NLP Assignment 4 Report

Word2Vec Training (for pretrained embeddings):

SST:

```
# Model Compilation, Parameters and Metrics

sst_model.compile(loss      =  'binary_crossentropy',
                  # optimizer =  'adam',
                  optimizer = tf.keras.optimizers.Adam(),
                  metrics   =  ['accuracy'])

sst_word2vec_training = sst_model.fit(sst_x, sst_y, epochs=5, batch_size=256)
sst_word2vec_embeddings = sst_model.dense1.weights[0]
```

```
Epoch 1/5
260/260 [==============================] - 3s 4ms/step - loss: 0.4586 - accuracy: 0.8339
Epoch 2/5
260/260 [==============================] - 1s 4ms/step - loss: 0.0352 - accuracy: 0.9980
Epoch 3/5
260/260 [==============================] - 1s 4ms/step - loss: 0.0110 - accuracy: 0.9981
Epoch 4/5
260/260 [==============================] - 1s 4ms/step - loss: 0.0084 - accuracy: 0.9982
Epoch 5/5
260/260 [==============================] - 1s 5ms/step - loss: 0.0076 - accuracy: 0.9982
```

Multi NLI:

```
multi_nli_model = MyModel(128)
# Model Compilation, Parameters and Metrics

multi_nli_model.compile(loss      =  'binary_crossentropy',
                  # optimizer =  'adam',
                  optimizer = tf.keras.optimizers.Adam(),
                  metrics   =  ['accuracy'])

multi_nli_word2vec_training = multi_nli_model.fit(multi_nli_x, multi_nli_y, epochs=10, batch_size=256)
multi_nli_word2vec_embeddings = multi_nli_model.dense1.weights[0]
```

```
Epoch 1/10
981/981 [==============================] - 5s 4ms/step - loss: 0.5578 - accuracy: 0.7340
Epoch 2/10
981/981 [==============================] - 5s 5ms/step - loss: 0.4974 - accuracy: 0.7496
Epoch 3/10
981/981 [==============================] - 4s 4ms/step - loss: 0.4918 - accuracy: 0.7482
Epoch 4/10
981/981 [==============================] - 5s 5ms/step - loss: 0.4905 - accuracy: 0.7487
Epoch 5/10
981/981 [==============================] - 4s 4ms/step - loss: 0.4902 - accuracy: 0.7473
Epoch 6/10
981/981 [==============================] - 4s 4ms/step - loss: 0.4900 - accuracy: 0.7479
Epoch 7/10
981/981 [==============================] - 5s 5ms/step - loss: 0.4898 - accuracy: 0.7482
Epoch 8/10
981/981 [==============================] - 4s 4ms/step - loss: 0.4898 - accuracy: 0.7468
Epoch 9/10
981/981 [==============================] - 5s 5ms/step - loss: 0.4896 - accuracy: 0.7476
Epoch 10/10
981/981 [==============================] - 4s 4ms/step - loss: 0.4896 - accuracy: 0.7473
```

Elmo Model Training:

SST

```
elmo_model.compile(loss      =  'sparse_categorical_crossentropy',
                   # optimizer =  'adam',
                   optimizer = tf.keras.optimizers.Adam(),
                   metrics   =  ['accuracy'])

sst_elmo_training = elmo_model.fit(np.array(x_train), np.array(y_train), epochs=15, batch_size=256, validation_data=(x_valid,y_valid))
```

```
Epoch 1/15
34/34 [==============================] - 19s 296ms/step - loss: 5.3667 - accuracy: 0.6530 - val_loss: 2.4873 - val_accuracy: 0.6716
Epoch 2/15
34/34 [==============================] - 6s 187ms/step - loss: 2.4795 - accuracy: 0.6758 - val_loss: 2.3161 - val_accuracy: 0.6716
Epoch 3/15
34/34 [==============================] - 6s 171ms/step - loss: 2.2401 - accuracy: 0.6990 - val_loss: 2.0762 - val_accuracy: 0.7060
Epoch 4/15
34/34 [==============================] - 6s 160ms/step - loss: 2.0345 - accuracy: 0.7402 - val_loss: 1.9188 - val_accuracy: 0.7762
Epoch 5/15
34/34 [==============================] - 3s 85ms/step - loss: 1.9130 - accuracy: 0.7767 - val_loss: 1.8270 - val_accuracy: 0.7859
Epoch 6/15
34/34 [==============================] - 4s 134ms/step - loss: 1.8432 - accuracy: 0.7822 - val_loss: 1.7727 - val_accuracy: 0.7925
Epoch 7/15
34/34 [==============================] - 3s 84ms/step - loss: 1.7959 - accuracy: 0.7866 - val_loss: 1.7276 - val_accuracy: 0.7924
Epoch 8/15
34/34 [==============================] - 3s 78ms/step - loss: 1.7527 - accuracy: 0.7867 - val_loss: 1.6868 - val_accuracy: 0.7925
Epoch 9/15
34/34 [==============================] - 2s 60ms/step - loss: 1.7121 - accuracy: 0.7867 - val_loss: 1.6476 - val_accuracy: 0.7925
Epoch 10/15
34/34 [==============================] - 2s 64ms/step - loss: 1.6723 - accuracy: 0.7867 - val_loss: 1.6063 - val_accuracy: 0.7925
Epoch 11/15
34/34 [==============================] - 2s 61ms/step - loss: 1.6363 - accuracy: 0.7869 - val_loss: 1.5747 - val_accuracy: 0.7929
Epoch 12/15
34/34 [==============================] - 2s 70ms/step - loss: 1.6085 - accuracy: 0.7874 - val_loss: 1.5501 - val_accuracy: 0.7936
Epoch 13/15
34/34 [==============================] - 2s 59ms/step - loss: 1.5833 - accuracy: 0.7878 - val_loss: 1.5269 - val_accuracy: 0.7941
Epoch 14/15
34/34 [==============================] - 2s 65ms/step - loss: 1.5578 - accuracy: 0.7885 - val_loss: 1.5034 - val_accuracy: 0.7942
Epoch 15/15
34/34 [==============================] - 2s 61ms/step - loss: 1.5306 - accuracy: 0.7890 - val_loss: 1.4769 - val_accuracy: 0.7950
```

Multi NLI

```
v_size = len(multi_nli_int_to_vocab)
multi_nli_elmo_model = ELMO_Model(multi_nli_word2vec_embeddings, v_size)
multi_nli_elmo_model.compile(loss      =  'sparse_categorical_crossentropy',
                   # optimizer =  'adam',
                   optimizer = tf.keras.optimizers.Adam(),
                   metrics   =  ['accuracy'])

multi_nli_elmo_training = multi_nli_elmo_model.fit(tf.constant(x_train), tf.constant(y_train_model), epochs=5, batch_size=256, validatio
```

```
Epoch 1/5
246/246 [==============================] - 166s 643ms/step - loss: 1.2232 - accuracy: 0.9420 - val_loss: 0.5437 - val_accuracy: 0.9373
Epoch 2/5
246/246 [==============================] - 143s 579ms/step - loss: 0.4143 - accuracy: 0.9525 - val_loss: 0.5196 - val_accuracy: 0.9396
Epoch 3/5
246/246 [==============================] - 134s 543ms/step - loss: 0.3992 - accuracy: 0.9534 - val_loss: 0.4970 - val_accuracy: 0.9405
Epoch 4/5
246/246 [==============================] - 135s 551ms/step - loss: 0.3805 - accuracy: 0.9537 - val_loss: 0.4748 - val_accuracy: 0.9409
Epoch 5/5
246/246 [==============================] - 130s 527ms/step - loss: 0.3650 - accuracy: 0.9541 - val_loss: 0.4551 - val_accuracy: 0.9415
```

ELMO Classifier training:

SST:

```
elmo_classifier.compile(loss      =  'binary_crossentropy',
                optimizer = tf.keras.optimizers.Adam(),
                metrics   =  ['accuracy'])
sst_elmo_training = elmo_classifier.fit(tf.constant(x_train), tf.constant(y_train_classifier), epochs=15, batch_size=256, validation_da
```

```
Epoch 1/15
Tensor("elmo__classifier/concat:0", shape=(3,), dtype=int32)
WARNING:tensorflow:Gradients do not exist for variables ['elmo__model/time_distributed/kernel:0', 'elmo__model/time_distributed/bias:0'
WARNING:tensorflow:Gradients do not exist for variables ['elmo__model/time_distributed/kernel:0', 'elmo__model/time_distributed/bias:0'
Tensor("elmo__classifier/concat:0", shape=(3,), dtype=int32)
WARNING:tensorflow:Gradients do not exist for variables ['elmo__model/time_distributed/kernel:0', 'elmo__model/time_distributed/bias:0'
WARNING:tensorflow:Gradients do not exist for variables ['elmo__model/time_distributed/kernel:0', 'elmo__model/time_distributed/bias:0'
34/34 [==============================] - ETA: 0s - loss: 0.6249 - accuracy: 0.6555Tensor("elmo__classifier/concat:0", shape=(3,), dtype
34/34 [==============================] - 15s 194ms/step - loss: 0.6249 - accuracy: 0.6555 - val_loss: 0.6993 - val_accuracy: 0.5967
Epoch 2/15
34/34 [==============================] - 5s 137ms/step - loss: 0.5182 - accuracy: 0.7498 - val_loss: 0.8351 - val_accuracy: 0.5241
Epoch 3/15
34/34 [==============================] - 4s 118ms/step - loss: 0.4235 - accuracy: 0.8136 - val_loss: 0.9703 - val_accuracy: 0.4723
Epoch 4/15
34/34 [==============================] - 3s 72ms/step - loss: 0.3633 - accuracy: 0.8509 - val_loss: 1.0240 - val_accuracy: 0.5114
Epoch 5/15
34/34 [==============================] - 2s 70ms/step - loss: 0.3157 - accuracy: 0.8771 - val_loss: 1.1146 - val_accuracy: 0.5186
Epoch 6/15
34/34 [==============================] - 2s 48ms/step - loss: 0.2766 - accuracy: 0.8982 - val_loss: 1.1685 - val_accuracy: 0.5050
Epoch 7/15
34/34 [==============================] - 2s 49ms/step - loss: 0.2559 - accuracy: 0.9106 - val_loss: 1.4061 - val_accuracy: 0.5041
Epoch 8/15
34/34 [==============================] - 3s 82ms/step - loss: 0.2344 - accuracy: 0.9175 - val_loss: 1.3433 - val_accuracy: 0.4877
Epoch 9/15
34/34 [==============================] - 2s 68ms/step - loss: 0.2112 - accuracy: 0.9318 - val_loss: 1.3884 - val_accuracy: 0.5186
Epoch 10/15
34/34 [==============================] - 1s 30ms/step - loss: 0.1981 - accuracy: 0.9355 - val_loss: 1.5369 - val_accuracy: 0.5204
Epoch 11/15
34/34 [==============================] - 1s 43ms/step - loss: 0.1715 - accuracy: 0.9479 - val_loss: 1.6022 - val_accuracy: 0.5150
Epoch 12/15
34/34 [==============================] - 1s 21ms/step - loss: 0.1562 - accuracy: 0.9542 - val_loss: 1.5688 - val_accuracy: 0.5241
Epoch 13/15
34/34 [==============================] - 1s 29ms/step - loss: 0.1395 - accuracy: 0.9600 - val_loss: 1.8308 - val_accuracy: 0.4850
Epoch 14/15
34/34 [==============================] - 1s 31ms/step - loss: 0.1372 - accuracy: 0.9589 - val_loss: 1.8029 - val_accuracy: 0.4886
Epoch 15/15
34/34 [==============================] - 1s 27ms/step - loss: 0.1319 - accuracy: 0.9614 - val_loss: 1.7591 - val_accuracy: 0.5241
```

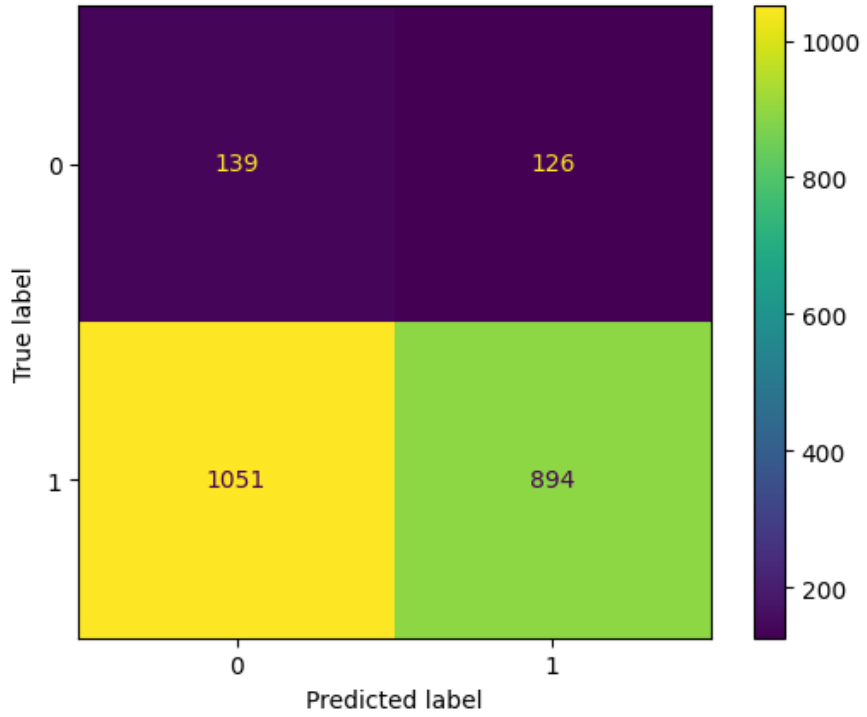
Multi NLI:

```
multi_nli_elmo_classifier.compile(loss      =  'sparse_categorical_crossentropy',
                optimizer = tf.keras.optimizers.Adam(),
                metrics   =  ['accuracy'])
multi_nli_elmo_training = multi_nli_elmo_classifier.fit([x_train_premise, x_train_hypothesis], tf.constant(y_train_classifier), epochs=1
```

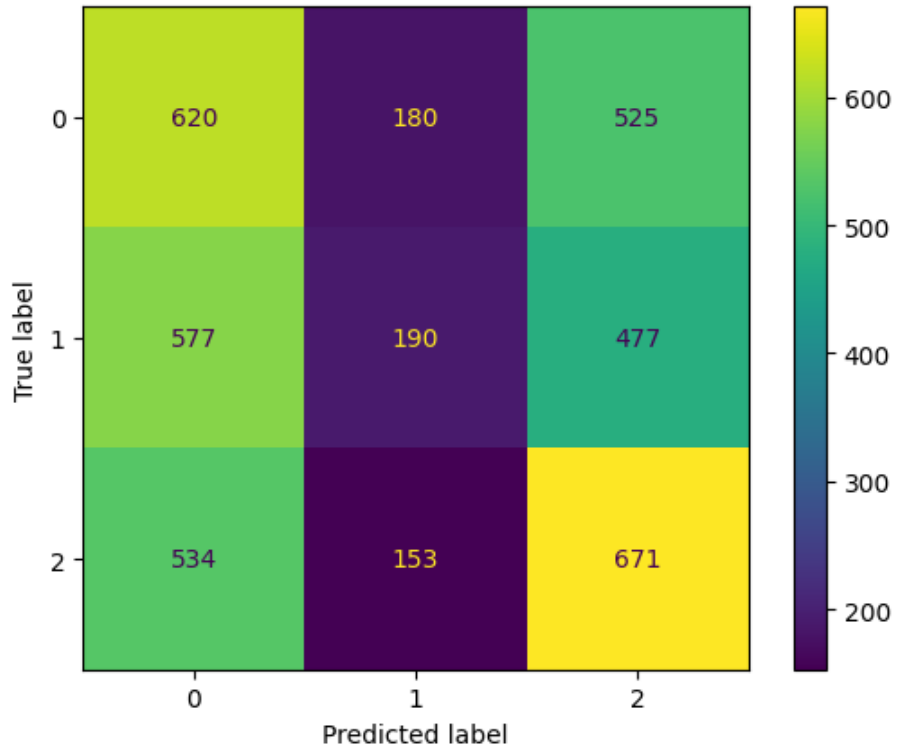
```
Epoch 1/15
Tensor("elmo__classifier_multi_nli/concat:0", shape=(3,), dtype=int32)
(None, 128, 3)
Tensor("elmo__classifier_multi_nli/concat_2:0", shape=(3,), dtype=int32)
WARNING:tensorflow:Gradients do not exist for variables ['elmo__model_1/time_distributed_1/kernel:0', 'elmo__model_1/time_distributed_1/b
WARNING:tensorflow:Gradients do not exist for variables ['elmo__model_1/time_distributed_1/kernel:0', 'elmo__model_1/time_distributed_1/b
Tensor("elmo__classifier_multi_nli/concat:0", shape=(3,), dtype=int32)
(None, 128, 3)
Tensor("elmo__classifier_multi_nli/concat_2:0", shape=(3,), dtype=int32)
WARNING:tensorflow:Gradients do not exist for variables ['elmo__model_1/time_distributed_1/kernel:0', 'elmo__model_1/time_distributed_1/b
WARNING:tensorflow:Gradients do not exist for variables ['elmo__model_1/time_distributed_1/kernel:0', 'elmo__model_1/time_distributed_1/b
123/123 [==============================] - ETA: 0s - loss: 1.0963 - accuracy: 0.3570Tensor("elmo__classifier_multi_nli/concat:0", shape=(
(None, 128, 3)
Tensor("elmo__classifier_multi_nli/concat_2:0", shape=(3,), dtype=int32)
123/123 [==============================] - 46s 263ms/step - loss: 1.0963 - accuracy: 0.3570 - val_loss: 1.0996 - val_accuracy: 0.3400
Epoch 2/15
123/123 [==============================] - 21s 174ms/step - loss: 1.0951 - accuracy: 0.3586 - val_loss: 1.0986 - val_accuracy: 0.3275
Epoch 3/15
123/123 [==============================] - 17s 142ms/step - loss: 1.0927 - accuracy: 0.3678 - val_loss: 1.0994 - val_accuracy: 0.3547
Epoch 4/15
123/123 [==============================] - 16s 125ms/step - loss: 1.0881 - accuracy: 0.3895 - val_loss: 1.0980 - val_accuracy: 0.3606
Epoch 5/15
123/123 [==============================] - 13s 106ms/step - loss: 1.0832 - accuracy: 0.3973 - val_loss: 1.0944 - val_accuracy: 0.3700
Epoch 6/15
123/123 [==============================] - 13s 104ms/step - loss: 1.0725 - accuracy: 0.4181 - val_loss: 1.1013 - val_accuracy: 0.3690
Epoch 7/15
123/123 [==============================] - 11s 86ms/step - loss: 1.0673 - accuracy: 0.4290 - val_loss: 1.1114 - val_accuracy: 0.3491
Epoch 8/15
123/123 [==============================] - 11s 90ms/step - loss: 1.0673 - accuracy: 0.4325 - val_loss: 1.0988 - val_accuracy: 0.3695
Epoch 9/15
123/123 [==============================] - 11s 93ms/step - loss: 1.0579 - accuracy: 0.4438 - val_loss: 1.1126 - val_accuracy: 0.3802
Epoch 10/15
123/123 [==============================] - 10s 85ms/step - loss: 1.0448 - accuracy: 0.4612 - val_loss: 1.1026 - val_accuracy: 0.3888
Epoch 11/15
123/123 [==============================] - 11s 90ms/step - loss: 1.0332 - accuracy: 0.4738 - val_loss: 1.1139 - val_accuracy: 0.3901
Epoch 12/15
123/123 [==============================] - 11s 89ms/step - loss: 1.0255 - accuracy: 0.4801 - val_loss: 1.1166 - val_accuracy: 0.3845
Epoch 13/15
123/123 [==============================] - 12s 98ms/step - loss: 1.0212 - accuracy: 0.4870 - val_loss: 1.1190 - val_accuracy: 0.3883
Epoch 14/15
123/123 [==============================] - 11s 91ms/step - loss: 1.0069 - accuracy: 0.5016 - val_loss: 1.1111 - val_accuracy: 0.3886
Epoch 15/15
123/123 [==============================] - 11s 92ms/step - loss: 0.9935 - accuracy: 0.5142 - val_loss: 1.1440 - val_accuracy: 0.3932
```

Confusion Matrix:

SST:



Multi NLI:



ROC Curve (Only for SST):