

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
B-tim	0.9042	0.8847	0.8943	3018
I-art	0.0000	0.0000	0.0000	51
I-eve	0.2963	0.4103	0.3441	39
I-geo	0.8129	0.8511	0.8316	1108
I-gpe	0.9444	0.5484	0.6939	31
I-nat	0.1250	0.2500	0.1667	4
I-org	0.7644	0.7949	0.7793	2477
I-per	0.8637	0.8515	0.8576	2613
I-tim	0.7849	0.7765	0.7806	935
O	0.9891	0.9907	0.9899	132638
accuracy			0.9671	156741
macro avg	0.6433	0.6110	0.6192	156741
weighted avg	0.9664	0.9671	0.9667	156741

KẾT QUẢ CHI TIẾT THEO TỪNG NHÃN THỰC THỂ		
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
I-per	0.8515	0.8576
I-tim	0.7765	0.7806
O	0.9907	0.9899

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:

KẾT QUẢ CHI TIẾT THEO TỪNG NHÂN THỰC THỂ			
Tag	Accuracy	F1-Score	
B-AGE	0.9668	0.9381	
B-DATE	0.9788	0.9647	
B-GENDER	0.9624	0.9315	
B-JOB	0.5549	0.6214	
B-LOCATION	0.9411	0.9120	
B-NAME	0.5981	0.6835	
B-ORGANIZATION	0.7990	0.8398	
B-PATIENT_ID	0.9561	0.9303	
B-SYMPTOM_AND_DISEASE	0.8169	0.8681	
B-TRANSPORTATION	0.8601	0.8760	
I-AGE	0.0000	0.0000	
I-DATE	0.9948	0.9824	
I-JOB	0.3596	0.4824	
I-LOCATION	0.9056	0.9104	
I-NAME	0.3846	0.5263	
I-ORGANIZATION	0.8361	0.8649	
I-PATIENT_ID	0.0000	0.0000	
I-SYMPTOM_AND_DISEASE	0.7355	0.8175	
I-TRANSPORTATION	0.5942	0.6721	
O	0.9840	0.9811	

CLASSIFICATION REPORT				
	precision	recall	f1-score	support
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-SYMPTOM_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-SYMPTOM_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 nhân thực thể:

			CLASSIFICATION REPORT				
				precision	recall	f1-score	support
KẾT QUẢ CHI TIẾT THEO TỪNG NHÃN THỰC THỂ							
Tag	Accuracy	F1-Score					
B-art	0.0000	0.0000	B-art	0.0000	0.0000	0.0000	56
B-eve	0.3778	0.4857	B-eve	0.6800	0.3778	0.4857	45
B-geo	0.9077	0.8675	B-geo	0.8307	0.9077	0.8675	5729
B-gpe	0.9266	0.9420	B-gpe	0.9579	0.9266	0.9420	2330
B-nat	0.3333	0.3333	B-nat	0.3333	0.3333	0.3333	21
B-org	0.6686	0.7217	B-org	0.7838	0.6686	0.7217	3042
B-per	0.7611	0.7977	B-per	0.8381	0.7611	0.7977	2604
B-tim	0.8774	0.8896	B-tim	0.9022	0.8774	0.8896	3018
I-art	0.0000	0.0000	I-art	0.0000	0.0000	0.0000	51
I-eve	0.1795	0.2500	I-eve	0.4118	0.1795	0.2500	39
I-geo	0.8348	0.8182	I-geo	0.8023	0.8348	0.8182	1108
I-gpe	0.5484	0.6939	I-gpe	0.9444	0.5484	0.6939	31
I-nat	0.5000	0.3333	I-nat	0.2500	0.5000	0.3333	4
I-org	0.7602	0.7650	I-org	0.7698	0.7602	0.7650	2477
I-per	0.8462	0.8488	I-per	0.8514	0.8462	0.8488	2613
I-tim	0.7455	0.7605	I-tim	0.7762	0.7455	0.7605	935
0	0.9907	0.9888	0	0.9869	0.9907	0.9888	132638
			accuracy			0.9646	156741
			macro avg	0.6540	0.6034	0.6174	156741
			weighted avg	0.9634	0.9646	0.9638	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	support
KẾT QUẢ CHI TIẾT THEO TỪNG NHÃN THỰC THỂ								
	Tag	Accuracy	F1-Score					
	B-AGE	0.9615	0.9394	B-AGE	0.9182	0.9615	0.9394	572
	B-DATE	0.9769	0.9705	B-DATE	0.9641	0.9769	0.9705	1648
	B-GENDER	0.9668	0.9152	B-GENDER	0.8688	0.9668	0.9152	452
	B-JOB	0.7110	0.5721	B-JOB	0.4786	0.7110	0.5721	173
	B-LOCATION	0.9337	0.9230	B-LOCATION	0.9127	0.9337	0.9230	4432
	B-NAME	0.6234	0.7177	B-NAME	0.8455	0.6234	0.7177	316
	B-ORGANIZATION	0.8314	0.8479	B-ORGANIZATION	0.8650	0.8314	0.8479	771
	B-PATIENT_ID	0.9712	0.9441	B-PATIENT_ID	0.9184	0.9712	0.9441	1982
B-SYMPTOM_AND_DISEASE		0.8565	0.8734	B-SYMPTOM_AND_DISEASE	0.8910	0.8565	0.8734	1136
B-TRANSPORTATION		0.8860	0.8814	B-TRANSPORTATION	0.8769	0.8860	0.8814	193
	I-AGE	0.0000	0.0000	I-AGE	0.0000	0.0000	0.0000	6
	I-DATE	0.9925	0.9883	I-DATE	0.9840	0.9925	0.9883	1738
	I-JOB	0.4474	0.5204	I-JOB	0.6220	0.4474	0.5204	114
	I-LOCATION	0.9305	0.9133	I-LOCATION	0.8968	0.9305	0.9133	4903
	I-NAME	0.4615	0.6000	I-NAME	0.8571	0.4615	0.6000	13
	I-ORGANIZATION	0.8491	0.8771	I-ORGANIZATION	0.9072	0.8491	0.8771	2014
	I-PATIENT_ID	0.0000	0.0000	I-PATIENT_ID	0.0000	0.0000	0.0000	27
I-SYMPTOM_AND_DISEASE		0.7719	0.8318	I-SYMPTOM_AND_DISEASE	0.9017	0.7719	0.8318	1486
I-TRANSPORTATION		0.7246	0.7353	I-TRANSPORTATION	0.7463	0.7246	0.7353	69
	O	0.9813	0.9815	O	0.9817	0.9813	0.9815	63339
				accuracy			0.9624	85384
				macro avg	0.7718	0.7439	0.7516	85384
				weighted avg	0.9622	0.9624	0.9620	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
			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
			0	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

KẾT QUẢ CHI TIẾT THEO TỪNG NHÃN THỰC THỂ		
Tag	Accuracy	F1-Score
B-art	0.0660	0.1217
B-eve	0.2258	0.3043
B-geo	0.8921	0.8801
B-gpe	0.9474	0.9555
B-nat	0.2500	0.2796
B-org	0.7113	0.7479
B-per	0.8587	0.8516
B-tim	0.8918	0.8594
I-art	0.0137	0.0263
I-eve	0.0000	0.0000
I-geo	0.7703	0.7801
I-gpe	0.7358	0.8478
I-nat	0.3750	0.3000
I-org	0.6999	0.7516
I-per	0.9402	0.8829
I-tim	0.7889	0.7685
0	0.9918	0.9916

***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:

