**3. 평가 ( Evalutaion )**

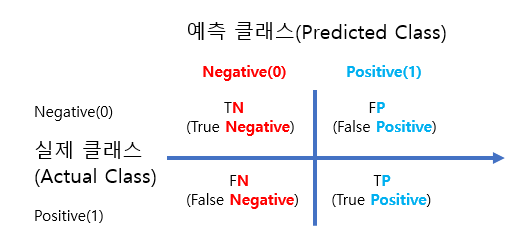
**- 분류(Classification) 성능 평가 지표**

**정확도(Accuracy) =**

* 이진 분류의 경우 불균형한 레이블 분포이면 ML 모델성능이 왜곡됨. **적합한 평가지표X**

(즉, 10000건 중 사기 당한 데이터 건수는 1건) (매우 불균형)

**오차행렬(Confusion Matrix)** **이진 분류**의 예측오류를 보여주는 지표



\*정확도 = 예측결과와 실제 값이 같은 건 수 / 전체데이터 수 =

**정밀도(Precision)** : 예측을 Posivite 한 대상 중 예측 == 실제인 비율 = precision\_score()

* 정밀도가 더 중요한 지표인 경우 ?

실제 Negative인 데이터 예측을 Positive로 잘못 판단 🡪 큰 영향이 발생 : ex) 스팸메일

**재현율(Recall) :** 실제 값을 Positive 한 대상 중 예측 == 실제인 비율 = recall\_score()

* 재현율이 더 중요한 지표인 경우 ?

실제 Positive인 데이터 예측을 Negative로 잘못 판단 🡪 큰 영향이 발생 : ex) 암 진단, 금융사기

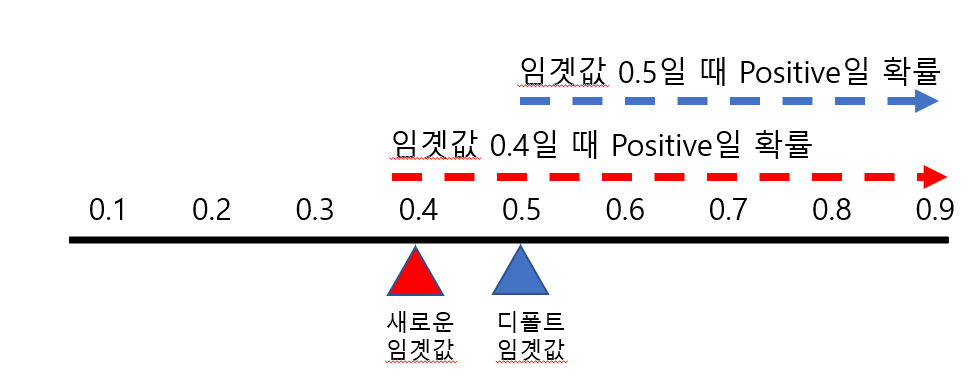
\* 불균형한 레이블 클래스의 이진 분류 모델은 매우 적은 수의 결과 값에 Positive(1)을, 그렇지 않은 경우 Negative (0)을 부여함

**정밀도(Precision)/ 재현율(Recall) Trade-Off :**

정밀도 또는 재현율이 특별히 강조돼야 할 경우 **분류의 결정 임계값(Threshold)을 조정해** 정밀도 또는 재현율의 수치를 높일 수 있음.

정밀도와 재현율은 상호 보완적인 평가지표 🡪 하나 높이면 하나는 떨어짐 : **Trade-Off**

**분류 결정 임계값(Threshold)에 따른 Positive 예측 확률 변화 : precision\_recall\_curve()**



**정밀도 =** **정확도 =**  ( *FP TP 는 증가 FN 이 감소 )*

**분류 결정 임계값이 낮아질수록 Positive로 예측할 확률이 높아짐. 재현율 증가 !**

**정밀도(Precision)/ 재현율(Recall) 의 맹점 :**

**F1 스코어 :** 정밀도와 재현율 어느 한쪽으로 치우치지 않아야 높은 값을 가짐. **f1\_score()**

**F1 = =**

**ROC곡선과 AUC :** 머신러닝의 이진 분류 모델 예측 성능을 판단하는 중요한 평가지표

**ROC곡선 :** FPR(False Positive Rate)가 변할 때 TPR(True Positive Rate)가 어떻게 변하는지

\* TPR : 재현율(민감도) =

\* FPR : 실제 Negative를 잘못 예측한 비율 =

**AUC :** ROC곡선 밑의 면적을 구한 것 **1 에 가까울수록 좋은 수치**

**roc\_curve(y\_true, y\_score**) : fpr, tpr ,threholds (array)로 반환

**roc\_auc\_score(y\_true, y\_score)** : Auc 스코어 값 반환

y\_true : 실제 클래스 값 array

y\_score : **predict\_prob() 의 반환 값** array에서 Positive 컬럼의 예측확률이 보통 사용됨.