CS769 Assignment1

1. I have implemented the DAN network using –
    a. Embedding layer – (|vocab| * 300)
    b. FC layer
    c. FC layer
    d. FC layer – (300 * |tags|)

2. I tried with different activation functions: tanh, relu, leakyRelu :
    - LeakyRelu gave me the best performance.

3. To initialize weights, I have used Xavier initializations.

4. I sorted the input sequences based on their length to have minimal padding in each batch.

5. I am using the fasttext - 'wiki-news-300d-1M' embeddings –
    - For the words of vocab which are not in pre-trained embeddings, I am initializing them with Xavier values.
    - I have saved my embeddings as a .npy file and I am loading them based on the args. I have copied those embeddings in the zip folder.
    - Files – 'embs_cfimdb.npy' , 'embs_sst.npy'

6. I implemented word drop out with 0.2 prob, where I drop out the words randomly. But this wasn't giving me much accuracy boost, so I have commented it out.

7. Last Run Accuracy results -

```
-Accuracy: 0.4457 (985/2210)
-Save predictions to 9082943094/sst-test-output.txt
-Accuracy: 0.4223 (465/1101)
-Save predictions to 9082943094/sst-dev-output.txt


-Accuracy: 0.4795 (234/488)
-Save predictions to 9082943094/cfimdb-test-output.txt
-Accuracy: 0.9388 (230/245)
-Save predictions to 9082943094/cfimdb-dev-output.txt


Best achieved around 0.454 (sst-test), 0.43 (sst-dev), 0.945
(cfimdb_dev)
```