

로지스틱 회귀분석

ToBig's 11기 이영전

# Logistic Regression

Week2 과제

## 과제 01 | 로지스틱 개념 이해하기

# ★ Assignment2

- 실습코드에서 Q1. Q2.채우기

1번 문제는 로지스틱 개념이해를 목적으로 출제했습니다.

- 1) P값의 변화량에 따른 Odds의 변화량을 설명해주세요.
- 2) 입력 값이 주어졌을 때, 로지스틱 함수를 통해 Classification하는 방법에 대해 설명해주세요.(개념)

“Logit값을 통해 추정된 회귀식을 바탕으로 회귀식을 추정한 뒤 000~~”

## 과제 02 | 로지스틱 모델 구현하기

# ★ Assignment2

- 로지스틱 회귀 모델 구현하기(Feat. Money ball Data)
- 2번 과제의 목표는 로지스틱 회귀 모델 구현입니다.
- sklearn에서 제공하는 Library를 활용하면 비교적 간단하게 구현하실 수 있습니다.
- 전처리의 어려움을 도와드리고자 EDA와 전처리를 한 파일을 첨부해 드립니다.(이미 하신 분들은 자신의 방식으로 하셔도 무관합니다!)
- **시간적 여유가 있으시다면 Tutorial을 따라가면서 하나하나 구현해보세요.**

\* MoneyBall\_EDA\_Logistic\_Help.ipynb

\* Moneyball\_Explanation.ipynb (머니볼과 빅데이터) <- 영화와 이글을 읽고 영감을 받았습니다.

## 과제 02 | 로지스틱 모델 구현하기

# ★ Assignment2

- 첨부된 Ipython 파일에 주어진 답들에 대한
- 정답을 적어주세요.
- Logistic Model 구현과
- Data 교차검증을 하시면 됩니다.

Q1. train\_test\_split, data를 **trainset**과 **testset**로 나눠주세요.

train\_test\_split() 함수에 들어가는 각각의 인자 값은 무엇을 의미하는지  
설명해주세요

- arrays
- test\_size
- random\_state

Q2. **Scaling**

Scaling을 통해 우리가 하고자 하는 것은 무엇인가요?

Q3. **LogisticRegression()** 모델을 만들어주세요. 그리고 만든 모델 인자값에 들어  
가는 값들의 의미를 설명해주세요.

- e.g LogisticRegression(random\_state=0, solver="", multi\_class=")
- \*\*모델명은 model로 만듭니다\*\*
- random\_state
- solver
- multi\_class

Q4. data를 **교차검증** 해주세요.(10-fold cross\_validation)

- 10-fold cross\_validation을 위한 인자값을 입력해주세요.
- kfold = selection.KFold("교차검증을 위한 인자 만들기")
- 교차검증 결과를 출력하고 해석합니다.

## 과제 02 | 로지스틱 모델 구현하기

### ★ 이번 과제를 통해...

- Logistic Regression에 대한 개념적인 이해와
- 데이터 분석 전반적인 프로세스를 배우셨으면 합니다.
- Library를 활용하면 Model구현은 크게 어려운 일이 아닙니다.
- 그럼 화이팅...!!

Q & A

들어주셔서 감사합니다.