

Họ Tên: Phạm Quốc Thắng

MSSV: 23521431

BTTH: 2

Câu 1: Sử dụng mô hình BiLSTM để giải quyết bài toán NER cho tiếng Anh:

- Đánh giá mô hình theo độ đo F1-score (micro và macro):

F1-Score (Micro): 0.9671

F1-Score (Macro): 0.6192

- Kết quả F1-score và Accuracy cho từng nhãn thực thể:

| CLASSIFICATION REPORT | | | | | |
|-----------------------|-----------------------|-----------|--------|--------------|---------|
| | | precision | recall | f1-score | support |
| | B-art | 0.0000 | 0.0000 | 0.0000 | 56 |
| | B-eve | 0.5263 | 0.4444 | 0.4819 | 45 |
| | B-geo | 0.8630 | 0.9003 | 0.8813 | 5729 |
| | B-gpe | 0.9632 | 0.9313 | 0.9470 | 2330 |
| | B-nat | 0.5000 | 0.2381 | 0.3226 | 21 |
| | B-org | 0.7743 | 0.7094 | 0.7404 | 3042 |
| | B-per | 0.8247 | 0.8057 | 0.8151 | 2604 |
| ===== | Tag Accuracy F1-Score | | | | |
| | B-art | 0.0000 | 0.0000 | | |
| | B-eve | 0.4444 | 0.4819 | | |
| | B-geo | 0.9003 | 0.8813 | | |
| | B-gpe | 0.9313 | 0.9470 | | |
| | B-nat | 0.2381 | 0.3226 | | |
| | B-org | 0.7094 | 0.7404 | | |
| | B-per | 0.8057 | 0.8151 | | |
| | B-tim | 0.8847 | 0.8943 | | |
| | I-art | 0.0000 | 0.0000 | | |
| | I-eve | 0.4103 | 0.3441 | | |
| | I-geo | 0.8511 | 0.8316 | | |
| | I-gpe | 0.5484 | 0.6939 | | |
| | I-nat | 0.2500 | 0.1667 | | |
| | I-org | 0.7949 | 0.7793 | accuracy | |
| | I-per | 0.8515 | 0.8576 | macro avg | |
| | I-tim | 0.7765 | 0.7806 | weighted avg | |
| | O | 0.9907 | 0.9899 | 0.9671 | 156741 |
| | O | 0.9907 | 0.9899 | 0.6433 | 0.6110 |
| | O | 0.9907 | 0.9899 | 0.9664 | 0.9671 |
| | O | 0.9907 | 0.9899 | 0.9667 | 0.6192 |
| | O | 0.9907 | 0.9899 | 0.9667 | 156741 |

Câu 2: Sử dụng mô hình BiLSTM để giải quyết bài toán NER cho tiếng Việt:

- Đánh giá mô hình theo độ đo F1-score (micro và macro):

F1-Score (Micro): 0.9606

F1-Score (Macro): 0.7401

- Kết quả F1-score và Accuracy cho từng nhãn thực thể:

| CLASSIFICATION REPORT | | | | | |
|--|-----------------------|-----------|--------|----------|---------|
| | | precision | recall | f1-score | support |
| KẾT QUẢ CHI TIẾT THEO TỪNG NHÃN THỰC THỂ | B-AGE | 0.9110 | 0.9668 | 0.9381 | 572 |
| | B-DATE | 0.9511 | 0.9788 | 0.9647 | 1648 |
| | B-GENDER | 0.9025 | 0.9624 | 0.9315 | 452 |
| | B-JOB | 0.7059 | 0.5549 | 0.6214 | 173 |
| | B-LOCATION | 0.8846 | 0.9411 | 0.9120 | 4432 |
| | B-NAME | 0.7975 | 0.5981 | 0.6835 | 316 |
| | B-ORGANIZATION | 0.8851 | 0.7990 | 0.8398 | 771 |
| | B-PATIENT_ID | 0.9058 | 0.9561 | 0.9303 | 1982 |
| | B-SYMPOTM_AND_DISEASE | 0.9261 | 0.8169 | 0.8681 | 1136 |
| | B-TRANSPORTATION | 0.8925 | 0.8601 | 0.8760 | 193 |
| | I-AGE | 0.0000 | 0.0000 | 0.0000 | 6 |
| | I-DATE | 0.9703 | 0.9948 | 0.9824 | 1738 |
| | I-JOB | 0.7321 | 0.3596 | 0.4824 | 114 |
| | I-LOCATION | 0.9153 | 0.9056 | 0.9104 | 4903 |
| | I-NAME | 0.8333 | 0.3846 | 0.5263 | 13 |
| | I-ORGANIZATION | 0.8957 | 0.8361 | 0.8649 | 2014 |
| | I-PATIENT_ID | 0.0000 | 0.0000 | 0.0000 | 27 |
| | I-SYMPOTM_AND_DISEASE | 0.9200 | 0.7355 | 0.8175 | 1486 |
| | I-TRANSPORTATION | 0.7736 | 0.5942 | 0.6721 | 69 |
| | O | 0.9782 | 0.9840 | 0.9811 | 63339 |
| | accuracy | | | 0.9606 | 85384 |
| | macro avg | 0.7890 | 0.7114 | 0.7401 | 85384 |
| | weighted avg | 0.9597 | 0.9606 | 0.9596 | 85384 |

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

*Cải thiện câu 1 bằng GRU:

- Đánh giá mô hình theo độ đo F1-score (micro và macro):

F1-Score (Micro): 0.9646
F1-Score (Macro): 0.6174

- Kết quả F1-score và Accuracy cho từng nhãn thực thể:

| CLASSIFICATION REPORT | | | | | |
|--|----------|----------|--------------|--------|----------|
| | | | precision | recall | f1-score |
| KẾT QUẢ CHI TIẾT THEO TỪNG NHÃN THỰC THỂ | | | | | support |
| Tag | Accuracy | F1-Score | | | |
| B-art | 0.0000 | 0.0000 | B-art | 0.0000 | 0.0000 |
| B-eve | 0.3778 | 0.4857 | B-eve | 0.6800 | 0.3778 |
| B-geo | 0.9077 | 0.8675 | B-geo | 0.8307 | 0.9077 |
| B-gpe | 0.9266 | 0.9420 | B-gpe | 0.9579 | 0.9266 |
| B-nat | 0.3333 | 0.3333 | B-nat | 0.3333 | 0.3333 |
| B-org | 0.6686 | 0.7217 | B-org | 0.7838 | 0.6686 |
| B-per | 0.7611 | 0.7977 | B-per | 0.8381 | 0.7611 |
| B-tim | 0.8774 | 0.8896 | B-tim | 0.9022 | 0.8774 |
| I-art | 0.0000 | 0.0000 | I-art | 0.0000 | 0.0000 |
| I-eve | 0.1795 | 0.2500 | I-eve | 0.4118 | 0.1795 |
| I-geo | 0.8348 | 0.8182 | I-geo | 0.8023 | 0.8348 |
| I-gpe | 0.5484 | 0.6939 | I-gpe | 0.9444 | 0.5484 |
| I-nat | 0.5000 | 0.3333 | I-nat | 0.2500 | 0.5000 |
| I-org | 0.7602 | 0.7650 | I-org | 0.7698 | 0.7602 |
| I-per | 0.8462 | 0.8488 | I-per | 0.8514 | 0.8462 |
| I-tim | 0.7455 | 0.7605 | I-tim | 0.7762 | 0.7455 |
| O | 0.9907 | 0.9888 | O | 0.9869 | 0.9907 |
| | | | accuracy | | 0.9646 |
| | | | macro avg | 0.6540 | 0.6034 |
| | | | weighted avg | 0.9634 | 0.9646 |
| | | | | | 156741 |

*Cải thiện câu 2 bằng GRU:

- Đánh giá mô hình theo độ đo F1-score (micro và macro):

F1-Score (Micro): 0.9624
F1-Score (Macro): 0.7516

- Kết quả F1-score và Accuracy cho từng nhãn thực thể:

| CLASSIFICATION REPORT | | | | | |
|--|----------|----------|-----------------------|--------|----------|
| | | | precision | recall | f1-score |
| KẾT QUẢ CHI TIẾT THEO TỪNG NHÃN THỰC THỂ | | | | | support |
| Tag | Accuracy | F1-Score | | | |
| B-AGE | 0.9615 | 0.9394 | B-AGE | 0.9182 | 0.9615 |
| B-DATE | 0.9769 | 0.9705 | B-DATE | 0.9641 | 0.9769 |
| B-GENDER | 0.9668 | 0.9152 | B-GENDER | 0.8688 | 0.9668 |
| B-JOB | 0.7110 | 0.5721 | B-JOB | 0.4786 | 0.7110 |
| B-LOCATION | 0.9337 | 0.9230 | B-LOCATION | 0.9127 | 0.9337 |
| B-NAME | 0.6234 | 0.7177 | B-NAME | 0.8455 | 0.6234 |
| B-ORGANIZATION | 0.8314 | 0.8479 | B-ORGANIZATION | 0.8650 | 0.8314 |
| B-PATIENT_ID | 0.9712 | 0.9441 | B-PATIENT_ID | 0.9184 | 0.9712 |
| B-SYMPTOM_AND_DISEASE | 0.8565 | 0.8734 | B-SYMPTOM_AND_DISEASE | 0.8910 | 0.8565 |
| B-TRANSPORTATION | 0.8860 | 0.8814 | B-TRANSPORTATION | 0.8769 | 0.8860 |
| I-AGE | 0.0000 | 0.0000 | I-AGE | 0.0000 | 0.0000 |
| I-DATE | 0.9925 | 0.9883 | I-DATE | 0.9840 | 0.9925 |
| I-JOB | 0.4474 | 0.5204 | I-JOB | 0.6220 | 0.4474 |
| I-LOCATION | 0.9305 | 0.9133 | I-LOCATION | 0.8968 | 0.9305 |
| I-NAME | 0.4615 | 0.6000 | I-NAME | 0.8571 | 0.4615 |
| I-ORGANIZATION | 0.8491 | 0.8771 | I-ORGANIZATION | 0.9072 | 0.8491 |
| I-PATIENT_ID | 0.0000 | 0.0000 | I-PATIENT_ID | 0.0000 | 0.0000 |
| I-SYMPTOM_AND_DISEASE | 0.7719 | 0.8318 | I-SYMPTOM_AND_DISEASE | 0.9017 | 0.7719 |
| I-TRANSPORTATION | 0.7246 | 0.7353 | I-TRANSPORTATION | 0.7463 | 0.7246 |
| O | 0.9813 | 0.9815 | O | 0.9817 | 0.9813 |
| | | | accuracy | | 0.9624 |
| | | | macro avg | 0.7718 | 0.7439 |
| | | | weighted avg | 0.9622 | 0.9624 |
| | | | | | 85384 |

*Cải thiện câu 1 bằng XLM-RoBERTa:

- Tên model: xlm-roberta-base
- Link tải model: <https://huggingface.co/FacebookAI/xlm-roberta-base>
- Đánh giá mô hình theo độ đo F1-score (micro và macro):
F1-Score (Micro): 0.9662
F1-Score (Macro): 0.6088
- Kết quả F1-score và Accuracy cho từng nhãn thực thể:

| CLASSIFICATION REPORT | | | | | |
|--|--------------|-----------|--------|----------|---------|
| | | precision | recall | f1-score | support |
| | B-art | 0.7778 | 0.0660 | 0.1217 | 106 |
| | B-eve | 0.4667 | 0.2258 | 0.3043 | 62 |
| | B-geo | 0.8685 | 0.8921 | 0.8801 | 9155 |
| | B-gpe | 0.9638 | 0.9474 | 0.9555 | 3901 |
| | B-nat | 0.3171 | 0.2500 | 0.2796 | 52 |
| | B-org | 0.7885 | 0.7113 | 0.7479 | 5095 |
| | B-per | 0.8446 | 0.8587 | 0.8516 | 4127 |
| | B-tim | 0.8292 | 0.8918 | 0.8594 | 3365 |
| KẾT QUẢ CHI TIẾT THEO TÙNG NHÃN THỰC THỂ | I-art | 0.3333 | 0.0137 | 0.0263 | 73 |
| | I-eve | 0.0000 | 0.0000 | 0.0000 | 48 |
| | I-geo | 0.7902 | 0.7703 | 0.7801 | 1354 |
| | I-gpe | 1.0000 | 0.7358 | 0.8478 | 53 |
| | I-nat | 0.2500 | 0.3750 | 0.3000 | 8 |
| | I-org | 0.8116 | 0.6999 | 0.7516 | 3489 |
| | I-per | 0.8322 | 0.9402 | 0.8829 | 4833 |
| | I-tim | 0.7491 | 0.7889 | 0.7685 | 1113 |
| | O | 0.9913 | 0.9918 | 0.9916 | 178168 |
| | accuracy | | | 0.9662 | 215002 |
| | macro avg | 0.6832 | 0.5976 | 0.6088 | 215002 |
| | weighted avg | 0.9655 | 0.9662 | 0.9654 | 215002 |

*Cải thiện câu 2 bằng XLM-RoBERTa:

- Tên model: xlm-roberta-base
- Link tải model: <https://huggingface.co/FacebookAI/xlm-roberta-base>
- Đánh giá mô hình theo độ đo F1-score (micro và macro):
F1-Score (Micro): 0.9693
F1-Score (Macro): 0.7404
- Kết quả F1-score và Accuracy cho từng nhãn thực thể:

| CLASSIFICATION REPORT | | | | | |
|---|----------|----------|-----------------------|--------|----------|
| | | | precision | recall | f1-score |
| | | | | | support |
| KẾT QUẢ CHI TIẾT THEO TỪNG NHÃN THỰC THỂ | | | | | |
| Tag | Accuracy | F1-Score | | | |
| B-AGE | 0.9702 | 0.9710 | B-AGE | 0.9719 | 0.9702 |
| B-DATE | 0.9865 | 0.9870 | B-DATE | 0.9874 | 0.9865 |
| B-GENDER | 0.9478 | 0.9275 | B-GENDER | 0.9080 | 0.9478 |
| B-JOB | 0.8962 | 0.6537 | B-JOB | 0.5145 | 0.8962 |
| B-LOCATION | 0.9674 | 0.9502 | B-LOCATION | 0.9336 | 0.9674 |
| B-NAME | 0.9338 | 0.9481 | B-NAME | 0.9629 | 0.9338 |
| B-ORGANIZATION | 0.8499 | 0.8596 | B-ORGANIZATION | 0.8696 | 0.8499 |
| B-PATIENT_ID | 0.9893 | 0.9811 | B-PATIENT_ID | 0.9730 | 0.9893 |
| B-SYMPOTM_AND_DISEASE | 0.8958 | 0.8934 | B-SYMPOTM_AND_DISEASE | 0.8910 | 0.8958 |
| B-TRANSPORTATION | 0.9718 | 0.9718 | B-TRANSPORTATION | 0.9718 | 0.9718 |
| I-AGE | 0.0000 | 0.0000 | I-AGE | 0.0000 | 0.0000 |
| I-DATE | 0.9919 | 0.9896 | I-DATE | 0.9873 | 0.9919 |
| I-JOB | 0.0933 | 0.1618 | I-JOB | 0.6098 | 0.0933 |
| I-LOCATION | 0.9400 | 0.9490 | I-LOCATION | 0.9582 | 0.9400 |
| I-NAME | 0.0000 | 0.0000 | I-NAME | 0.0000 | 0.0000 |
| I-ORGANIZATION | 0.9150 | 0.8962 | I-ORGANIZATION | 0.8783 | 0.9150 |
| I-PATIENT_ID | 0.0000 | 0.0000 | I-PATIENT_ID | 0.0000 | 0.0000 |
| I-SYMPOTM_AND_DISEASE | 0.8788 | 0.8902 | I-SYMPOTM_AND_DISEASE | 0.9020 | 0.8788 |
| I-TRANSPORTATION | 0.6742 | 0.7911 | I-TRANSPORTATION | 0.9570 | 0.6742 |
| O | 0.9846 | 0.9868 | O | 0.9891 | 0.9846 |
| | | | accuracy | | 0.9693 |
| | | | macro avg | 0.7633 | 0.7443 |
| | | | weighted avg | 0.9701 | 0.9693 |
| | | | | | 172252 |