

METRICS

1. Output:

Đối với mỗi sample, output có dạng là vector 6 chiều ứng với 6 khía cạnh, mỗi chiều nhận giá trị nguyên thuộc đoạn $[0, 5]$ tương ứng với 0: không có khía cạnh và 1-5 ứng với số sao đánh giá nếu khía cạnh đó xuất hiện.

Như vậy, output bài toán ứng với n samples:

$$y_i = [a_0, \dots, a_5], a_j \in [0, 5], i = 1 \dots n, j = 0 \dots 5.$$

Trong đó:

a_j : giá trị review của khía cạnh j .

(0: khía cạnh j không xuất hiện, 1-5: mức độ biểu cảm).

n : số samples.

y_i : output của sample thứ i .

Quy ước: khía cạnh 0-5 lần lượt là:

a_0 : Dịch vụ vui chơi giải trí.

a_1 : Dịch vụ lưu trú.

a_2 : Hệ thống nhà hàng.

a_3 : Dịch vụ ăn uống.

a_4 : Dịch vụ di chuyển.

a_5 : Dịch vụ mua sắm.

2. Metrics for Aspect Classification:

- Tính F1-score cho từng khía cạnh j :

$$F1_j = 2 \times \frac{Precision_j \times Recall_j}{Precision_j + Recall_j}$$

$$Precision_j = \frac{TP_j}{TP_j + FP_j}$$

$$Recall_j = \frac{TP_j}{TP_j + FN_j}$$

Lưu ý: Khi tính F1-score thì các giá trị 1-5 đều xem là 1, giá trị 0 vẫn giữ nguyên như output của bài toán Classification thông thường nhận giá trị đầu ra $\{0, 1\}$.

3. Metrics for 5-Stars Scale Rating:

- Xem như bài toán regression.
- Tính R2 score cho từng khía cạnh j:

$$R2_j = 1 - \frac{RSS_j}{K_j}$$

$$RSS_j = \text{Residual Sum of Squares} = \sum_{i=1}^{n_j} \left(\hat{y}_{ij} - y_{ij} \right)^2$$

$$K_j = \text{Total Sum of Squares of Max Distance} = \sum_{i=1}^{n_j} \left(\max_{sentiment} - \min_{sentiment} \right)^2$$

Trong đó:

- o \hat{y}_{ij} : giá trị dự đoán của khía cạnh j thuộc samples i.
- o \bar{y}_j : giá trị trung bình khía cạnh j của toàn bộ dataset.
- o y_{ij} : giá trị đúng của khía cạnh j thuộc sample i.
- o n_j : số samples dự đoán đúng có khía cạnh j.

Lưu ý: Chỉ tính R2-score đối với khía cạnh của samples mà ở đó giá trị dự đoán và giá trị thực tế đều khác 0. (Vì trường hợp nhận giá trị 0 đã xét ở bước tính F1-score cho từng khía cạnh).

4. Final metrics:

$$Score = \frac{1}{6} \sum_{j=0}^5 F1_j \times R2_j$$