2022-02-24, Boostcamp RECsys team 12

박정호

Boostcamp

Abstract

Accuracy, precision, recall에 binary 또는 multiclass에 대한 정의를 알아보고 비교한다.

1. 평가지표(score functions)

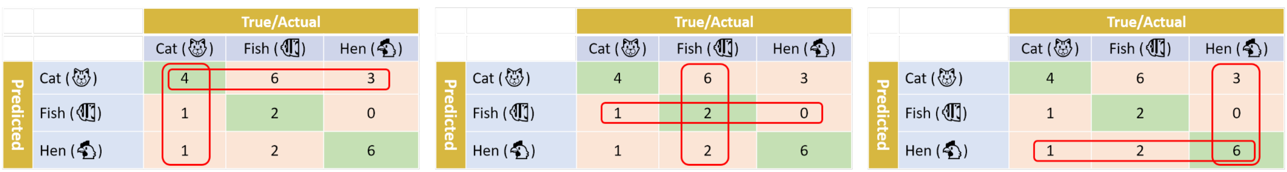

Figure 1: Example to comprehend precision(column) and recall(row).

Definition 1.1 (Accuracy)**.** 테스트 셋 $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$이 주어져있고 모델 $f : x_i \mapsto \hat{y}_i$이 주어져 있다. 이때, 모델의 **정확도**는 다음과 같다.

$$\text{Accuracy}(f, \mathcal{D}) := \frac{1}{N}|\{x_i \in \mathcal{D} : f(x_i) = y_i\}|. \tag{1}$$

이해하기 쉽게 기억하는 방법은 위의 그림의 *전체 경우에서 대각성분을 더한 값의 비율*이다. 즉, binary classification의 경우: P/N는 실제 데이터가 target/ non-target인 경우, T/F는 모델의 예측이 옳았다/ 틀렸다를 의미할 때,

$$\text{Accuracy} = \frac{TP + FN}{TP + TN + FP + FN}. \tag{2}$$

Definition 1.2 (Precision)**.** 테스트 셋 $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$이 주어져있고 모델 $f : x_i \mapsto \hat{y}_i$이 주어져 있다. 이때, 테스트 데이터 셋의 레이블이 $y_i \in \{1, 2, \cdots, K\}$ 라고 가정하자. 이때 모델의 **정밀도**는 다음과 같다.

$$\text{Precision}(f, \mathcal{D}) := \frac{1}{K}\sum_{j=1}^{K}\frac{|\{(x, k) \in \mathcal{D} \;:\; f(x) = k\}|}{|\{(x, y) \in \mathcal{D} \;:\; y = k\}|}. \tag{3}$$

Binary classification의 경우: P/N는 실제 데이터가 target/ non-target인 경우, T/F는 모델의 예측이 옳았다/ 틀렸다를 의미할 때

$$\text{Precision} = \frac{TP}{TP + FP}. \tag{4}$$

이해하기 쉽게 기억하려면, 위의 figure 1처럼 하나의 클래스를 기준으로 나머지를 모두 negative로 생각하고 precision을 계산한 뒤, 평균을 내는 것이다.

Definition 1.3 (Recall)**.** 테스트 셋 $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$이 주어져있고 모델 $f : x_i \mapsto \hat{y}_i$이 주어져 있다. 이때, 테스트 데이터 셋의 레이블이 $y_i \in \{1, 2, \cdots, K\}$ 라고 가정하자. 이때 모델의 **재현율**은 다음과 같다.

$$\text{Recall}(f, \mathcal{D}) := \frac{1}{K}\sum_{j=1}^{K}\frac{|\{(x, k) \in \mathcal{D} \;:\; f(x) = k\}|}{|\{(x, y) \in \mathcal{D} \;:\; f(x) = k\}|} \tag{5}$$

Binary classification의 경우: P/N는 실제 데이터가 target/ non-target인 경우, T/F는 모델의 예측이 옳았다/ 틀렸다를 의미할 때

$$\text{Recall} = \frac{TP}{TP + FN}. \tag{6}$$

이해하기 쉽게 기억하려면, 위의 figure 1처럼 하나의 클래스를 기준으로 나머지를 모두 negative로 생각하고 recall을 계산한 뒤, 평균을 내는 것이다.

2. Remind

- 벡터공간은 벡터공간 위에서 정의된 하나의 연산과 스칼라 곱 연산이 잘 호환되는 대수적 공간을 의미한다.
- n 차원 벡터 공간은 최대 n 개의 선형독립 집합을 가진다.
- n 개의 선형독립 집합을 **기저**라고 부른다.
$$\mathcal{B} = \{v_1, v_2, \cdots, v_n\}.$$
- 벡터공간의 모든 원소를 기저의 선형결합으로 **표현** 할 수 있다.
$$v = \sum_{i=1}^{n} c_i v_i.$$
- 기저는 일종의 좌표계로 이해하면 좋다.
$$v = \sum_{i=1}^{n} c_i v_i \quad \Longleftrightarrow \quad v = (c_1, c_2, \cdots, c_n)_{\mathcal{B}}.$$
- 선형변환은 행렬과 같다.
- 따라서 선형변환의 역변환은 역행렬을 왼쪽에 곱하는 것과 같다.
- 행렬을 분해하면 계산을 줄이고 해석에 편의성을 준다.

3. Matrix decomposition

3.1. Matrix Factorization

중학교 1학년 1학기로 기억을 더듬어 돌아가보자. 우리는 소수를 배우고 자연수의 소인수 분해를 배웠다. 그리고 시간이 좀 더 흘러서 다항식의 인수분해에 대해서 배운 기억이 있을 것이다. 이는 수학적인 물건을 더 알기 쉬운 물건으로 분해하는 것을 의미한다. 인수를 영어로 factor라고 부르는데, 이를 생각하면 Matrix Factorization은 행렬을 두개의 행렬의 곱으로 표현한 것이라고 할 수 있다. 행렬의 곱은 모양이 중요한데 여기에 재미난 특성이 있다.

$$(m \times n) = (m \times l) \times (l \times n). \tag{7}$$

이 때, l의 선택은 자유롭기 때문에 사이즈를 충분히 크게 만들어도 좋다. 추천시스템에서 사용되는 MF계열의 모델들은 이 l을 latent dimension이라고 부른다. 어떤 rating matrix $R \in \mathbb{R}^{m \times n}$에 대해서,

$$R = UI^t \tag{8}$$

를 만족하는 $U \in \mathbb{R}^{m \times l}$ 와 $I \in \mathbb{R}^{n \times l}$를 찾는 것이다. (9)번 식을 조금 뜯어보면 U의 행벡터를 유저의 latent vector로 생각할 수 있고, I의 행백터를 item의 latent vector라고 생각할 수 있다. 그리고 두 벡터의 내적($\hat{r} = u \cdot i = ui^t$)으로 rating의 추정값을 얻을 수 있다.

4. 미적분 들어가기

선형변환에 대한 역변환(역함수)는 선형변환이 되는 것을 배웠고 역변환이 존재하지 않을 때 가장 유사한 역변환이 무엇인지 공부했다. 일반적인 함수에 역함수가 존재하기 위해서 어떤 조건이 필요한지 알아보자.

Definition 4.1. 함수 $f : X \to Y$ 의 역함수 $f^{-1} : Y \to X$는 다음의 성질을 만족하는 함수이다.

$$f(f^{-1}(y)) = y, \quad f^{-1}(f(x)) = x \qquad \text{for} \quad x \in X, y \in Y. \tag{9}$$

Theorem 4.1. 일대일 대응 함수는 역함수를 가진다.

Example 4.1 (일대일 대응함수의 예)**.**
$$x, x^3, x^5, \cdots, x^{2n-1}$$

Example 4.2 (일대일 대응함수가 아닌 예)**.**
$$c, x^2, x^4, \cdots, x^{2n}$$

우리가 딥러닝 모델의 업데이트를 위해서 필요한 조건은 **미분가능성**이다. 정의역 전체에서 역함수를 가지지 않더라도 아주 작은 근방에선 역함수를 가진다. (접선을 생각하면된다.) 이 역함수의 미분을 간단하게 계산할 수 있는 공식이 있는데,

Theorem 4.2 (역함수 정리)**.** 일급함수 $f : \mathbb{R} \to \mathbb{R}$ 과 한 점 $x \in \mathbb{R} = \mathrm{dom}(f)$ 근방에 대해서 이에 대응하는 역함수 $f^{-1} : \mathbb{R} \to \mathbb{R}$ 가 존재하고,

$$(f^{-1})'(y) = \frac{1}{f'(x)} \qquad \text{for} \quad f(x) = y, f'(x) \neq 0 \tag{10}$$

이다.

Proof. $f^{-1}(y) = f^{-1}(f(x)) = x$ 이므로 양변을 x로 미분하면

$$\frac{d}{dx} f^{-1}(f(x)) = (f^{-1})'(f(x)) f'(x) = 1.$$

$f(x) = y$ 이고 $f'(x) \neq 0$이므로

$$(f^{-1})'(y) = \frac{1}{f'(x)}$$

임을 알 수 있다. □