NLU Assignment3

Vipul Kumar Rathore
rathorevipul28@gmail.com

1 Model 1 - LSTM

1.1 Model

- In my LSTM model,I am using embedding layer having input dimension of 11311(the no of words in dataset), output dimension of 100 and input length as maximum length which is 100.
- On top of this embedding layer, I am building a bidirectional LSTM layer having 128 hidden units and recurrent dropout of 0.1.
- In my output layer, I am using softmax activation.

1.2 cross-validation and Results

- I am splitting the original dataset into 80% training and 20% test.
- While training my model, I am again making a validation set with 10% split and using 5 epochs with batch size as 32 for cross-validation.At the end of 5th epoch, the training accuracy is 99.19% whereas the validation set accuracy is 98.51%.

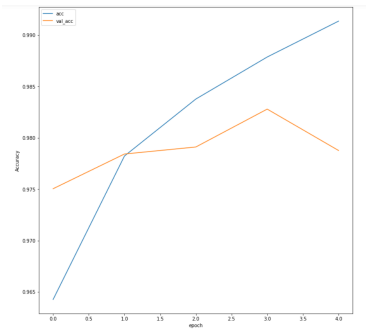

Figure 1: Training and validation accuracy for LSTM.

2 Model 2 - CRF

2.1 Model

- In my CRF model, I am making use of word identity, word suffix, word shape and word POS tag, also some information from nearby words is used.
- For training, I am using L-BFGS algorithm with Elastic Net(L1 + L2) regularization and maximum iteration as 1000.

2.2 Ablation study analysis

- I am trying to remove many features one by one like 'bias','word.lower()','word.isupper()','word.istitle()',etc.
- Using above technique, I observed that the f1-score dropped drastically when I tried to remove the feature 'word.lower()', while for other features it didn't vary much.This implies that the feature 'word.lower()' is the most important feature in our model.

2.3 Validation and Results

- I splitted my dataset into 80% training and 20% test.
- The f1-measure on test data is **0.7744**.
- The other measures on the test data are showing in figure below.

	precision	recall	f1-score	support
D	0.853	0.578	0.689	2291
O	0.802	0.966	0.876	10785
T	0.768	0.304	0.436	2432
avg / total	0.804	0.805	0.779	15508

Figure 2: F-1 score table for CRF