

BÁO CÁO HUẤN LUYỆN BẰNG LSTM

Huấn luyện dựa trên bộ ViHSD:

Điều kiện chọn hàm:

```
if f1_test_pos_1 > best_f1_score-0.03 and test_acc > best_acc_score-0.03 and  
f1_test_pos_1 + test_acc > best_f1_score+best_acc_score:
```

Hàm mất mát sử dụng weight:

```
def weighted_binary_cross_entropy(y_true, y_pred, pos_weight):  
    """  
    Weighted Binary Cross Entropy (WBCE) = - (w * y * log(p) + (1 - y) * log(1 - p))  
    Trong đó:  
    y: Nhãn thực tế (0 hoặc 1).  
    p: Xác suất dự đoán cho lớp positive (nhãn 1).  
    w: Trọng số cho lớp positive.  
    """  
    epsilon = 1e-7  
    y_pred = torch.clamp(y_pred, epsilon, 1 - epsilon) # giới hạn giá trị dự đoán tro  
    bce = - (pos_weight * y_true * torch.log(y_pred) +  
            | (1 - y_true) * torch.log(1 - y_pred)) # Binary Cross Entropy trong đ  
    return torch.mean(bce)  
  
from collections import Counter  
print(Counter(y_train))  
neg_count = Counter(y_train)[0]  
pos_count = Counter(y_train)[1]  
pos_weight = torch.tensor((neg_count/2) / pos_count)
```

✓ 0.0s

Counter({0: 18582, 1: 3960})

Chọn save model:

Train loss: 0.0170, Train acc: 0.7107, F1 train pos 1 score: 0.5744, F1 train
macro score: 0.7420 |
Test loss: 0.0205, Test acc: 0.7177, F1 test pos 1 score: 0.4613, F1 test
macro score: 0.6763
save model at this epoch

Kết quả train:

ViHSD					ViCTSD				
	precision	recall	f1-score	support		precision	recall	f1-score	support
0	0.89	0.90	0.90	5548	0	0.89	0.77	0.82	890
1	0.50	0.47	0.48	1132	1	0.11	0.25	0.16	110
accuracy			0.83	6680	accuracy			0.71	1000
macro avg	0.70	0.68	0.69	6680	macro avg	0.50	0.51	0.49	1000
weighted avg	0.83	0.83	0.83	6680	weighted avg	0.81	0.71	0.75	1000
Bộ dữ liệu có độ toxic cao, từng từ đơn đều có thể mang tính toxic					Bộ dữ liệu có độ toxic không cao, ý nghĩa cả câu hợp lại mới tạo nên sự toxic				

Huấn luyện trên tập ViCTSD

Train loss: 0.0269, Train acc: 0.7438, F1 train pos 1 score: 0.2017, F1 train macro score: 0.5308 |

Test loss: 0.0281, Test acc: 0.7842, F1 test pos 1 score: 0.1797, F1 test macro score: 0.5319

save model at this epoch

ViHSD					ViCTSD				
	precision	recall	f1-score	support		precision	recall	f1-score	support
0	0.84	0.94	0.89	5548	0	0.89	0.87	0.88	890
1	0.29	0.12	0.17	1132	1	0.12	0.15	0.13	110
accuracy			0.80	6680	accuracy			0.79	1000
macro avg	0.56	0.53	0.53	6680	macro avg	0.51	0.51	0.51	1000
weighted avg	0.75	0.80	0.77	6680	weighted avg	0.81	0.79	0.80	1000
Bộ dữ liệu có độ toxic cao, từng từ đơn đều có thể mang tính toxic					Bộ dữ liệu có độ toxic không cao, ý nghĩa cả câu hợp lại mới tạo nên sự toxic				

Mô hình BiLSTM huấn luyện trên tập ViHSD:

ViHSD					ViCTSD				
	precision	recall	f1-score	support		precision	recall	f1-score	support
0	0.91	0.82	0.86	5548	0	0.89	0.61	0.72	890
1	0.40	0.58	0.47	1132	1	0.11	0.41	0.18	110
accuracy			0.78	6680	accuracy			0.59	1000
macro avg	0.65	0.70	0.67	6680	macro avg	0.50	0.51	0.45	1000
weighted avg	0.82	0.78	0.79	6680	weighted avg	0.81	0.59	0.66	1000
Bộ dữ liệu có độ toxic cao, từng từ đơn đều có thể mang tính toxic					Bộ dữ liệu có độ toxic không cao, ý nghĩa cả câu hợp lại mới tạo nên sự toxic				