

정확도측정(딥러닝)

👤 생성자	👤 재환 김
🏷️ 태그	엔지니어링

1. 딥러닝에서의 정확도 측정 방법

(1) 분류 문제에서의 딥러닝 정확도

딥러닝 모델이 분류 문제를 다룰 때, 정확도는 모델의 예측값과 실제 라벨을 비교하는 방식으로 계산됩니다.

① 기본 정확도 (Accuracy)

- 분류 문제에서 딥러닝 모델의 정확도(Accuracy)는 다음과 같이 계산됩니다.

$$\text{Accuracy} = \frac{\text{정확히 분류된 샘플 수}}{\text{전체 샘플 수}}$$

예를 들어, 이미지 분류 문제에서 총 1000개의 이미지 중에서 950개가 정확히 분류되었다면, 정확도는 95%가 됩니다.

② 혼동 행렬(Confusion Matrix)

분류 문제에서 모델의 성능을 더 세부적으로 분석하려면 혼동 행렬(Confusion Matrix)을 사용합니다. 이 행렬을 사용하여 True Positive, False Positive, False Negative, True Negative 등을 계산할 수 있으며, 다음과 같은 추가 성능 지표를 도출할 수 있습니다.

- 정밀도(Precision):** 예측이 True로 된 것 중 실제로 True인 비율 $\text{Precision} = \frac{TP}{TP+FP}$

- 재현율(Recall):** 실제 True인 것 중에서 예측이 True로 된 비율 $\text{Recall} = \frac{TP}{TP+FN}$

- F1-score:** Precision과 Recall의 조화 평균

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

이러한 지표들은 불균형한 데이터셋을 다룰 때 유용합니다. 예를 들어, 특정 클래스가 데이터셋에서 매우 적은 비율을 차지할 경우, 단순한 정확도만으로는 모델 성능을 제대로 평가하기 어렵습니다.

(2) 회귀 문제에서의 딥러닝 성능 평가

회귀 문제는 분류 문제와 달리 연속된 값을 예측하므로, 정확도 대신 예측값과 실제 값의 차이를 측정하는 지표를 사용합니다. 대표적인 회귀 평가 지표는 다음과 같습니다.

① 평균제곱오차(Mean Squared Error, MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (y^{(i)} - \hat{y}^{(i)})^2$$

- 여기서 $y^{(i)}$ 는 실제 값, $\hat{y}^{(i)}$ 는 예측 값입니다.
- MSE는 예측 값과 실제 값의 차이의 제곱을 평균하여 구합니다. 값이 작을수록 모델의 예측이 정확하다는 것을 의미합니다.

② 평균 절대 오차(Mean Absolute Error, MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y^{(i)} - \hat{y}^{(i)}|$$

- MAE는 예측 값과 실제 값의 절대 차이의 평균을 구하는 방법으로, MSE와 달리 큰 오차가 발생해도 제곱하지 않기 때문에 안정적인 지표입니다.

③ 결정계수(R-squared, R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y^{(i)} - \hat{y}^{(i)})^2}{\sum_{i=1}^n (y^{(i)} - \bar{y})^2}$$

- R^2 값은 모델이 실제 데이터를 얼마나 잘 설명하는지를 나타내는 지표로, 1에 가까울수록 모델이 데이터를 잘 설명하는 것입니다.

2. 비지도 학습에서의 정확도 측정 방법

비지도 학습(Unsupervised Learning)은 정답 라벨이 없는 데이터를 사용하여 모델을 학습시킵니다. 그렇기 때문에 전통적인 의미의 정확도를 사용할 수 없으며, 모델이 학습한 결과를 평가하기 위한 다른 방법들이 사용됩니다. 대표적인 비지도 학습 방법인 군집화(Clustering)와 차원 축소(Dimensionality Reduction)에서의 성능 평가 방법을 살펴보겠습니다.

(1) 군집화(Clustering)에서의 성능 평가

① 실루엣 점수(Silhouette Score)

- 실루엣 점수는 데이터 포인트가 자신이 속한 군집과 얼마나 가까운지, 그리고 다른 군집과 얼마나 먼지를 측정하는 지표입니다. 실루엣 점수는 다음과 같이 계산됩니다.

$$S = \frac{b-a}{\max(a,b)}$$

- 여기서 a 는 한 데이터 포인트가 속한 군집 내에서의 평균 거리, b 는 그 포인트와 가장 가까운 다른 군집까지의 평균 거리입니다.
- 실루엣 점수는 -1에서 1 사이의 값을 가지며, 1에 가까울수록 해당 군집화가 잘 이루어진 것으로 봅니다.

② 엘보우 방법(Elbow Method)

- 군집화의 성능을 평가할 때, **엘보우 방법**은 최적의 군집 수를 찾기 위해 사용됩니다. 이 방법은 군집 수를 변경하면서 각 군집 내 데이터 간의 거리 제곱합(SSE: Sum of Squared Errors)을 계산합니다. SSE가 급격히 감소하는 지점을 찾아, 그 지점에서 최적의 군집 수를 선택합니다.

③ 군집의 일관성 평가(Davies-Bouldin Index, DBI)

- **Davies-Bouldin Index**는 각 군집 내의 데이터 응집도와 군집 간의 거리 비율을 계산하는 방법으로, 값이 작을수록 군집이 잘 분리된 것으로 간주됩니다.

(2) 차원 축소(Dimensionality Reduction)에서의 성능 평가

차원 축소는 데이터를 저차원 공간으로 투영하는 과정으로, 원본 데이터의 구조를 얼마나 잘 보존했는지를 평가해야 합니다.

① 재구성 오류(Reconstruction Error)

- 차원 축소 후 원본 데이터와 저차원 데이터로 재구성한 데이터 간의 차이를 측정하여, 얼마나 원본을 잘 보존했는지 평가할 수 있습니다. 예를 들어, 주성분 분석(PCA)에서는 원본 데이터를 저차원 주성분으로 변환한 뒤, 이를 다시 원래 차원으로 변환하여 원본 데이터와의 오차를 계산합니다.

② 설명 분산 비율(Explained Variance Ratio)

- 차원 축소 기법이 원본 데이터의 분산을 얼마나 잘 설명하는지를 나타내는 지표입니다. 예를 들어, PCA에서 주성분이 원본 데이터의 분산을 얼마나 설명하는지 확인하는데 사용되며, 높은 설명 분산 비율은 차원 축소가 성공적임을 나타냅니다.

요약:

- **딥러닝 분류**에서는 정확도, 정밀도, 재현율, F1-score 등과 같은 지표를 사용하며, **회귀**에서는 MSE, MAE, R2 같은 지표를 사용합니다.

R^2

- **비지도 학습**에서는 군집화 평가를 위해 실루엣 점수, 엘보우 방법, DBI 등을 사용하며, 차원 축소 평가를 위해 재구성 오류와 설명 분산 비율을 사용합니다.