**1-3 마켓과 머신러닝**

산점도(scatter plot)

-matplotlib

import matplotlib.pyplot as plt

plt.scatter(bream\_length,bream\_weight)

plt.xlabel('길이')

plt.ylabel('무게')

plt.show()

2가지 데이터 1그래프

plt.scatter(bream\_length,bream\_weight)

plt.scatter(smelt\_length,smelt\_weight)

plt.xlabel('길이')

plt.ylabel('무게')

plt.show()

k-최근접 이웃

2차원리스트

fish\_data = [[l,w] for l,w in zip(length,weight)]

모듈 불러오고 객체만들고, 훈련, 정확도평가, 예측

from sklearn.neighbors import KNeighborsClassifier

kn = KNeighborsClassifier()

kn.fit(fish\_data,fish\_target)

kn.score(fish\_data,fish\_target)

kn.predict([[33,176]])

최근접이웃 근접 확인 개수 조정(기본값=5)

KNeighborsClassifier(n\_neighbors=29)

+n\_jobs=-1 모든 CPU코어 사용

**2-1 훈련세트와 테스트세트**

지도학습: 정답이 주어짐

비지도학습: 정답이 없이 군집 분류

2차원배열로 만들기 위하여 numpy 이용

import numpy as np

input\_arr = np.array(fish\_data)

target\_arr = np.array(fish\_target)

인덱스를 사용하여 input과 target의 매칭이 깨지지 않도록 랜덤화

np.random.seed(42)

index = np.arange(49)

np.random.shuffle(index)

(seed=난수고정)

**2-2 데이터 전처리**

numpy column\_stack

테이블이(가) 표시된 사진

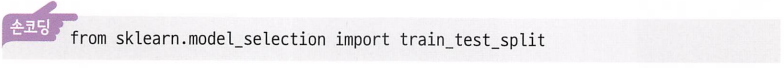
자동 생성된 설명

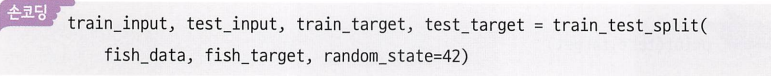
Concatenate

텍스트이(가) 표시된 사진

자동 생성된 설명

훈련세트와 테스트세트 나누기



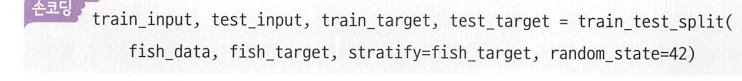


확인 numpy shape사용

테이블이(가) 표시된 사진

자동 생성된 설명

샘플링 편향 방지



k-최근접이웃 근처 점 확인

텍스트이(가) 표시된 사진

자동 생성된 설명

평균 표준편차 구하기

텍스트이(가) 표시된 사진

자동 생성된 설명

스케일을 맞추기



**03-1 k-최근접 이웃 회귀**