

투빅스 11기 정규과정

ToBig's 10기 박성진

# Logistic Regression & Classification

## Unit 05 | Tutorial & Assignment

### ◆ Tutorial

- Logistic Regression 맛보기 with Python(Logistic\_Regression\_01.ipynb)

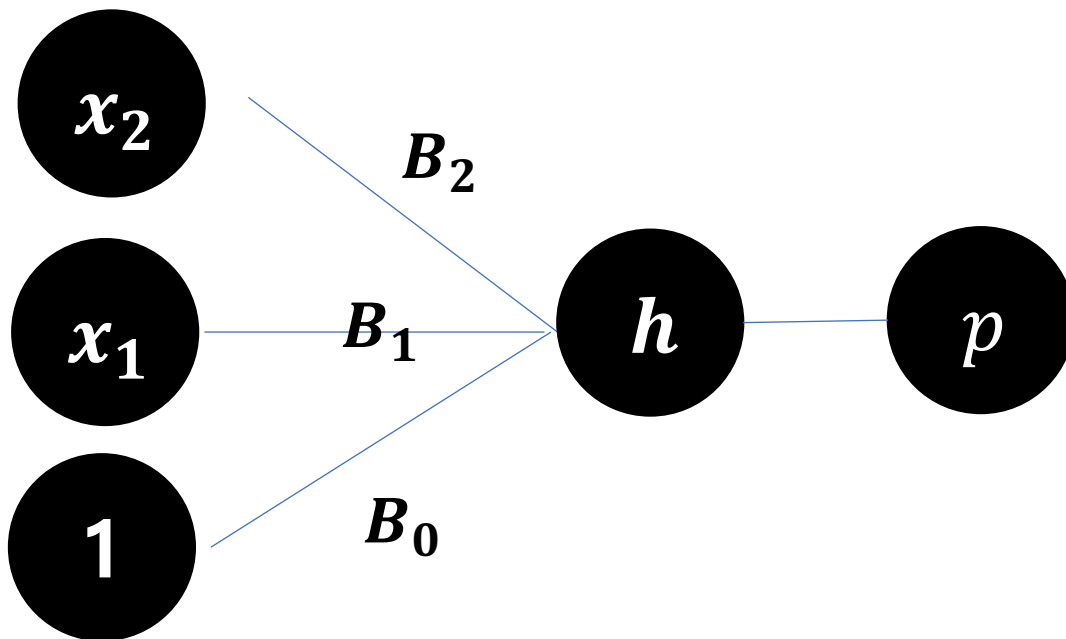
### ◆ Assignment 1

1. 튜토리얼 파일(Logistic\_Regression\_01.ipynb)을 수정하세요.
  - \* confusion matrix와 ROC curve를 활용하여 cut-off만 조정
  - \* 새로운 결과값 도출 및 의미해석

## Unit 05 | Tutorial & Assignment

### ◆ Assignment 2

- Gradient Descent Algorithm
  - ✓ Logistic Regression에서의 회귀 계수 추정
    - Gradient Descent Algorithm을 이용한 방법
    - 설명변수가 두개인 경우



$$h = \sum_{i=1}^2 B_i x_i + B_0$$

$$p = \frac{1}{1 + \exp(-h)}$$



## Unit 05 | Tutorial & Assignment

### ◆ Assignment 2

- Gradient Descent Algorithm

$$h = \sum_{i=1}^2 B_i x_i + B_0$$

$$p = \frac{1}{1 + \exp(-h)}$$

$$L^* = \log(L) = \sum_{i=1}^n \log(p^{x_i} (1-p)^{1-x_i}) = \sum_{i=1}^n \{x_i \log(p) + (1-x_i) \log(1-p)\}$$

$L^*$  는 우리가 Maximize 해야하는 likelihood

## Unit 05 | Tutorial &amp; Assignment

## ◆ Assignment 2

- Gradient Descent Algorithm

$$L^* = \log(L) = \sum_{i=1}^n \log(p^{x_i}(1-p)^{1-x_i}) = \sum_{i=1}^n \{x_i \log(p) + (1-x_i) \log(1-p)\}$$

이  $L^*$  에 마이너스를 붙이면 Convex 함수가 된다  
따라서 Likelihood를 Maximize하는 방법으로 푸는게 아니라

## Unit 05 | Tutorial & Assignment

### ◆ Assignment 2

- Gradient Descent Algorithm

$$L^* = \log(L) = \sum_{i=1}^n \log(p^{x_i}(1-p)^{1-x_i}) = \sum_{i=1}^n \{x_i \log(p) + (1-x_i) \log(1-p)\}$$

$J = -\log(L)$  ,  $\log(L)$ 은 앞서 말한 *Likelihood*

따라서 Convex Function 으로 바꿔준 Likelihood를 새로운 Object Function으로 보고, Gradient Descent 방법으로 최적화를 진행

## Unit 05 | Tutorial & Assignment

### ◆ Assignment 2

- Gradient Descent Algorithm

LABEL (남/여)	Bias	experience	salary
1	1	0.7	48000
0	1	1.9	48000
1	1	2.5	60000
...	...	...	...

- Object function을 최적화 한다는것은 결국 회귀계수를 추정한다는 의미
- 주어진 데이터를 본인이 구현한 최적화 알고리즘을 가지고 회귀계수를 추정하세요.
- Bias term의 사용은 선택사항