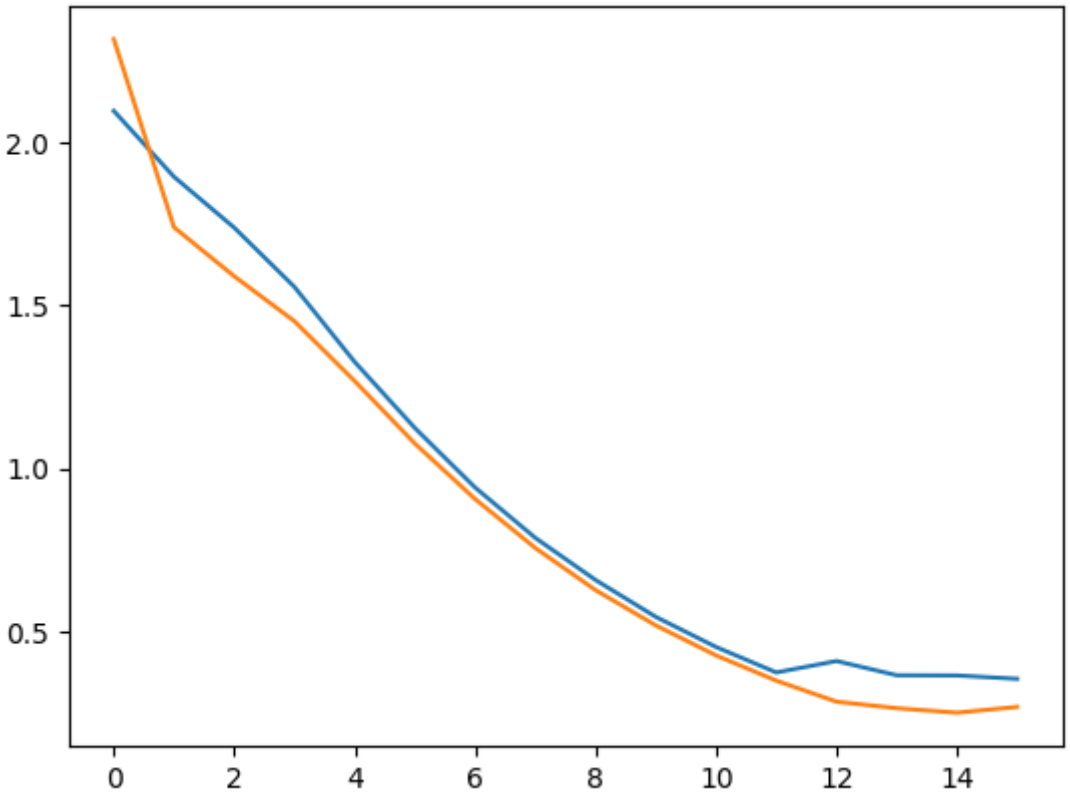
ELMO

The code is written in pytorch using pytorch lightning,

```
.
├── Assignment4.pdf
├── data
│   ├── sst_stoi.pkl
│   ├── sst_test.pkl
│   ├── sst_train.pkl
│   └── sst_validation.pkl
├── elmo
│   ├── configs
│   ├── data
│   │   ├── nli.py # nli dataset loader
│   │   └── sst.py # sst dataset loader
│   ├── __init__.py
│   ├── models
│   │   ├── components
│   │   │   ├── elmo.py # elmo embeddi
│   │   │   └── rnn.py # the lstm/rnn model that uses emo embedding
│   │   ├── elmo.py # the lightning datamodule for elmo
│   │   └── rnn.py # the lightning datamoduel for rnn
│   ├── pretrain.py
│   └── train.py
└── elmo.py
```

SST Dataset

Training ELMO

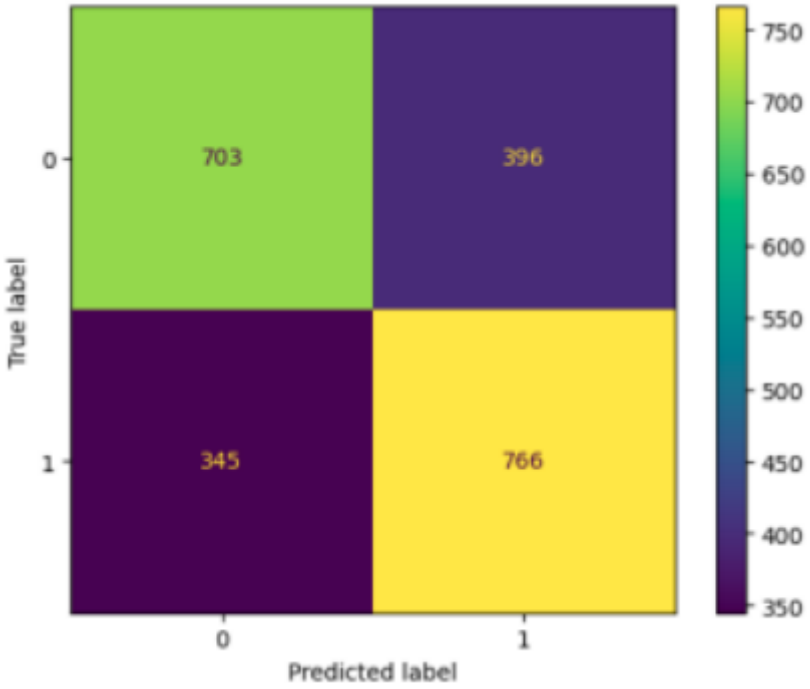


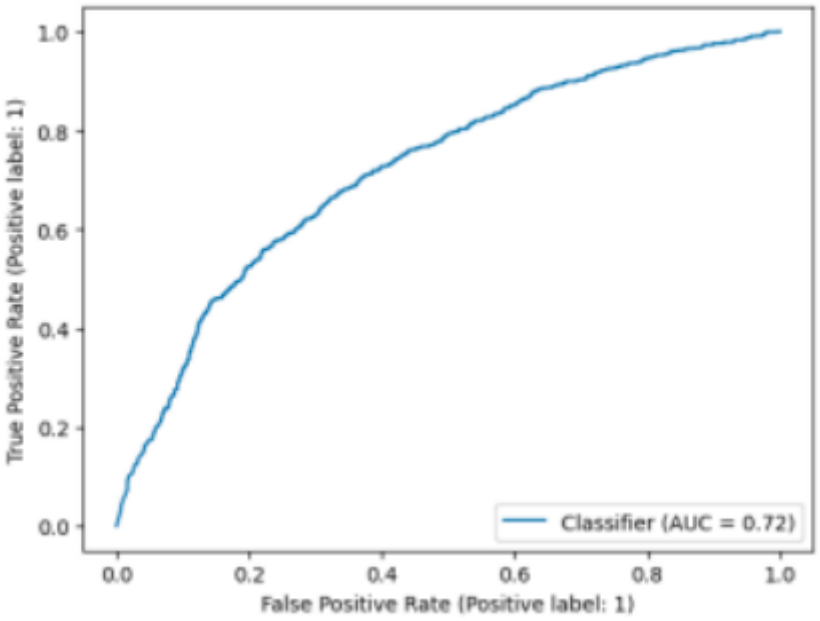
Hyperparameters(used pretrained glove embeddings)

```
BATCH_SIZE = 32
EPOCHS = 15
charcnn:
    char_embedding_dim: int = 16,
    kernel_sizes: list = [1, 2, 3, 4, 5, 6, 7],
    layer_sizes: list = [8, 8, 16, 32, 64, 128, 256],  # 512
elmo:
     char_embedding_dim: int = 16,
     input_size: int = 512,
     hidden_size: int = 256,
     dropout: float = 0.5,
     num_highway_layers: int = 2,
```

Accuracy

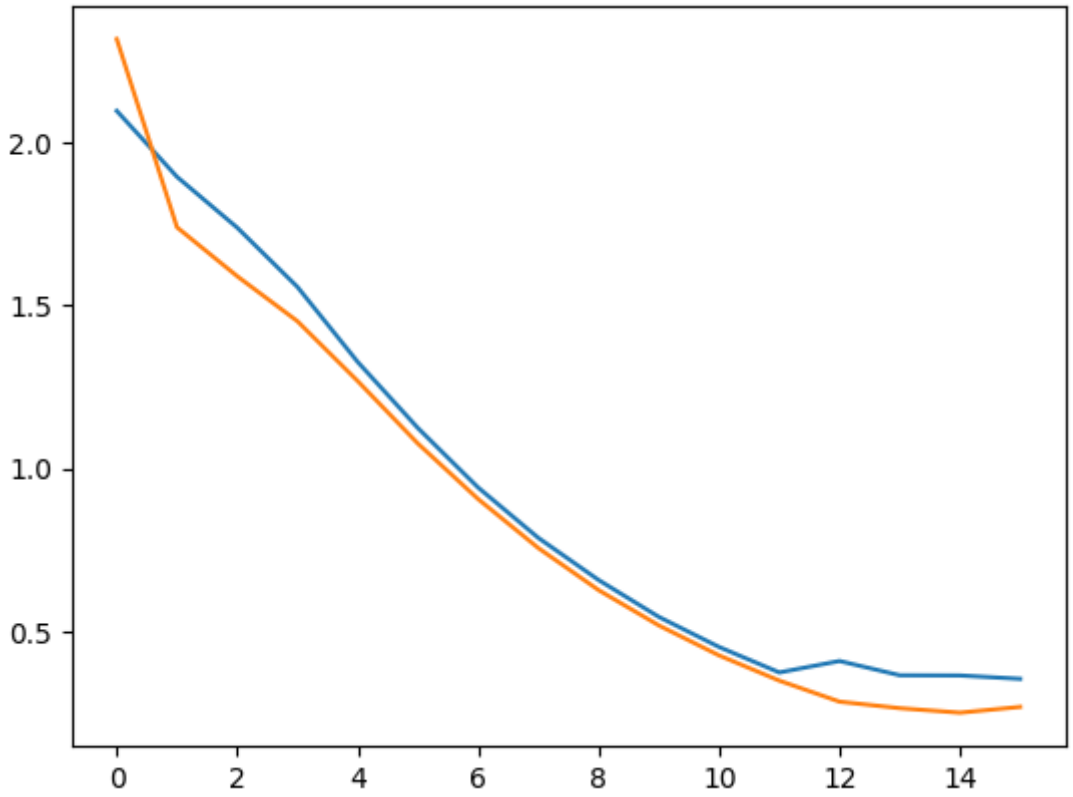
```
              precision    recall  f1-score   support

           1       0.69      0.64      0.66      1099
           2       0.66      0.68      0.67      1111

    accuracy                           0.69      2280
   macro avg       0.68      0.66      0.67      2280
weighted avg       0.69      0.69      0.69      2280
```





NLI Dataset

Training ELMO

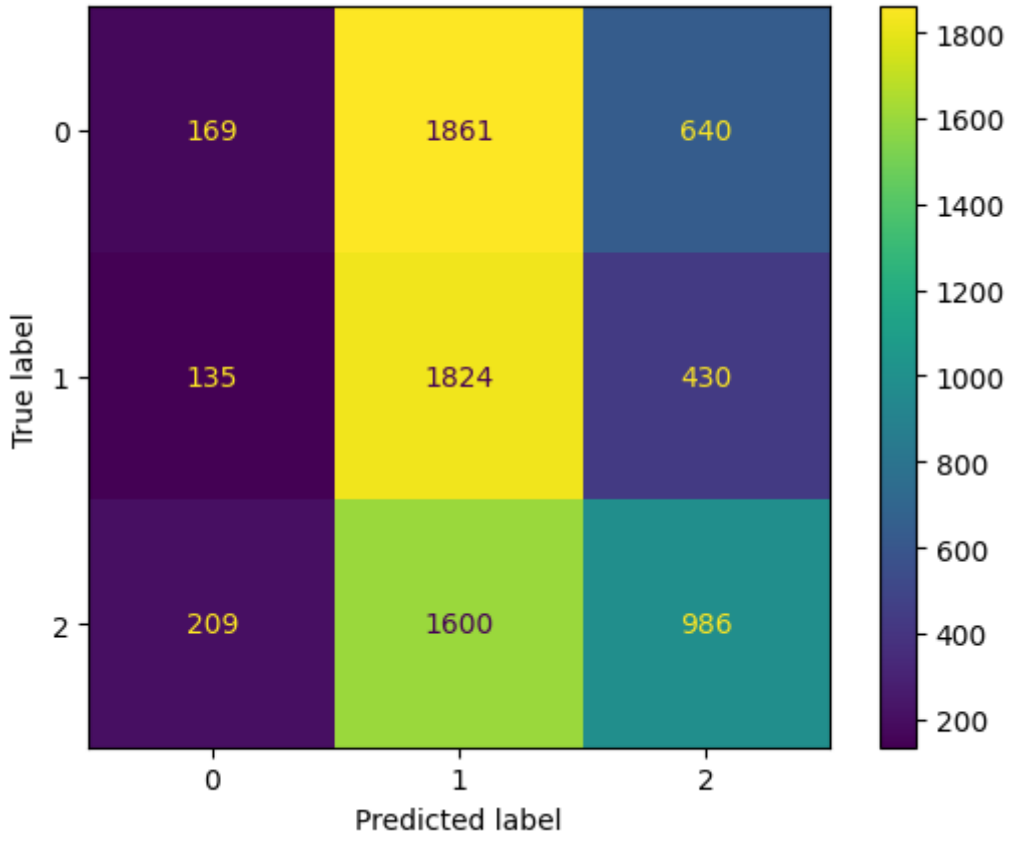


Hyperparameters(used pretrained glove embeddings)

```
BATCH_SIZE = 32
EPOCHS = 15
charcnn:
    char_embedding_dim: int = 16,
    kernel_sizes: list = [1, 2, 3, 4, 5, 6, 7],
    layer_sizes: list = [8, 8, 16, 32, 64, 128, 256],  # 512
elmo:
     char_embedding_dim: int = 16,
     input_size: int = 512,
     hidden_size: int = 256,
     dropout: float = 0.5,
     num_highway_layers: int = 2,
```

Accuracy

```
              precision    recall  f1-score   support

           0       0.43      0.16      0.21      2670
           1       0.45      0.76      0.48      2389
           2       0.48      0.45      0.41      2795

    accuracy                           0.48      7854
   macro avg       0.48      0.49      0.43      7854
weighted avg       0.49      0.48      0.43      7854
```


169
1861
640
135
1824
430
209
1600
986
0
1
2
True label
0
1
2
Predicted label
1800
1600
1400
1200
1000
800
600
400
200