INLP – Assignment 2

Course Coordinator: Manish Srivatsava
Name : Lakshmipathi Balaji
Roll No: 2021114007
Mail : lakshmipathi.balaji@gmail.com

Assignment details:
– Pre-training on sst and multi_nli datasets using Elmo architectures to get wrod embeddings.
–Using these word embeddings for training on sst and multi_nli datasets where these tasks are considered as classification tasks.

Procedure:
-Got the csv files of datasets from huggingface.
-Using elmo architecture hiddens_states of a token for every word are retrieved from every bILM layers.
-Using these hidden_states final elmo_word_embeddings are built by linearly summing the hidden_states and concatenating it with the wordtovec embedding.
- Note that, the baseline embeddings used are from wordtovec architecture with the help of the function nn.Embedding in PyTorch.
- After building these final_elmo_embeddings for every word these are used to train and test on the sst and multi_nli datasets.

Things learnt:
- Making a pretrained model for word_embeddings.
- Dealing with pretrained models.
- Building some simple architectures for NLP tasks.

Results:

```
Epoch: 0, Loss: 81646.8980255127        Epoch: 0, Loss: 28573.364916801453
Epoch: 1, Loss: 56909.94713973999       Epoch: 1, Loss: 19674.288639068604
Epoch: 2, Loss: 49188.96164703369       Epoch: 2, Loss: 17197.329107284546
Epoch: 3, Loss: 44612.11266326904       Epoch: 3, Loss: 15785.196628570557
Epoch: 4, Loss: 41435.01380157471       Epoch: 4, Loss: 14690.77370262146
Epoch: 5, Loss: 38945.197368621826      Epoch: 5, Loss: 13836.60588645935
Epoch: 6, Loss: 36966.24389076233       Epoch: 6, Loss: 13156.581552505493
Epoch: 7, Loss: 35268.80284881592       Epoch: 7, Loss: 12520.411077976227
Epoch: 8, Loss: 33659.39066696167       Epoch: 8, Loss: 11986.980777740479
Epoch: 9, Loss: 32285.30918121338       Epoch: 9, Loss: 11525.59891462326
Epoch: 10, Loss: 31063.810062408447     Epoch: 10, Loss: 11119.788970947266
Epoch: 11, Loss: 29899.337154388428     Epoch: 11, Loss: 10792.89938735962
Epoch: 12, Loss: 28876.67477798462      Epoch: 12, Loss: 10493.37898015976
Epoch: 13, Loss: 27989.037992477417     Epoch: 13, Loss: 10205.899777412415
Epoch: 14, Loss: 27084.67932510376      Epoch: 14, Loss: 9947.063751220703
Epoch: 15, Loss: 26116.95749282837      Epoch: 15, Loss: 9722.292019605637
Epoch: 16, Loss: 25265.8355178833       Epoch: 16, Loss: 9502.305861473083
Epoch: 17, Loss: 24569.103590011597     Epoch: 17, Loss: 9282.042249202728
Epoch: 18, Loss: 23996.46361541748      Epoch: 18, Loss: 9035.144765377045
Epoch: 19, Loss: 23378.246620178223     Epoch: 19, Loss: 8821.214429140091
Epoch: 20, Loss: 22647.462894439697     Epoch: 20, Loss: 8628.688705921173
Epoch: 21, Loss: 21956.284419059753     Epoch: 21, Loss: 8449.707573413849
Epoch: 22, Loss: 21334.254138946533     Epoch: 22, Loss: 8285.101417541504
Epoch: 23, Loss: 20793.886729240417     Epoch: 23, Loss: 8135.375336647034
Epoch: 24, Loss: 20308.077460289        Epoch: 24, Loss: 7996.411991119385
```

- Pretraining on train_data and test_data as I assumed to have seperate pipeline for training and testing data even at the word embeddings level as usually pretraining happens at a large scale for the SST_model. I have only done the assignment for SST_dataset because of time constraints.

- Accuracy and classification report with the help of sklearn.

	precision	recall	f1-score	support
0	0.68	0.99	0.81	171
1	1.00	0.84	0.91	234
2	0.78	0.97	0.87	186
3	1.00	0.70	0.82	380
4	0.99	0.85	0.92	191
5	1.00	0.75	0.86	20
6	0.93	0.96	0.94	336
7	0.85	1.00	0.92	335
8	0.99	0.97	0.98	241
9	0.98	0.89	0.93	101
10	0.88	1.00	0.94	15
accuracy			0.90	2210
macro avg	0.92	0.90	0.90	2210
weighted avg	0.92	0.90	0.90	2210