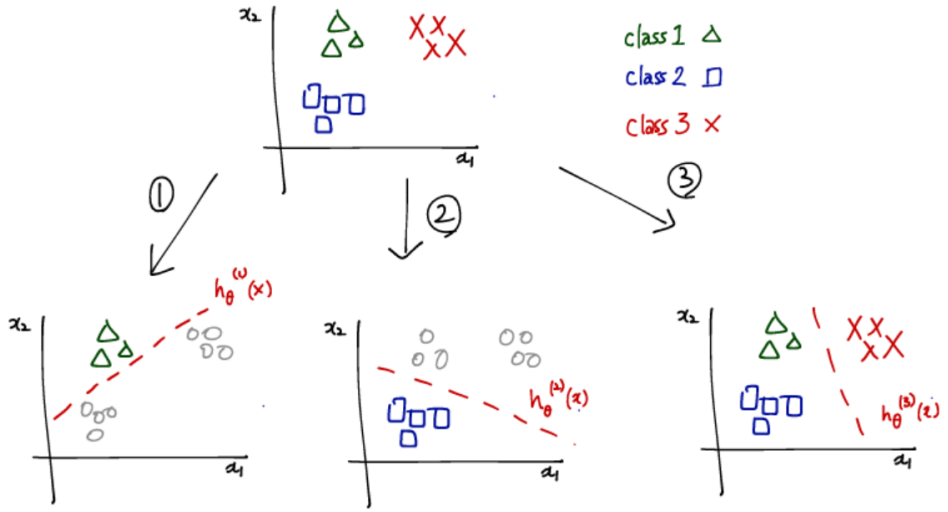
✓ **Gradient descent algorithm**
(= linear Regression)

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



Logistic regression : Multi-class classification

- ✓ Multi-class Classification
  - ✓ Example
    - Azure Cognitive service 를 이용한 multi-class classification
    - Cat(1) Dog(2) Lion(3) 3가지의 범주로 분류



$$h_\theta^{(i)}(x) = P(y = i|x;\theta), \quad i \in \{1,2,3\}$$

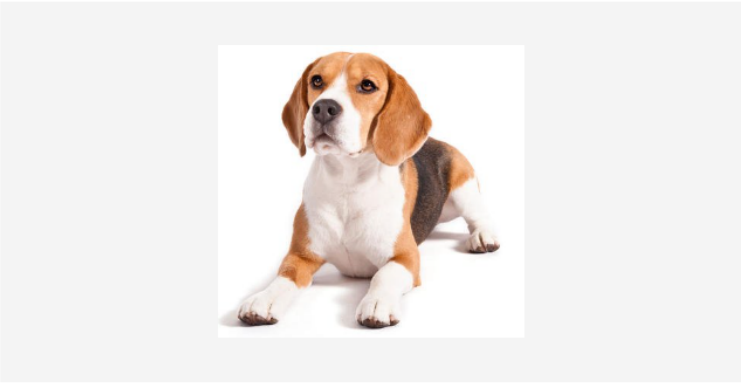
출처: https://wikidocs.net/4291#one-vs-all-one-vs-rest

Logistic regression : Multi-class classification

- ✓ Multi-class Classification
  - ✓ Example
    - Azure Cognitive service 를 이용한 multi-class classification
    - Cat(1) Dog(2) Lion(3) 3가지의 범주로 분류

〈Trainning set〉

〈Result〉



Results

Tag	Probability
dog	100%
cat	0%
lion	0%

Neural Network

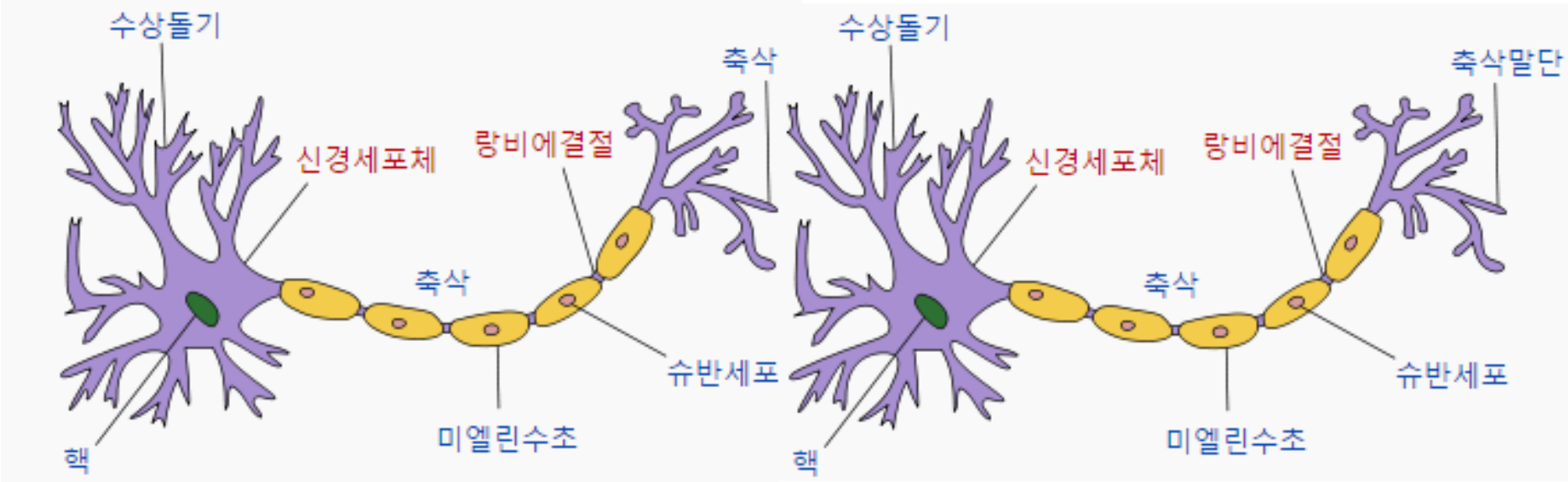
Neural Network : What is Neural Network?

✓Neural Network 란?

: 생물학의 신경망에서 영감을 얻은 알고리즘

-> 인공신경망

✓뉴런의 구조



Neural Network : What is Neural Network?

✓뉴런 -> 퍼셉트론

수학 모델

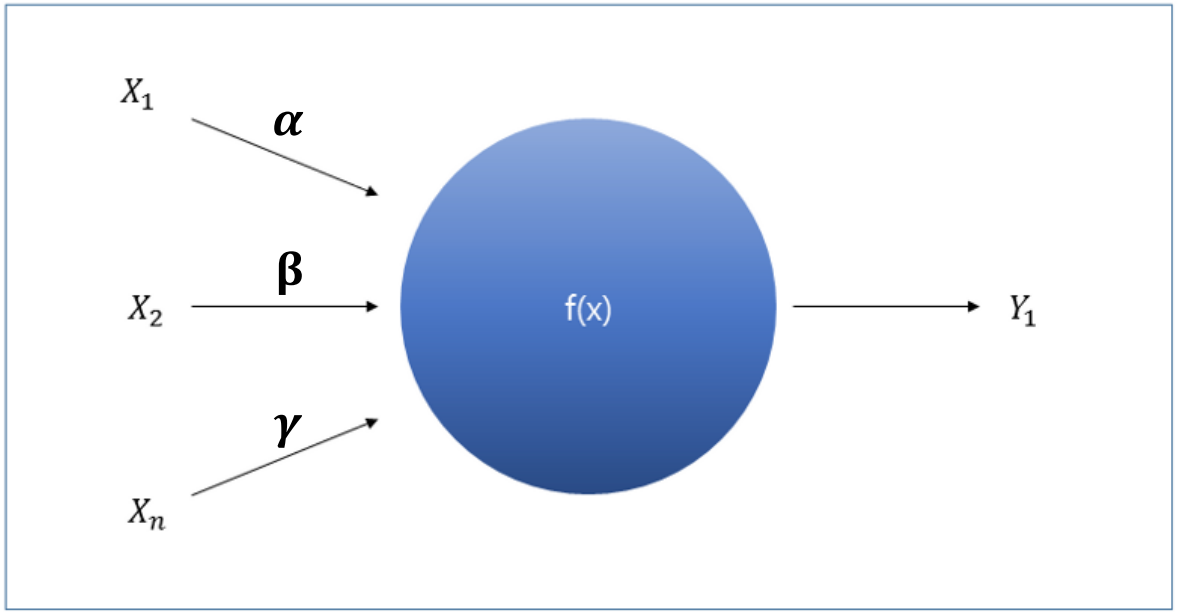
입력 → 함수 → 출력

✓퍼셉트론(perceptron)

: 인간 신경 세포 하나를 흉내 낸 것

- 하나의 퍼셉트론에 대해 학습 가능

도식



y = 목표치

$$x_1 \times \alpha + x_2 \times \beta + x_3 \times \gamma = y_0$$

y와 y_0 를 비교-> 가중치 α, β, γ 조정

뉴런

신호 → 수상돌기 → 축색돌기 → 신호

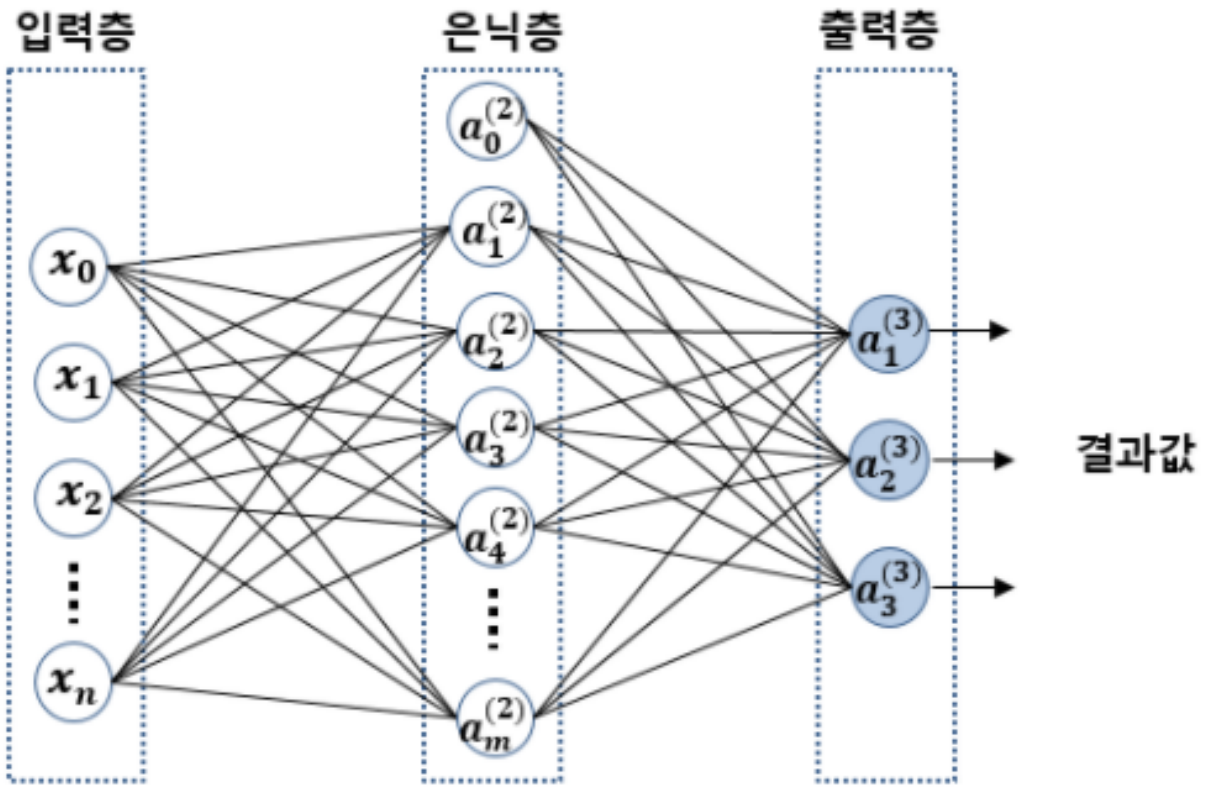
Neural Network : What is Neural Network?

✓Neural Network

: 하나의 퍼셉트론을 병렬로 배치하고

이 하나의 layer를 여러 개로 배치한

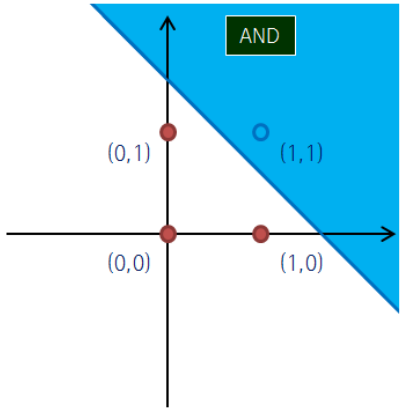
다층 퍼셉트론

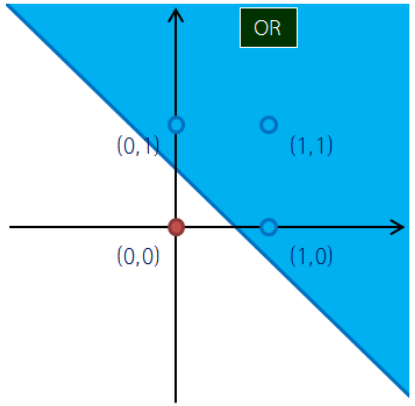


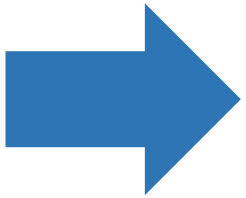
Neural Network : XOR Example

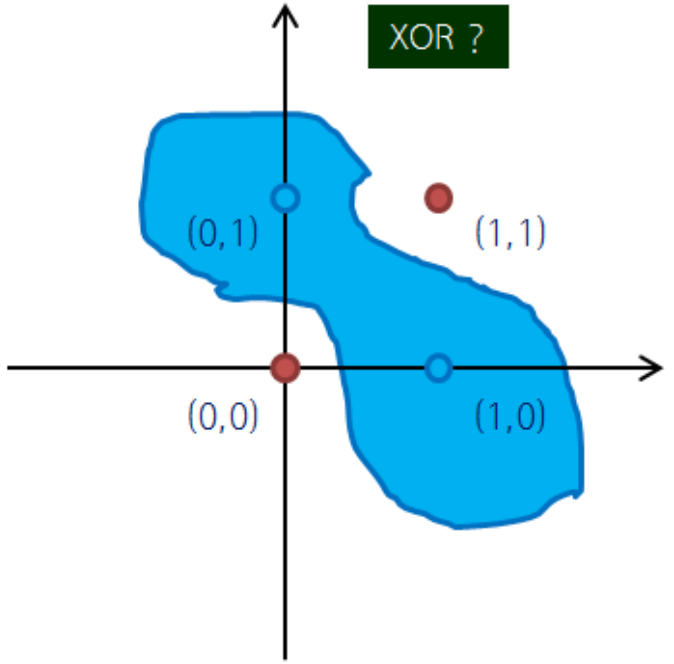
✓XOR example

A	B	A **XOR** B
0	0	0
0	1	1
1	0	1
1	1	0





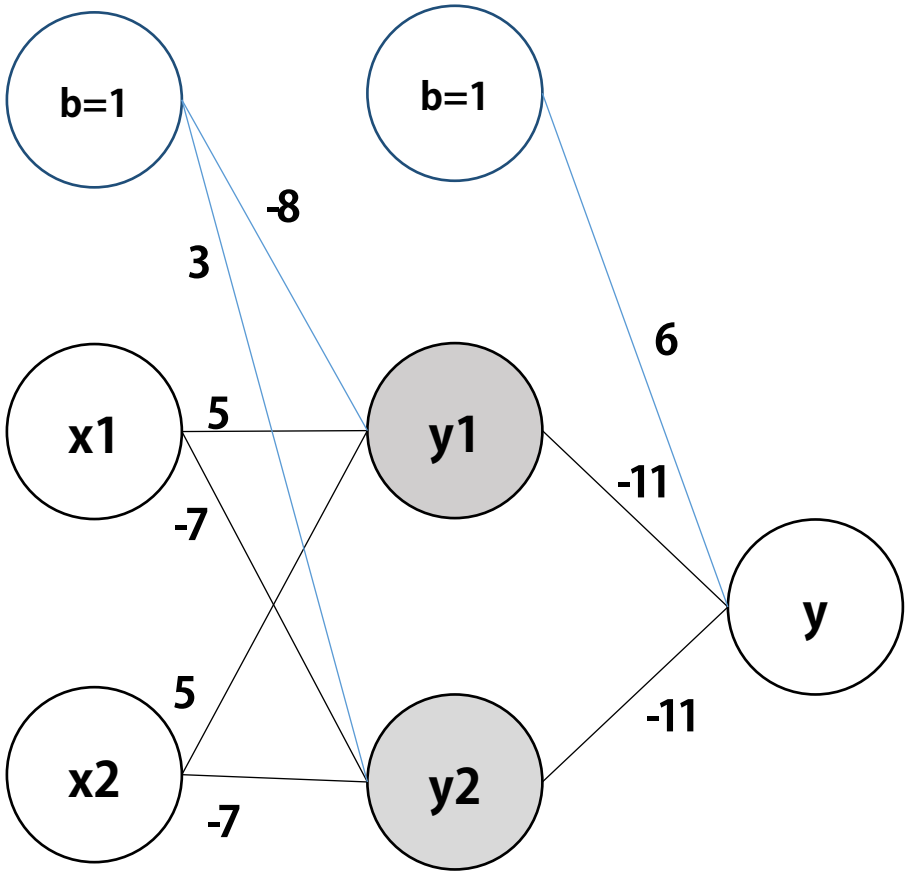




Neural Network : XOR Example

✓XOR example

x1	x2	y1	y2	y
0	0	0	1	0
0	1	0	0	1
1	0	0	0	1
1	1	1	0	0



y1 : 0*5+0*5-8=sigmoid(-8)=0
y2 : 0*-7+0*-7+3=sigmoid(3)=1
y=0*-11+ 1*-11+6=sigmoid(-5)=0

감사합니다