Assignment 4: ELMO

Srijan Chakraborty
2020115001

Introduction:

The advancements in natural language processing (NLP) have been significant in recent years due to the development of powerful deep learning models. One such model is ELMo, which can capture complex language structures and contextual information, improving performance on various NLP tasks. The key strength of ELMo is its ability to create contextualized word embeddings, providing a more detailed representation of words in a sentence. This project aims to create an ELMo model from scratch and assess its performance on two benchmark datasets: Multi-NLI and SST.

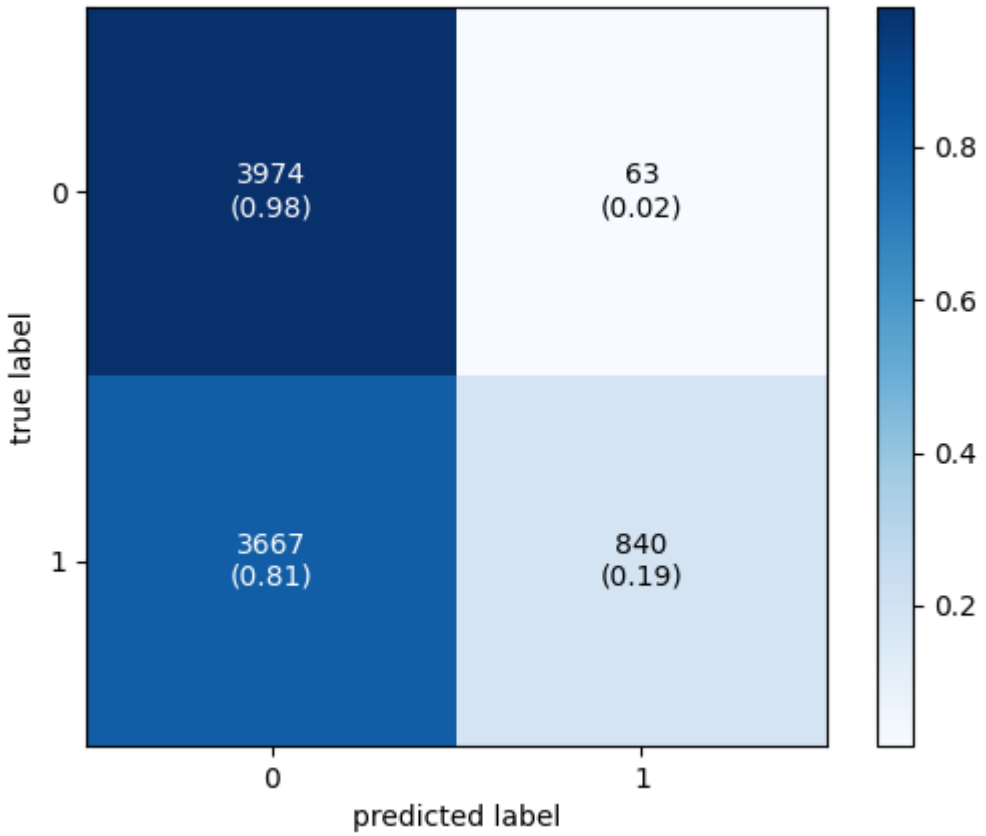
Our ELMo model starts with pre-trained GloVe embeddings to understand word semantics and then uses a two-layer LSTM architecture to capture contextual information. The final embedding is a combination of the original GloVe embeddings and the LSTM hidden states, blending semantic and contextual information.

By evaluating the ELMo model on the Multi-NLI (a collection of sentence pairs labelled for entailment, contradiction, and neutral relationships) and SST (focused on sentiment analysis of movie reviews) datasets, we aim to demonstrate its potential for enhancing natural language understanding.

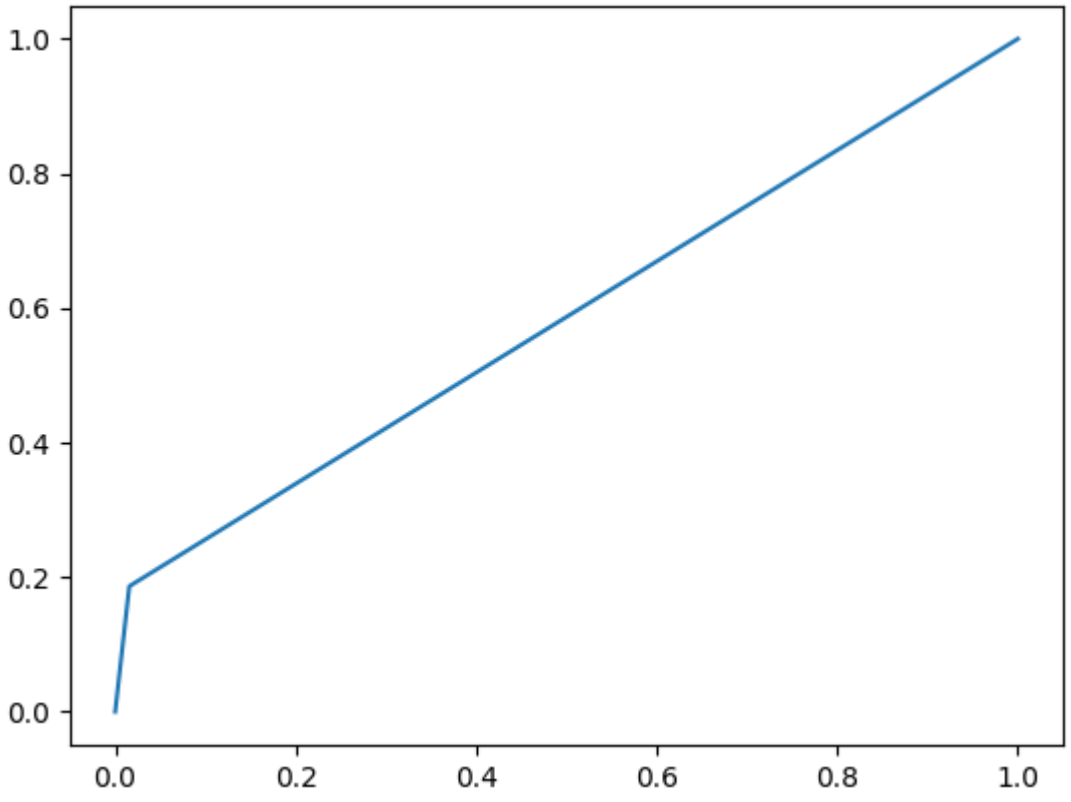
Analysis:

For the SST Dataset:

```
Classification Report:
              precision    recall  f1-score   support

           0       0.52      0.98      0.68      4037
           1       0.93      0.19      0.31      4507

    accuracy                           0.56      8544
   macro avg       0.73      0.59      0.50      8544
weighted avg       0.74      0.56      0.49      8544


Confusion Matrix:
[[3974   63]
 [3667  840]]
AUC:  0.5853855497618847
Done
```

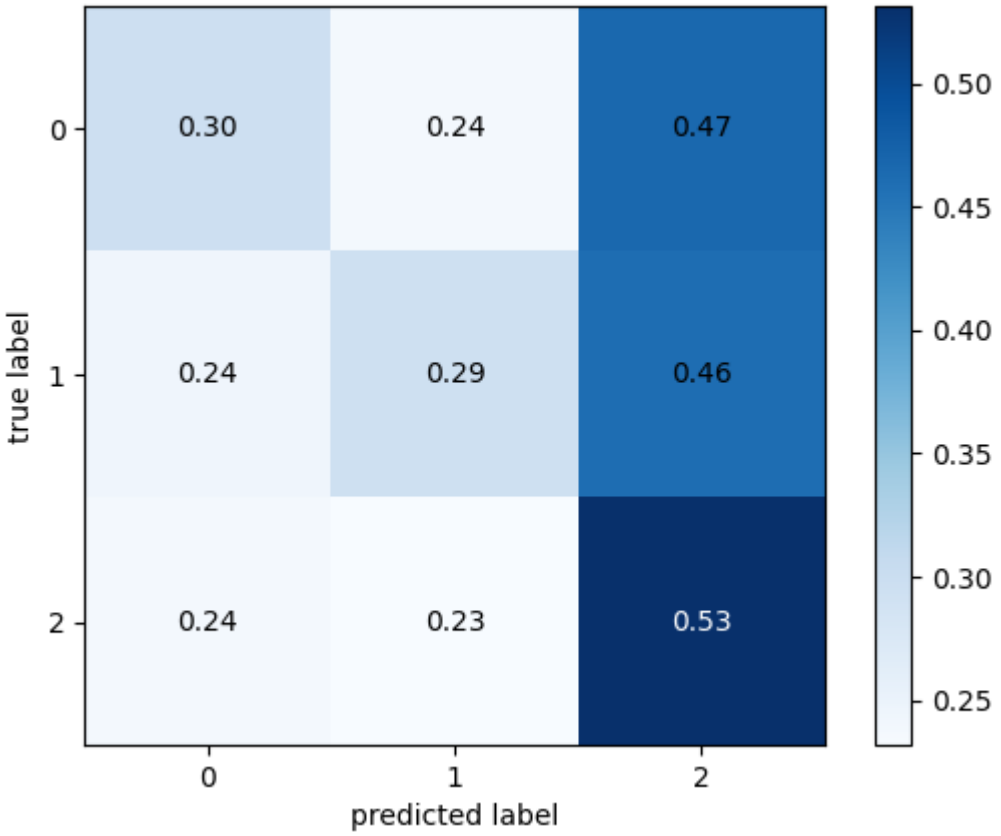

3974
(0.98)
63
(0.02)
3667
(0.81)
840
(0.19)
0
1
true label
0
1
predicted label
0.8
0.6
0.4
0.2

ROC Curve:


1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

For the NLI Dataset:

```
Classification Report:
              precision    recall  f1-score   support

           0       0.38      0.34      0.36     13525
           1       0.35      0.33      0.34     12122
           2       0.40      0.45      0.42     14353

    accuracy                           0.38     40000
   macro avg       0.37      0.37      0.37     40000
weighted avg       0.38      0.38      0.38     40000


Confusion Matrix:
[[4621 3735 5169]
 [3463 4051 4608]
 [4018 3934 6401]]
```



Due to a time constraint and the difficulty of multi-class plotting of the ROC curve, this has been skipped.

It should be noted that the performance of our ELMo model is a reflection of its ability to understand word semantics and contextual information, which is particularly advantageous for NLP tasks. Although the results may not be at the forefront of the field, they still serve as a starting point for further improvements and optimizations.

One major factor to consider when interpreting these results is the limited computational resources and training time that were available for our project. Advanced NLP models often require extensive training periods and vast amounts of computing power to achieve top-tier performance. In our case, these constraints may have restricted the model's potential, and with additional resources and training time, we could potentially achieve improved results.

Possible avenues for future research include fine-tuning the model architecture, adjusting hyperparameters, or incorporating additional training data to enhance the model's performance. Additionally, allocating more compute resources and extending training time could also lead to further improvements.

In conclusion, our ELMo model has demonstrated its potential for contributing to natural language understanding by capturing contextual information in word embeddings. Despite the limitations in computational resources and training time, the results obtained on the SST and Multi-NLI datasets provide valuable insights into the model's performance and pave the way for future research and improvements in the field of NLP.