Assignment 3 Report

Solution 1:

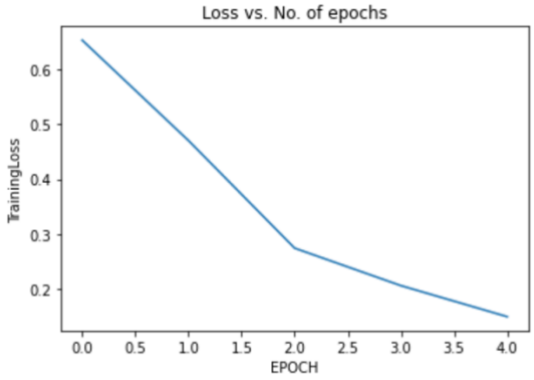
With Attention:

A. Choice of Hyperparameters:

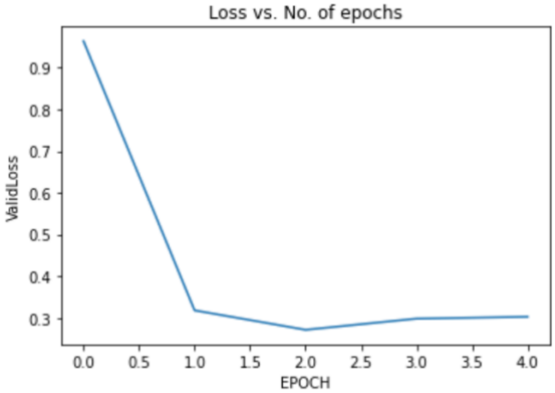
- **a. Epochs = 5**
- **b. Dropout = 0.4**

B. Training Loss & Validation Loss vs No. of Epochs curve:

- **a. Training Loss Curve:**



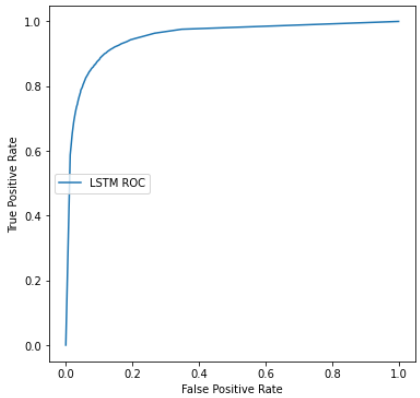
- **b. Validation Loss Curve:**



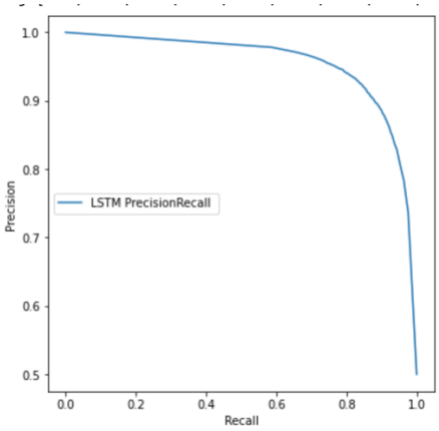
C. Testing Accuracy + Confusion Matrix:

- **a. 88.87%**
- **b. [11515. 985.]**
  [1797. 10703.]

D. ROC Curve:



E. PR Curve:



F. Conclusion from curves D and E:

ROC curve shifted more towards top left which tells that it gives better performance. **PR** curve also gives better visualization that it is tilted in right top corner. This tells that model is predicting classes near to the ground truth. High precision and high recall is always welcomed. This means model is giving accurate results and giving all positive results.

Without Attention:

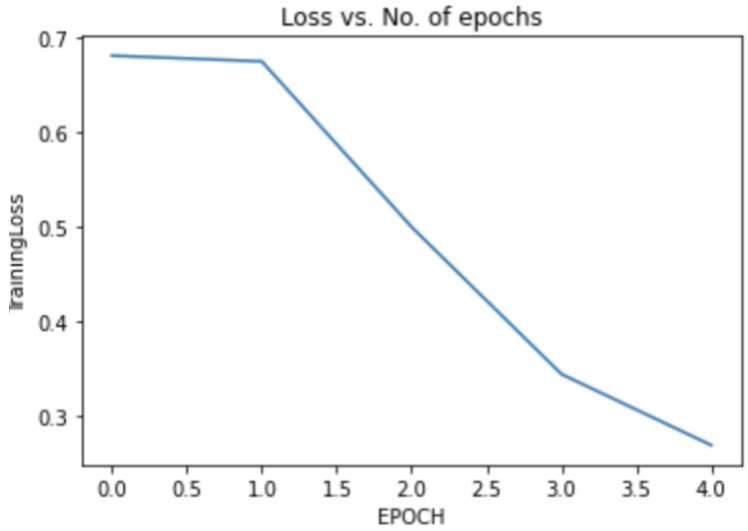
A. Choice of Hyperparameters:

a. **Epochs = 5**

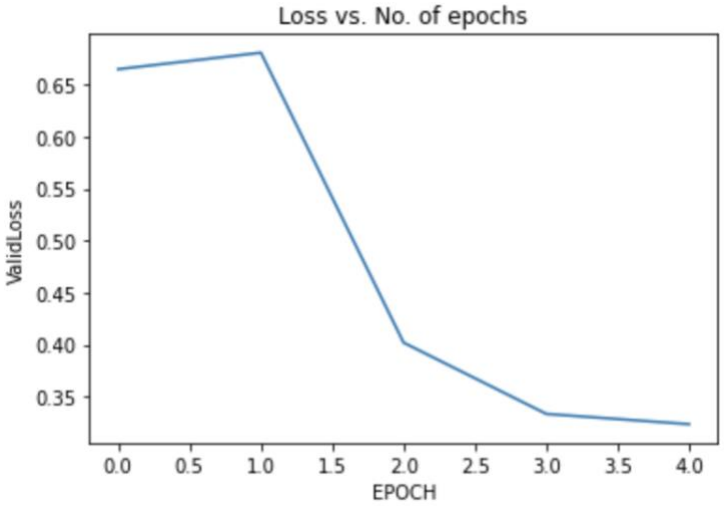
b. **Dropout = 0.4**

B. Training Loss & Validation Loss vs No. of Epochs curve:

a. **Training Loss Curve:**



b. **Validation Loss Curve:**

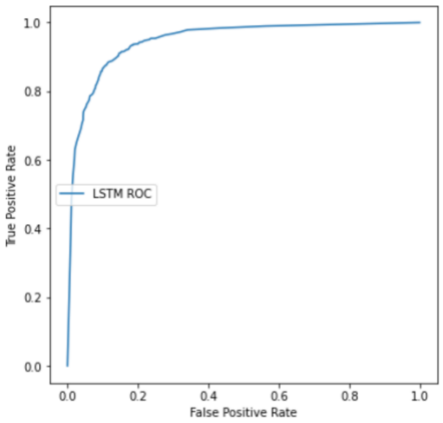


C. Testing Accuracy + Confusion Matrix:

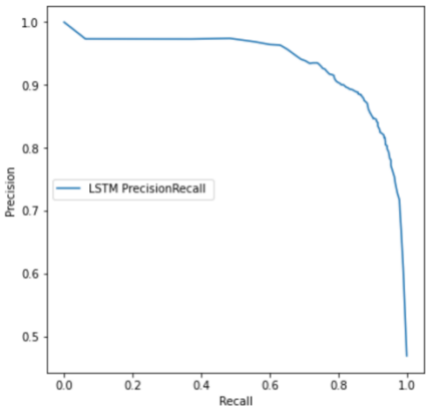
a. **88.21%**

b. `[1161.  167.]`
`[ 131. 1041.]`

D. ROC Curve:



E. PR Curve:



Supporting Theory for Question 1

1. **Confusion matrix:** Confusion Matrix tells about the true positives, true negatives, false positive, false negatives. In simple terms, Confusion matrix tells about how many times “class A” is classified as class A and class B actually. And same for “class B”. This is best way to evaluate your model. That is How many times our model was confusion with pullover but it was coat in reality. This is where we need to create a predicted label array along with the original labels. So as to identify the parameters.
2. **TP,FP,TN,FN, TPR, FPR,TNR,FNR**: for better performance, TPR and TNR should be high and rest low. Because we want prediction to be similar to the ground truth.

   a. True Positive: # of times coat is predicted as coat
   b. False Positive: # of times pullover is predicted as coat
   c. True Negative: # of times pullover is predicted as pullover d. False Negative: # of times coat is predicted as pullover.
   e. True Positive Rate: #tp/#tp+#fn
   f. False Positive Rate: #fp/#tn+#fp
   g. True Negative Rate: #tn/#tn+#fp
   h. False negative Rate: #fn/#tp+#fn

3. **ROC Curve and its importance**: ROC Curve is simpler measure of confusion matrices. Receiver Operator Characteristics curve is simple way to summarize. The vertical axis of the graph is sensitivity (TPR) and x axis is (1-specificity) FPR with varying thresholds. Threshold is a kind of decision boundary which enables the samples to classify into their respective classes. With increasing threshold you allow more samples of class X. The importance of ROC curve is that, Classifiers that give curves closer to the top-left corner indicate a better performance. We envision higher TPR because we want more predictions to match with ground truth. ROC curve is suitable for balanced datasets. Because ROC curve may give you wrong hopes of model being accurate.

An **ROC curve** (**receiver operating characteristic curve**) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters:

- True Positive Rate
- False Positive Rate

True Positive Rate (**TPR**) is a synonym for recall and is therefore defined as follows:

$$TPR = \frac{TP}{TP + FN}$$

False Positive Rate (**FPR**) is defined as follows:

$$FPR = \frac{FP}{FP + TN}$$

4. **Precision Recall and curve importance**: Suitable for imbalanced data. In more simpler words, **Precision** is the number of times the class A got correctly classified divided by number of times the model assumed it was class A. However (TPR) **Recall** is number of times model classified it was class A divided by the number of times it was actually class A.

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

5. PR Curve is a plot of Precision (vertical axis) and Recall (horizontal axis) measures at different thresholds. Because of this we can rely on PR Curve for imbalanced dataset. Even if the dataset is not balanced or the dataset is biased, we will not deal with different proportions of the data. We will check how our model performed within that particular class. As precision checks for correctness of the classification to the assumption for that class of classifier whereas recall checks for correctness of the classification to actual ground truth of that class of classifier. In general PR Curve overlaps for balanced and imbalanced datasets. However, the PR curve drop point tells the proportion of the classes you incorporated. In simpler terms, PR Curve tells about the biasness by the droppage of the curve to some value very high or very low according to the class your envision. However for balanced dataset it drops at 0.5. Additionally, we can say that when recall = 1, we shall check the precision such that we get to know about balanced and imbalanced datasets.
6. Extra Theory for biasness: Whereas it is not a possible way to detect biasness in ROC Curves. Because of the biased dataset you will see a beautiful and perfect ROC Curve such that the curve for balanced dataset will be lower top left than to the curve of imbalanced dataset which will be higher top left. Thus we would first look at the recall values, then precision values. We are more cantered towards the actual values.

Actual	Predicted	
0	0	TN
0	1	FP
1	0	FN
1	1	TP

Solution 3:

Input Sentence 1: `Giraffen an einem sonnigen Tag im Busch.`

Output Sentence 1: `English`
`['the', '<unk>', 'on', 'a', 'sunny', 'day', 'in', 'the', '<unk>', '.', '<eos>']`

Input Sentence 2: `Alle die wandern sind nicht verloren`
`# English Original: All those who wander are not lost`

Output Sentence 3: `English`
`['all', 'the', 'back', 'of', 'the', 'skate', '.', '<eos>']`

BLEU SCORE: `0.3256714642047882`

Solution 4:

Images:

EPOCH (2):

```
Example 1 CORRECT: Man in red jacket

Example 1 OUTPUT: <SOS> a man in a red shirt is standing on a <UNK> . <EOS>

Example 2 CORRECT: Elephent in grass

Example 2 OUTPUT: <SOS> a man in a red shirt is standing on a <UNK> . <EOS>

Example 3 CORRECT: bears playing on stones

Example 3 OUTPUT: <SOS> a man in a red shirt is standing on a <UNK> . <EOS>

Example 4 CORRECT: A parked bus

Example 4 OUTPUT: <SOS> a man in a red shirt is standing on a <UNK> . <EOS>

Example 5 CORRECT: A red car on road

Example 5 OUTPUT: <SOS> a man in a red shirt is standing on a <UNK> . <EOS>
```

EPOCH (18):

```
Example 1 CORRECT: Man in red jacket

Example 1 OUTPUT: <SOS> a man in a red shirt is standing on a bench . <EOS>

Example 2 CORRECT: Elephent in grass

Example 2 OUTPUT: <SOS> a man in a red shirt is standing on a bench . <EOS>

Example 3 CORRECT: bears playing on stones

Example 3 OUTPUT: <SOS> a dog is running through the grass . <EOS>

Example 4 CORRECT: A parked bus

Example 4 OUTPUT: <SOS> a man in a red shirt is standing on a bench . <EOS>

Example 5 CORRECT: A red car on road

Example 5 OUTPUT: <SOS> a man in a red shirt is standing on a rock wall . <EOS>
```

EPOCH (25):

```
Example 1 CORRECT: Man in red jacket

Example 1 OUTPUT: <SOS> a man in a red jacket and a backpack stands on a rock .
<EOS>

Example 2 CORRECT: Elephent in grass

Example 2 OUTPUT: <SOS> a man in a red shirt and a woman in a white shirt and
jeans stands in front of a brick building . <EOS>



Example 3 CORRECT: bears playing on stones

Example 3 OUTPUT: <SOS> a black dog is running through the grass . <EOS>

Example 4 CORRECT: A parked bus

Example 4 OUTPUT: <SOS> a man in a red shirt and a woman in a white shirt and
jeans is standing in front of a brick building . <EOS>

Example 5 CORRECT: A red car on road

Example 5 OUTPUT: <SOS> a man in a red shirt and a white shirt is riding a bicycle
. <EOS>
```

EPOCH (50) - (Ran approx. 6 hours):

```
Example 1 CORRECT: Man in red jacket

Example 1 OUTPUT: <SOS> a man in a red jacket and a backpack stands on a beach.
<EOS>

Example 2 CORRECT: Elephent in grass

Example 2 OUTPUT: <SOS> a dog is standing on grass . <EOS>

Example 3 CORRECT: bears playing on stones

Example 3 OUTPUT: <SOS> a black dog is running through the grass . <EOS>

Example 4 CORRECT: A parked bus

Example 4 OUTPUT: <SOS> a bus on road. <EOS>

Example 5 CORRECT: A red car on road

Example 5 OUTPUT: <SOS> a man in a red shirt and a white shirt is riding a bicycle
. <EOS>
```

References:

1. https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html
2. https://www.youtube.com/channel/UCkzW5JSFwvKRjXABI-UTAkQ
3. https://acutecaretesting.org/en/articles/precision-recall-curves-what-are-they-and-how-are-they-used
4. http://www.davidsbatista.net/blog/2018/08/19/NLP_Metrics/
5. https://github.com/aladdinpersson/Machine-Learning-Collection/tree/master/ML/Pytorch/more_advanced
6. https://github.com/gucci-j/imdb-classification-gru