

관련 용어

재현율 (Recall)

재현율은 실제로 맞는 정답 중에서 모델이 얼마나 많이 찾아냈는지를 나타내는 비율입니다. 즉, 놓치는 사례(False Negative)를 얼마나 줄였는지를 평가하는 지표입니다. 예를 들어 스팸메일 탐지에서 재현율이 높다는 것은, 대부분의 스팸을 성공적으로 잡아냈다는 의미입니다. 재현율은 모델이 가능한 한 많은 긍정 사례를 포착해야 하는 보안 탐지나 질병 진단과 같은 영역에서 특히 중요합니다. 다만 재현율이 높으면 불필요한 긍정 판단이 늘어날 수 있어, 정밀도와 균형 있게 해석해야 합니다. 두 지표는 보통 F1 점수(F1-score)로 함께 평가되어 모델의 종합적 성능을 판단하는 데 사용됩니다.

정확도 vs 정밀도 vs 재현율

이 세 지표는 모두 모델의 분류 성능을 측정하지만, 평가 관점이 서로 다릅니다. 정확도는 전체 예측 중 맞춘 비율을, 정밀도는 모델이 '긍정'이라고 판단한 결과 중 실제로 맞은 비율을, 재현율은 실제 긍정 사례 중 모델이 맞게 찾아낸 비율을 의미합니다. 예를 들어 스팸메일 탐지에서 정밀도가 높으면 정상 메일을 스팸으로 잘못 분류하지 않지만, 재현율이 높으면 대부분의 스팸을 놓치지 않고 잡아냅니다. 정밀도와 재현율은 상충관계에 있어, 둘 중 하나만 높이는 것은 어렵습니다. 따라서 모델의 목적에 따라 어떤 지표를 우선시할지가 달라집니다. 의료 진단처럼 놓치면 안 되는 경우에는 재현율을, 금융 사기 탐지처럼 오탐이 문제인 경우에는 정밀도를 중시합니다.

		실제 정답	
		True	False
분류 결과	True	옳은 긍정 예측(A) (True Positive)	잘못된 긍정 예측 (False Positive) 크면 재현율 ↓
	False	잘못된 부정 예측 (False Negative) 크면 정밀도 ↓	옳은 부정 예측(B) (True Negative)

A + B = 정확도

관련 용어

F1 점수 (F1-score)

F1 점수는 정밀도와 재현율의 조화를 수치로 표현한 성능 지표로, 두 지표의 조화평균을 사용해 한쪽으로 치우치지 않는 균형 잡힌 평가를 제공합니다. 예를 들어 정밀도는 높지만 재현율이 낮거나, 그 반대인 경우 모두 F1 점수가 낮게 계산되어 모델의 전반적 신뢰성을 판단할 수 있습니다. 이 지표는 특히 불균형 데이터셋에서 정확도만으로는 성능을 평가하기 어려울 때 유용합니다. F1 점수가 높을수록 모델이 긍정 사례를 정확하고 끝넓게 탐지한다는 의미로, 의료 진단, 스팸 탐지 등 오탐·누락이 중요한 분야에서 널리 사용됩니다.