

Xử lý ngôn ngữ tự nhiên cho Khoa học dữ liệu

Bài tập thực hành 3 - DS310.M11.1

Nguyễn Lưu Thùy Ngân, Nguyễn Văn Kiệt, Huỳnh Văn Tín,
Lưu Thanh Sơn, Nguyễn Đức Vũ, Nguyễn Thành Luân
Sinh viên: Phạm Đức Thế - 19522253
Ngôn ngữ lập trình: Python

Thứ 5, ngày 28 tháng 10 năm 2021

Bài tập: Named Entity Recognition (NER)

Bài 1: Sử dụng mô hình LSTM để giải quyết bài toán NER cho tiếng Anh.

1. Đánh giá mô hình theo độ đo Accuracy và F1-score (micro và macro) .

- Kết quả mô hình theo độ đo Accuracy là: 98.54%
- Kết quả mô hình theo độ đo F1-score micro là: 80.3%
- Kết quả mô hình theo độ đo F1-score macro là: 53.76%

2. Kết quả của F1-score và Accuracy cho từng nhãn thực thể.

- Kết quả của F1-score cho từng nhãn thực thể là:

	precision	recall	f1-score	support
B-art	0.00	0.00	0.00	96
B-eve	0.00	0.00	0.00	79
B-geo	0.78	0.89	0.83	7380
B-gpe	0.94	0.94	0.94	3173
B-nat	0.00	0.00	0.00	38
B-org	0.66	0.57	0.61	3908
B-per	0.86	0.76	0.81	3359
B-tim	0.91	0.80	0.85	4102
I-art	0.00	0.00	0.00	86
I-eve	0.00	0.00	0.00	73
I-geo	0.74	0.80	0.77	1449
I-gpe	0.90	0.27	0.42	33
I-nat	0.00	0.00	0.00	12
I-org	0.74	0.74	0.74	3254
I-per	0.84	0.86	0.85	3424
I-tim	0.82	0.65	0.73	1339
O	0.99	1.00	0.99	447795
accuracy			0.98	479600
macro avg	0.54	0.49	0.50	479600
weighted avg	0.98	0.98	0.98	479600

Hình 1: Kết quả F1-score của mô hình LSTM cho từng nhãn thực thể

- Kết quả của Accuracy cho từng nhãn thực thể là:

	Label	Accuracy
0	B-art	0.000000
1	B-eve	0.000000
2	B-geo	0.892683
3	B-gpe	0.940120
4	B-nat	0.000000
5	B-org	0.569345
6	B-per	0.759452
7	B-tim	0.799122
8	I-art	0.000000
9	I-eve	0.000000
10	I-geo	0.802622
11	I-gpe	0.272727
12	I-nat	0.000000
13	I-org	0.737246
14	I-per	0.859813
15	I-tim	0.653473
16	O	0.995900

Hình 2: Kết quả Accuracy của mô hình LSTM cho từng nhãn thực thể

Bài 2: Áp dụng mô hình trên để giải quyết bài toán NER cho tiếng Việt.

1. Đánh giá mô hình theo độ đo Accuracy và F1-score (micro và macro).

- Kết quả mô hình theo độ đo Accuracy là: 98.15%
- Kết quả mô hình theo độ đo F1-score micro là: 82.67%
- Kết quả mô hình theo độ đo F1-score macro là: 75.54%

2. Kết quả của F1-score và Accuracy cho từng nhãn thực thể.

- Kết quả của F1-score cho từng nhãn thực thể là:

	precision	recall	f1-score	support
B-AGE	0.89	0.93	0.91	568
B-DATE	0.97	0.96	0.97	1640
B-GENDER	0.90	0.90	0.90	447
B-JOB	0.61	0.37	0.46	172
B-LOCATION	0.93	0.87	0.90	4425
B-NAME	0.83	0.53	0.65	317
B-ORGANIZATION	0.84	0.82	0.83	769
B-PATIENT_ID	0.92	0.86	0.89	1978
B-SYMPOM_AND_DISEASE	0.90	0.81	0.85	1134
B-TRANSPORTATION	0.93	0.59	0.72	193
I-AGE	0.00	0.00	0.00	6
I-DATE	0.95	0.99	0.97	1726
I-JOB	0.57	0.04	0.07	114
I-LOCATION	0.94	0.81	0.87	4892
I-NAME	0.00	0.00	0.00	13
I-ORGANIZATION	0.88	0.82	0.85	2007
I-PATIENT_ID	0.00	0.00	0.00	27
I-SYMPOM_AND_DISEASE	0.91	0.73	0.81	1479
I-TRANSPORTATION	0.81	0.38	0.51	69
O	0.99	1.00	0.99	214584
accuracy			0.98	236560
macro avg	0.74	0.62	0.66	236560
weighted avg	0.98	0.98	0.98	236560

Hình 3: Kết quả F1-score của mô hình LSTM cho từng nhãn thực thể

- Kết quả của Accuracy cho từng nhãn thực thể là:

	Label	Accuracy
0	B-AGE	0.926056
1	B-DATE	0.957317
2	B-GENDER	0.903803
3	B-JOB	0.366279
4	B-LOCATION	0.866893
5	B-NAME	0.533123
6	B-ORGANIZATION	0.816645
7	B-PATIENT_ID	0.857432
8	B-SYMPTOM_AND_DISEASE	0.806878
9	B-TRANSPORTATION	0.590674
10	I-AGE	0.000000
11	I-DATE	0.990151
12	I-JOB	0.035088
13	I-LOCATION	0.814186
14	I-NAME	0.000000
15	I-ORGANIZATION	0.822123
16	I-PATIENT_ID	0.000000
17	I-SYMPTOM_AND_DISEASE	0.732252
18	I-TRANSPORTATION	0.376812
19	O	0.996370

Hình 4: Kết quả Accuracy của mô hình LSTM cho từng nhãn thực thể

Bài 3: Cải thiện kết quả của câu 1, 2 bằng cách cài đặt thêm các mô hình khác.

1. Cài đặt mô hình GRU cho Bài 1

(a) Đánh giá mô hình theo độ đo Accuracy và F1-score (micro và macro).

- Kết quả mô hình theo độ đo Accuracy là: 98.58%
- Kết quả mô hình theo độ đo F1-score micro là: 81.06%
- Kết quả mô hình theo độ đo F1-score macro là: 54.06%

(b) Kết quả của F1-score và Accuracy cho từng nhãn thực thể.

- Kết quả của F1-score cho từng nhãn thực thể là:

	precision	recall	f1-score	support
B-art	0.00	0.00	0.00	96
B-eve	0.86	0.08	0.14	79
B-geo	0.89	0.85	0.87	7380
B-gpe	0.95	0.94	0.95	3173
B-nat	0.60	0.08	0.14	38
B-org	0.76	0.74	0.75	3908
B-per	0.82	0.83	0.82	3359
B-tim	0.90	0.88	0.89	4102
I-art	0.00	0.00	0.00	86
I-eve	0.67	0.03	0.05	73
I-geo	0.84	0.75	0.79	1449
I-gpe	0.89	0.52	0.65	33
I-nat	0.00	0.00	0.00	12
I-org	0.82	0.76	0.79	3254
I-per	0.86	0.84	0.85	3424
I-tim	0.83	0.73	0.78	1339
O	0.99	1.00	1.00	447795
accuracy			0.99	479600
macro avg	0.69	0.53	0.56	479600
weighted avg	0.98	0.99	0.99	479600

Hình 5: Kết quả F1-score của mô hình GRU cho từng nhãn thực thể

- Kết quả của Accuracy cho từng nhãn thực thể là:

	Label	Accuracy
0	B-art	0.000000
1	B-eve	0.075949
2	B-geo	0.850678
3	B-gpe	0.941065
4	B-nat	0.078947
5	B-org	0.736182
6	B-per	0.829116
7	B-tim	0.882253
8	I-art	0.000000
9	I-eve	0.027397
10	I-geo	0.746032
11	I-gpe	0.515152
12	I-nat	0.000000
13	I-org	0.759988
14	I-per	0.842874
15	I-tim	0.729649
16	O	0.997774

Hình 6: Kết quả Accuracy của mô hình GRU cho từng nhãn thực thể

2. Cài đặt mô hình GRU cho Bài 2

- (a) Đánh giá mô hình theo độ đo Accuracy và F1-score (micro và macro).

- Kết quả mô hình theo độ đo Accuracy là: 98.27%
- Kết quả mô hình theo độ đo F1-score micro là: 83.95%
- Kết quả mô hình theo độ đo F1-score macro là: 77.35%

- (b) Kết quả của F1-score và Accuracy cho từng nhãn thực thể.

- Kết quả của F1-score cho từng nhãn thực thể là:

	precision	recall	f1-score	support
B-AGE	0.94	0.92	0.93	568
B-DATE	0.98	0.96	0.97	1640
B-GENDER	0.97	0.88	0.92	447
B-JOB	0.82	0.35	0.49	172
B-LOCATION	0.93	0.89	0.91	4425
B-NAME	0.94	0.54	0.69	317
B-ORGANIZATION	0.88	0.81	0.84	769
B-PATIENT_ID	0.96	0.87	0.91	1978
B-SYMPTOM_AND_DISEASE	0.91	0.79	0.85	1134
B-TRANSPORTATION	0.94	0.58	0.72	193
I-AGE	0.00	0.00	0.00	6
I-DATE	0.97	0.99	0.98	1726
I-JOB	0.67	0.09	0.16	114
I-LOCATION	0.94	0.81	0.87	4892
I-NAME	0.00	0.00	0.00	13
I-ORGANIZATION	0.91	0.80	0.85	2007
I-PATIENT_ID	0.00	0.00	0.00	27
I-SYMPTOM_AND_DISEASE	0.93	0.70	0.80	1479
I-TRANSPORTATION	0.96	0.39	0.56	69
O	0.99	1.00	0.99	214584
accuracy			0.98	236560
macro avg	0.78	0.62	0.67	236560
weighted avg	0.98	0.98	0.98	236560

Hình 7: Kết quả F1-score của mô hình GRU cho từng nhãn thực thể

- Kết quả của Accuracy cho từng nhãn thực thể là:

	Label	Accuracy
0	B-AGE	0.924296
1	B-DATE	0.960366
2	B-GENDER	0.881432
3	B-JOB	0.348837
4	B-LOCATION	0.887232
5	B-NAME	0.539432
6	B-ORGANIZATION	0.807542
7	B-PATIENT_ID	0.865521
8	B-SYMPTOM_AND_DISEASE	0.791887
9	B-TRANSPORTATION	0.580311
10	I-AGE	0.000000
11	I-DATE	0.986095
12	I-JOB	0.087719
13	I-LOCATION	0.814800
14	I-NAME	0.000000
15	I-ORGANIZATION	0.802691
16	I-PATIENT_ID	0.000000
17	I-SYMPTOM_AND_DISEASE	0.704530
18	I-TRANSPORTATION	0.391304
19	O	0.997684

Hình 8: Kết quả Accuracy của mô hình GRU cho từng nhãn thực thể