

CS772 - DL4NLP

Assignment 3

Group 24

170050003 - Jigar Makwana
170050031 - Shashank Kumar
170050052 - Devki Nandan Malav
170050057 - Arpit Menaria

Problem Statement

- Use of RNN, LSTM, Bi-LSTM, GRU, and Bi-GRU for solving the same problem of Assignment 1 and Assignment 2
- Compare Accuracy, Precision, Recall, F1-Score for above-mentioned models

The models should be trained with:

- Data imbalance addressed (which you have already done in assignment 2)
- Scores with and without pre-trained embeddings (perform experiments with only one (any one of word2vec, fast-text or glove))

Architecture

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 15, 300)	5404200
GRU_1 (GRU)	(None, 15, 256)	428544
dropout_1 (Dropout)	(None, 15, 256)	0
flatten (Flatten)	(None, 3840)	0
dense_1 (Dense)	(None, 5)	19205
softmax (Activation)	(None, 5)	0

=====

Total params: 5,851,949
Trainable params: 447,749
Non-trainable params: 5,404,200

Library functions used

- sklearn and tensorflow.keras for building, training and testing model
- csv, numpy, nltk, os, random, re for preprocessing and other calculative stuff.
- gensim for the embedding (glove)

Model Comparisons

We experimented with number of model and dense layers.

MODEL	WITH PRE-TRAINED EMBEDDINGS (GLOVE)				WITHOUT PRE-TRAINED EMBEDDINGS			
	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1-score
1 LSTM + 1 Dense	0.57	0.64	0.57	0.60	0.62	0.62	0.62	0.61
1 GRU + 1 Dense	0.61	0.64	0.61	0.62	0.61	0.64	0.61	0.62
1 RNN + 1 Dense	0.55	0.62	0.55	0.58	0.50	0.63	0.50	0.54
1 Bi-LSTM + 1 Dense	0.54	0.64	0.54	0.57	0.45	0.65	0.45	0.50
1 Bi-GRU + 1 Dense	0.57	0.65	0.57	0.60	0.54	0.66	0.54	0.57

MODEL	WITH PRE-TRAINED EMBEDDINGS (GLOVE)				WITHOUT PRE-TRAINED EMBEDDINGS			
	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1-score
1 LSTM + 2 Dense	0.57	0.63	0.57	0.59	0.59	0.63	0.59	0.60
1 GRU + 2 Dense	0.57	0.54	0.57	0.55	0.52	0.65	0.52	0.56
1 RNN + 2 Dense	0.55	0.63	0.55	0.57	0.58	0.62	0.58	0.59
1 Bi-LSTM + 2 Dense	0.62	0.48	0.62	0.53	0.61	0.61	0.61	0.61
1 Bi-GRU + 2 Dense	0.16	0.03	0.16	0.05	0.22	0.07	0.22	0.10

MODEL	WITH PRE-TRAINED EMBEDDINGS (GLOVE)				WITHOUT PRE-TRAINED EMBEDDINGS			
	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1-score
2 LSTM + 1 Dense	0.60	0.65	0.60	0.61	0.55	0.65	0.55	0.57
2 GRU + 1 Dense	0.58	0.65	0.58	0.61	0.61	0.64	0.61	0.62
2 RNN + 1 Dense	0.52	0.64	0.52	0.56	0.50	0.66	0.50	0.55
2 Bi-LSTM + 1 Dense	0.56	0.66	0.56	0.58	0.57	0.63	0.57	0.59
2 Bi-GRU + 1 Dense	0.59	0.64	0.59	0.61	0.51	0.66	0.51	0.56

MODEL	WITH PRE-TRAINED EMBEDDINGS (GLOVE)				WITHOUT PRE-TRAINED EMBEDDINGS			
	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1-score
2 LSTM + 2 Dense	0.56	0.65	0.56	0.58	0.56	0.66	0.56	0.60
2 GRU + 2 Dense	0.56	0.66	0.56	0.59	0.58	0.65	0.58	0.61
2 RNN + 2 Dense	0.57	0.63	0.57	0.59	0.52	0.65	0.52	0.56
2 Bi-LSTM + 2 Dense	0.50	0.66	0.50	0.55	0.53	0.64	0.53	0.56
2 Bi-GRU + 2 Dense	0.58	0.65	0.58	0.60	0.58	0.66	0.58	0.61

Confusion Matrix

```
[[ 774   156   206    50    85]
 [ 219   102   192    53    64]
 [ 141    93   413   130   134]
 [  79    52   349   380   544]
 [ 173    94   401   642  4474]]
```

This is our best confusion matrix when we trained our model with 1 GRU and 1 Dense layer without any pre-trained embeddings (the one with pre-trained embeddings is also very close to this). This has best balance of true positives and one of the highest accuracy among other models.

Some qualitative examples

- This product is amazing but received broken.

Predicted class - 1

Probabilities - [9.8662829e-01, 1.3318885e-02, 3.9498471e-05, 2.7885380e-06, 1.0511705e-05]

- London is very expensive but very comfortable to live in.

Predicted class - 3

Probabilities - [0.19348499, 0.00388968, 0.40457958, 0.18472451, 0.21332121]

- Delhi is very polluted, crowded and that makes life hard.

Predicted class - 1

Probabilities - [0.9756824, 0.02074364, 0.0015215, 0.00107307, 0.00097942]