# Confusion Matrix

Binary classification 문제에서 target(Y)의 값은 0 또는 1이다. 이 상황에서 모델의 성능을 평가하기 위한 지표로 정확도(accuracy)를 이용할 수 있다. 즉, 50개의 데이터 중 35개를 올바르게 분류하였다면 정확도(accuracy)는 70%이다. 정확도 개념은 target의 각 클래스의 비율이 동일한 경우 합리적일 수 있지만, 클래스 비율의 비대칭성이 심한경우 문제가 될 수 있다. 예를 들어, 지폐의 진품(0) 또는 위조(1) 여부를 분류한다고 가정해본다. 이때, 지폐의 99%가 진품이고 1%가 위조 지폐라면 모든 지폐를 진품으로 분류하는 naïve algorithm으로 모델이 학습될 수 있다. 결과적으로, 모든 지폐를 진품으로 분류하더라도 99%라는 높은 정확도를 달성하지만 유용한 정보는 아니다. 이처럼 데이터의 비율이 비대칭적인 경우 모델의 성능을 판단하기 위해 정확도 외에 다른 척도를 사용할 필요가 있다.

1) Confusion matrix

| Confusion matrix | | predicted | |
|---|---|---|---|
| | | Negative(0) | Positive(1) |
| True (observed) | Negative(0) | TN, $N_{00}$ | FP, $N_{01}$ |
| | Positive(1) | FN, $N_{10}$ | TP, $N_{11}$ |

NOTE ) FP = type 1 error, FN = type 2 error

Confusion matrix는 위와 같이 table의 형식으로 정리한다. Confusion matrix에서 각각의 요소가 의미하는 바는 다음과 같다.

- TP(정답) : 실제로 1인 class이며, 모델의 예측값이 class 1으로 올바르게 분류
- FP(오답) : 실제로 0인 class이며, 모델의 예측값이 class 1으로 잘못 분류
- TN(정답) : 실제로 0인 class이며, 모델의 예측값이 class 0으로 올바르게 분류
- FN(오답) : 실제로 1인 class이며, 모델의 예측값이 clss 0으로 잘못 분류

위의 개념을 이용하여 분류 모델을 평가하기 위한 지표를 정의한다.

2) Glossary

분류 모델의 성능을 측정하는 두 가지 일반적인 지표는 민감도(sensitivity, recall)와 정밀도(precision)이다. 위의 예시에서, 민감도가 높다는 것은 실제 위조 지폐를 많이 식별하고 있다는 것을 의미한다. 또한, 정밀도가 높다는 것은 위조지폐를 예측할 때의 예측이 정확할 가능성이 높다는 것을 의미한다.

- Recall(민감도) = $\frac{TP}{TP+FN} = \frac{N_{11}}{N_{10}+N_{11}}$ : $y = 1$ class가 속한 자료 중 정분류된 자료의 비율
- Specificity(특이도) = $\frac{TN}{TN+FP} = \frac{N_{00}}{N_{00}+N_{01}}$ : $y = 0$ class가 속한 자료 중 정분류된 자료의 비율
- Precision(정밀도) = $\frac{TP}{TP+FP} = \frac{N_{11}}{N_{01}+N_{11}}$ : $y = 1$이라고 예측된 자료 중 실제 True 값이 $y = 1$인 자료의 비율

지폐 분류 예시에서, 모델이 모든 데이터를 $y = 1$로 분류하는 경우 recall = 1, precision = 0이 된다. 마찬가지로, 모델이 모든 데이터를 $y = 0$으로 분류하는 경우 recall = 0, precision = 1로 분류한다.

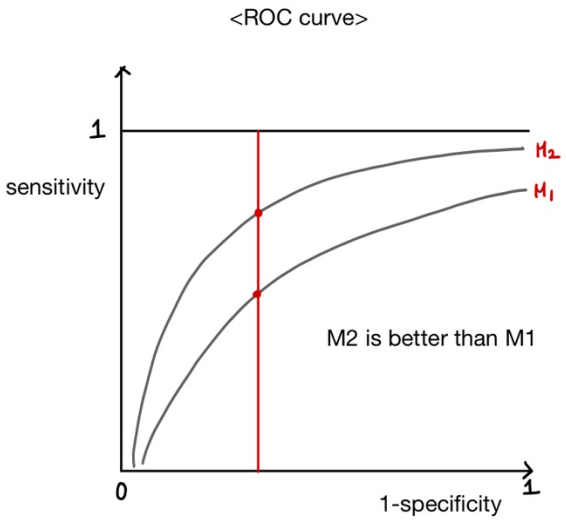
Recall과 precision이 모두 높은상황이 가장 이상적이지만, 두 지표를 모두 높이는 것은 불가능하다. Recall이 증가하면 precision이 감소하고, precision이 증가하면 recall이 감소한다. 따라서, 모델을 평가하는 경우 두 값의 조화평균인 F1-score를 이용한다. F1-score는 다음과 같이 정의된다.

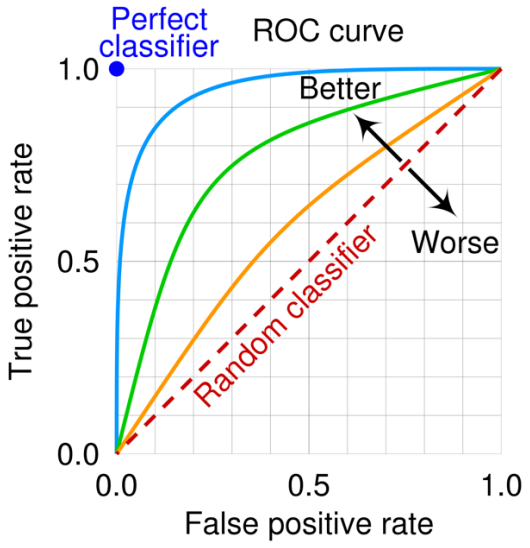
$$F_1\ score = \frac{2 \cdot recall \cdot precision}{recall + precision}$$

F1-score가 높을수록 좋은 모델이라고 판단한다.

3) ROC-curve & AUC

ROC-curve는 x축과 y축을 각각 sensitivity와 (1-specificity)로 정의한다. ROC-curve 또한 모델의 성능을 판단하는데 이용할 수 있다. ROC-curve에서 (0,1)에 가까울수록 좋은 모델이라고 판단한다.




NOTE ) sensitivity = True positive rate, (1-specificity) = False positive rate

만약 curve가 서로 교차하게 되면 어떤 모델의 성능이 더 우수한지 판단하기 어렵다. 따라서, curve 아래의 면적(Area Under the Curve, AUC)가 큰 모델의 성능이 더 우수하다고 할 수 있다.

<ROC curve>