

분류, 군집 연습문제

문제 1

아래의 주소는 오렌지와 자몽의 특성을 조사한 데이터셋을 내려받을 수 있는 웹 페이지이다.

<https://www.kaggle.com/datasets/joshmcdams/oranges-vs-grapefruit>

변수	설명	구분
name	오렌지, 자몽 여부	종속
diameter	지름	독립
weight	무게	독립
red	붉은 색상 정도	독립
green	초록 색상 정도	독립
blue	파란 색상 정도	독립

이 데이터셋을 활용하여 오렌지와 자몽을 분류할 수 있는 모델을 구성하고 성능 측정 결과를 제시하시오.

대부분의 데이터는 허구로 작성된 잘 정제된 것이므로 높은 스코어가 나타날 것 입니다.

문제 2

위 데이터에 대한 군집 모델을 구현하시오.

그런 후 임의의 데이터에 대해 분류 결과와 군집 결과가 동일한지 확인하세요.