백화점 데이터 기반 성별 예측

## 팀원 : 김지훈, 박주희, 박민우

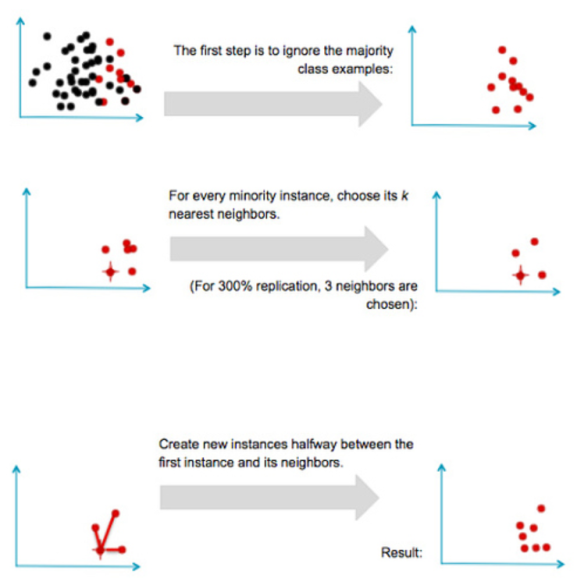
## 기간 : 2020.04.15~2020.04.23

1. **DATA 전처리**
   1. **Remove Missing(NULL) Values** & ONE-HOT Incoding

Null값 제거 및 (‘주구매코너’) 원핫인코딩

* 1. Target Label Check

성별자료 데이터 비대칭 확인



1. **Data Analysis**

2-1. 데이터 상관관계 확인 및 시각화

1. **Modeling**

3-1. SVC

3-2. MLP

3-3. Logistic Regression

3-4. KNeighborsClassifier

3-5. DecisionTreeClassifier

3-6. GradientBoostingClassifier

결과 confusion matrix로 나타내기

1. **예측률 높이기(추가실험)**

4-1. 다른 모델 사용해보기

(bagging,Ada Boost)

4-2.SMOTETomek 사용하기

4-3.Oversampling,Undersampling,

Nearsampling 해보기

4-4.pipline

4-5.GridSearchCV