

06

로지스틱 회귀분석 해석 방법

[로지스틱 회귀분석 해석]

00:08

로지스틱 회귀 방정식 중 회귀계수를 해석하는게 중요합니다!

로지스틱 회귀분석 해석방법

오즈(Odds)

$$odds = \frac{p}{1-p}$$

참일 확률 / 거짓일 확률
참일 확률이 거짓일 확률
보다 얼마나 큰가?

로짓(logit)변환

$$logit(p) = \ln\left(\frac{p}{1-p}\right)$$

오즈에
자연로그를 취한다.

로지스틱(logistic)

$$\ln\left(\frac{p}{1-p}\right) = a + bX$$

$$p = f(x) = \frac{e^{a+bX}}{1 + e^{a+bX}}$$

앞에서 구한 로짓 변환한 식을
직선의 방정식과 같다고 놓고
방정식을 재정리 하면 나오는 함수

[분류분석 평가지표]

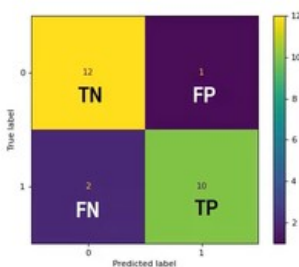
06:04

분류분석의 가장 기본은 혼동행렬을 이해하는 것입니다.

- 분류분석은 ADP 단골 문제입니다. confusion matrix를 이해하고 외워주세요!!

분류분석 평가지표

• confusion matrix (혼동행렬)



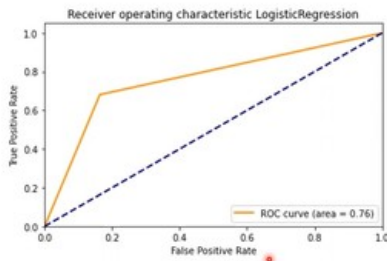
- TP(True Positive) : 긍정예측을 성공 즉, 환자라고 예측해서 실제 환자임을 맞춤
- TN(True Negative) : 부정예측을 성공 즉, 비환자라고 예측하여 실제 비환자임을 맞춤
- FP(False Positive) : 긍정예측을 실패 즉, 환자라고 예측했지만 비환자임
- FN(False Negative) : 부정예측을 실패 즉, 비환자라고 예측했지만 실제 환자임

[ROC 곡선과 ROC_AUC_SCORE]

10:25

ROC_AUC_SCORE는 분류분석의 대표적인 평가지표입니다.

분류분석 평가지표



ROC 곡선

- x축은 FPR(1-특이도), y축은 재현율을 가지는 그래프
- TPR(민감도)은 1 이고, FPR(1-특이도)는 0인 경우 가장 이상적
- 거짓 양성 비율(False Positive Rate;FPR) 의미

ROC_AUC

- ROC 곡선의 아래 면적을 계산한 것으로 값이 클수록 모델의 성능이 좋다고 판단 가능