Metrics for Model Evaluation

Metrics 또는 지표는 머신러닝 및 데이터 분석에서 모델의 성능을 평가하는 데 사용되는 기준을 의미합니다. 이들은 모델이 주어진 작업에서 얼마나 잘 수행되었는지를 정량적으로 평가하는 데 중요합니다.

# 1. 분류(Classification) 문제에서의 평가 지표

분류 문제에서 사용되는 주요 평가 지표는 다음과 같습니다:

정확도(Accuracy): 전체 예측 중 올바르게 예측된 비율입니다.

정밀도(Precision): 모델이 긍정 클래스로 예측한 것 중 실제로 긍정 클래스인 비율입니다.

재현율(Recall) 또는 민감도(Sensitivity): 실제 긍정 클래스 중 모델이 올바르게 예측한 비율입니다.

F1-스코어(F1-Score): 정밀도와 재현율의 조화평균입니다.

ROC-AUC: 분류 모델의 성능을 평가하기 위한 곡선 아래의 면적으로, 값이 1에 가까울수록 좋은 성능을 의미합니다.

# 2. 회귀(Regression) 문제에서의 평가 지표

회귀 문제에서 사용되는 주요 평가 지표는 다음과 같습니다:

평균 절대 오차(MAE, Mean Absolute Error): 예측값과 실제값 간의 절대적인 차이의 평균입니다.

평균 제곱 오차(MSE, Mean Squared Error): 예측값과 실제값 간의 차이의 제곱 평균입니다.

루트 평균 제곱 오차(RMSE, Root Mean Squared Error): MSE의 제곱근으로, 실제값과 예측값 간의 차이의 평균적인 크기를 나타냅니다.

결정 계수(R^2 Score): 예측값이 실제값과 얼마나 일치하는지를 나타내는 지표로, 값이 1에 가까울수록 예측이 잘 되었음을 의미합니다.

# 3. 클러스터링(Clustering) 문제에서의 평가 지표

클러스터링 문제에서 사용되는 주요 평가 지표는 다음과 같습니다:

실루엣 계수(Silhouette Score): 클러스터링의 품질을 측정하며, 값이 1에 가까울수록 잘 분류된 클러스터링을 의미합니다.

ARI (Adjusted Rand Index): 클러스터링이 실제 그룹과 얼마나 잘 일치하는지를 나타냅니다.

다빈도 균등점수(Purity Score): 각 클러스터에서 가장 큰 클래스를 찾고 그 클래스의 데이터 비율을 측정하는 지표입니다.

# 4. 사용 시 고려사항

문제의 특성, 데이터의 불균형, 모델의 목적에 따라 적합한 평가 지표를 선택해야 합니다. 하나의 지표만을 사용하는 것보다 여러 지표를 함께 사용하는 것이 모델의 성능을 더 잘 평가할 수 있습니다.