

## TỔNG HỢP KẾT QUẢ BÀI THỰC HÀNH LAB3

### Named Entity Recognition (NER)

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

	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	0.86	0.82	7380
B-gpe	0.95	0.90	0.92	3173
B-nat	0.00	0.00	0.00	38
B-org	0.73	0.53	0.61	3908
B-per	0.80	0.72	0.76	3359
B-tim	0.90	0.75	0.82	4102
I-art	0.00	0.00	0.00	86
I-eve	0.00	0.00	0.00	73
I-geo	0.84	0.56	0.67	1449
I-gpe	0.00	0.00	0.00	33
I-nat	0.00	0.00	0.00	12
I-org	0.70	0.67	0.68	3254
I-per	0.81	0.82	0.81	3424
I-tim	0.90	0.41	0.56	1339
O	0.99	1.00	0.99	447795
accuracy			0.98	479600
macro avg	0.49	0.42	0.45	479600
weighted avg	0.98	0.98	0.98	479600

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

	precision	recall	f1-score	support
B-AGE	0.00	0.00	0.00	77
B-DATE	0.00	0.00	0.00	272
B-GENDER	0.00	0.00	0.00	65
B-JOB	0.00	0.00	0.00	37
B-LOCATION	0.60	0.40	0.48	615
B-NAME	0.00	0.00	0.00	39
B-ORGANIZATION	0.00	0.00	0.00	107
B-PATIENT_ID	0.00	0.00	0.00	333
B-SYMPTOM_AND_DISEASE	0.00	0.00	0.00	155
B-TRANSPORTATION	0.00	0.00	0.00	24
I-AGE	0.00	0.00	0.00	1
I-DATE	0.00	0.00	0.00	244
I-JOB	0.00	0.00	0.00	20
I-LOCATION	0.47	0.71	0.57	661
I-NAME	0.00	0.00	0.00	2
I-ORGANIZATION	0.53	0.24	0.33	263
I-PATIENT_ID	0.00	0.00	0.00	3
I-SYMPTOM_AND_DISEASE	0.00	0.00	0.00	199
I-TRANSPORTATION	0.00	0.00	0.00	7
O	0.94	1.00	0.97	26996
accuracy			0.92	30120
macro avg	0.13	0.12	0.12	30120
weighted avg	0.87	0.92	0.89	30120

**Câu 3: Cải thiện kết quả của câu 1, 2 bằng cách cài đặt thêm các mô hình khác. (Mô hình GRU+CNN)**

**a) Câu 1:**

	precision	recall	f1-score	support
B-art	1.00	0.09	0.17	96
B-eve	0.75	0.42	0.54	79
B-geo	0.91	0.96	0.94	7380
B-gpe	0.98	0.95	0.96	3173
B-nat	0.94	0.42	0.58	38
B-org	0.89	0.85	0.87	3908
B-per	0.93	0.92	0.93	3359
B-tim	0.96	0.94	0.95	4102
I-art	0.96	0.27	0.42	86
I-eve	0.60	0.33	0.42	73
I-geo	0.93	0.92	0.92	1449
I-gpe	0.91	0.61	0.73	33
I-nat	0.27	0.25	0.26	12
I-org	0.92	0.93	0.93	3254
I-per	0.91	0.97	0.94	3424
I-tim	0.90	0.90	0.90	1339
0	1.00	1.00	1.00	447795
accuracy			0.99	479600
macro avg	0.87	0.69	0.73	479600
weighted avg	0.99	0.99	0.99	479600

**b) Câu 2:**

	precision	recall	f1-score	support
B-AGE	0.89	0.94	0.91	77
B-DATE	0.97	0.96	0.97	272
B-GENDER	0.92	0.72	0.81	65
B-JOB	0.75	0.08	0.15	37
B-LOCATION	0.94	0.95	0.95	615
B-NAME	0.70	0.54	0.61	39
B-ORGANIZATION	0.82	0.87	0.85	107
B-PATIENT_ID	0.93	0.97	0.95	333
B-SYMPTOM_AND_DISEASE	0.88	0.94	0.91	155
B-TRANSPORTATION	0.90	0.79	0.84	24
I-AGE	0.00	0.00	0.00	1
I-DATE	0.96	0.99	0.98	244
I-JOB	0.56	0.25	0.34	20
I-LOCATION	0.95	0.93	0.94	661
I-NAME	0.00	0.00	0.00	2
I-ORGANIZATION	0.88	0.92	0.90	263
I-PATIENT_ID	0.00	0.00	0.00	3
I-SYMPTOM_AND_DISEASE	0.91	0.89	0.90	199
I-TRANSPORTATION	1.00	0.43	0.60	7
0	1.00	1.00	1.00	26996
accuracy			0.99	30120
macro avg	0.75	0.66	0.68	30120
weighted avg	0.99	0.99	0.99	30120