logistic regression

# 라이브러리 링크

https://github.com/scikit-learn/scikit-learn/blob/main/sklearn/linear\_model/\_base.py

# 기초 설명

로지스틱 회귀의 목적은 일반적인 회귀 분석의 목표와 동일하게 종속 변수와 독립 변수간의 관계를 구체적인 함수로 나타내어 향후 예측 모델에 사용하는 것이다. 이는 독립 변수의 선형 결합으로 종속 변수를 설명한다는 관점에서는 선형 회귀 분석과 유사하다. 하지만 로지스틱 회귀는 선형 회귀 분석과는 다르게 종속 변수가 범주형 데이터를 대상으로 하며 입력 데이터가 주어졌을 때 해당 데이터의 결과가 특정 분류로 나뉘기 때문에 일종의 분류 (classification) 기법으로도 볼 수 있다.

<https://ko.wikipedia.org/wiki/%EB%A1%9C%EC%A7%80%EC%8A%A4%ED%8B%B1_%ED%9A%8C%EA%B7%80>

# 버전 정보

* NumPy >= 1.14.6 (pip install numpy)
* pandas >= 1.2.4 (pip install pandas)
* matplotlib == 3.22 (pip install matplotlib)
* sklearn == 1.0.2 (pip install sklearn)
* seaborn == 10.11.2 (pip install seaborn)

# 데이터셋 설명 및 출처

* sklearn내장 데이터 사용(아래 홈페이지 참조)

<https://scikit-learn.org/stable/install.html>

# 코드 설명

* sklearn에 있는 위스콘신 유방암 환자 데이터를 가지고 로지스틱 회귀를 구현하는 코드이다.

# 검증 방법

* f1\_score을 통한 검증