Mohammad Rehan Khan 2020MT10822

Saksham Singh 2020MT10841

Assignment 4

1. CNN

   Train Accuracy: 94.7
   Test accuracy: 12.8
   Test accuracy with dropout(0.5)- 13.87

   Hyperparameters
   . Number of epochs: 20
   . Batch size: 8
   . Learning rate: $5* 10^{-5}$
   . Optimizer: Adam

2. Bidirectional RNN

   Train accuracy with **dropout**: 82.9767
   Test accuracy with dropout: 46.8421

   Train accuracy **without dropout**: 97.1608
   Test accuracy without dropout: 42.9473

   Hyperparameters
   . Number of epochs: 25
   . Batch size: 100
   . Vocabulary size: 27894
   . Dropout: 0.3
   . Learning rate: 0.001
   . Optimizer: Adam

   Dropout prevents the model from overfitting. Also, the test accuracy obtained with a dropout is better than the accuracy obtained without a dropout.

3. Competitive Part

   Non comp test loss: 1.6
   Training loss: 0.6

   Hyperparameters

Number of epochs: 3
. Batch size: 8
. Dropout: 0.1
. Learning rate: 0.00002
. Optimizer: Adam

Architecture – Transformer
Pretrained model- distilBERT

Some observations and models tried

1. Combining the image representation with the text representation and passing through the standard neural network (fully connected layers) did not improve the accuracy much. (approximately 46%)
2. Passing the image representation as “h0” (for the RNN/LSTM layer) gave an accuracy of around 47%.
3. Using the weight matrix(globe embedding) obtained from the pre-trained embedding along with the CNN with max pooling improved the accuracy. The accuracy obtained in this case was around 57%.
4. Adam optimizer performed better than SGD in all the models we tried.