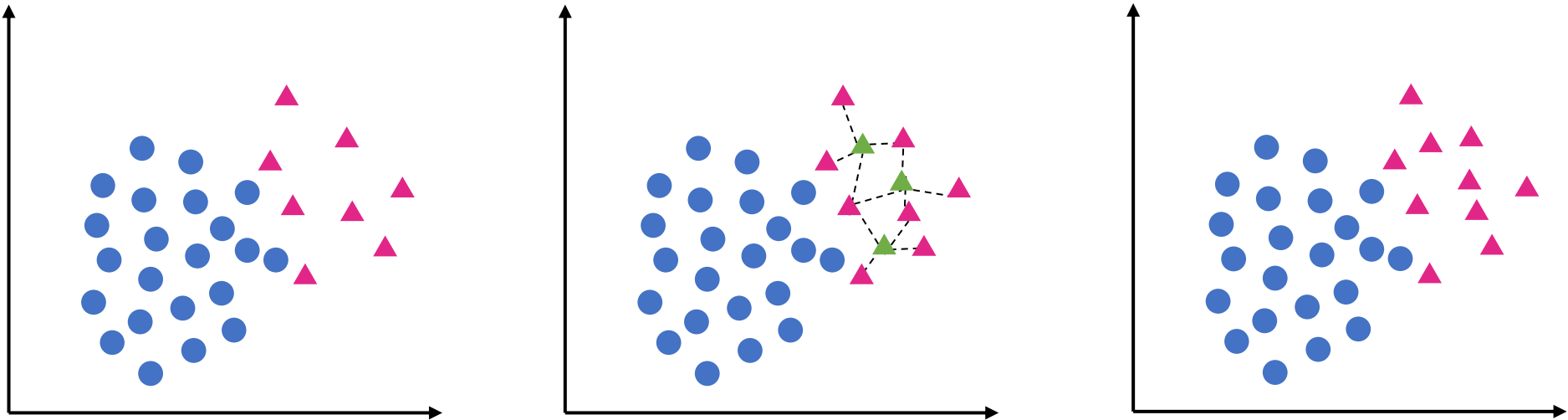


SMOTE (synthetic minority oversampling technique)

소수의 데이터를 인위적으로 오버샘플링하는 방법



공식문서 : https://imbalanced-learn.org/stable/references/generated/imblearn.over_sampling.SMOTE.html

GridSearchCV

각 하이퍼 파라미터에 입력한 값들의 경우의 수마다 교차검증을 통해 최적의 하이퍼 파라미터 값을 알려줌

주의사항 값을 많이 넣으면 그만큼 시간이 오래 걸림. 또한 지정 범위 밖의 최적의 하이퍼 파라미터 값을 놓칠 수 있음

param_grid : 딕셔너리 형태로 각 파라미터 명과 파라미터 값(리스트 형태로)을 둔다.

cv : Cross-Validation(교차 검증). 데이터를 몇 개의 부분집합으로 나눌 것인지 설정.
각 부분집합을 사용하여 모델을 여러 번 학습하고 평가하여 성능을 평균적으로 측정함.

n_jobs : 모델 학습 시에 병렬로 실행할 작업의 개수를 지정.
n_jobs=-1은 모든 CPU코어를 사용하여 병렬처리하여 학습시간 단축.

공식문서 : https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html