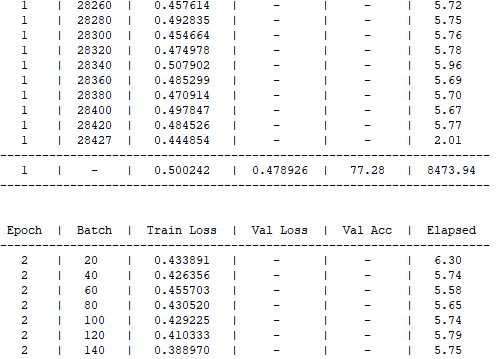
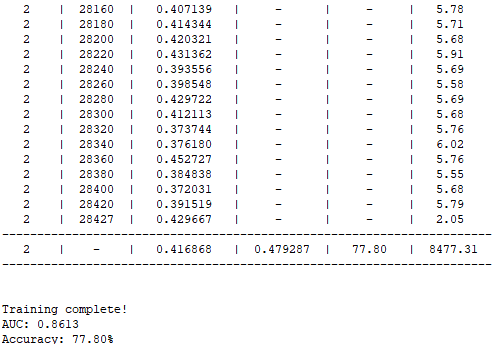
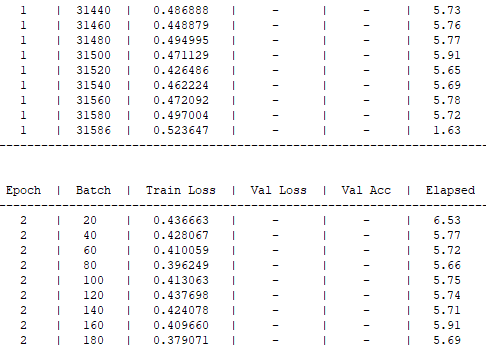
BERT base uncased with 1 Hidden Feed Forward Layer

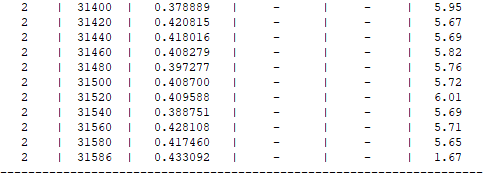
**Training on sarcasm\_train.csv and testing on sarcasm\_valid.csv for 2 epoch, LR = 5e-5:**

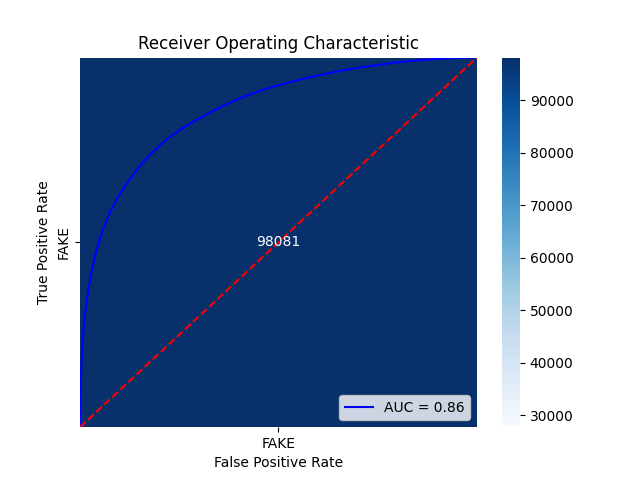
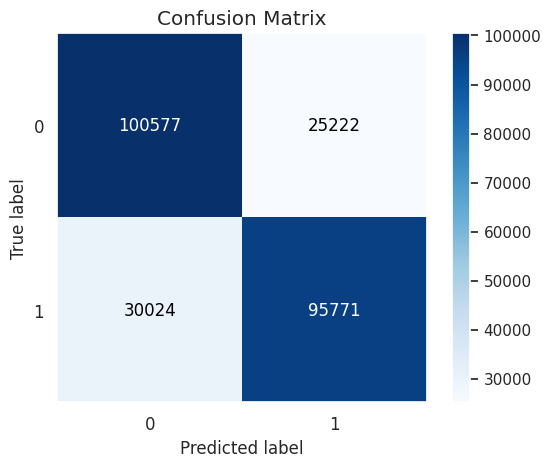
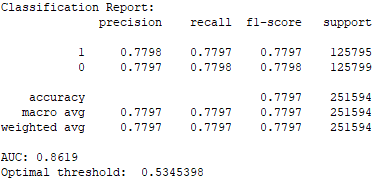


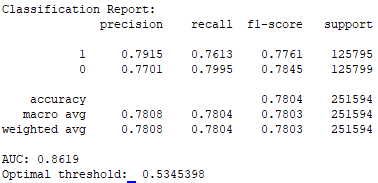


**Training on sarcasm\_train\_valid.csv(combined in the code) and testing on sarcasm\_test.csv for 2 epoch, LR = 5e-5:**

****

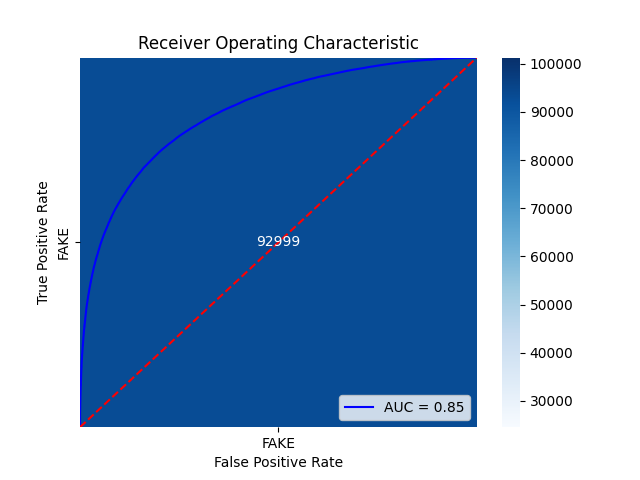
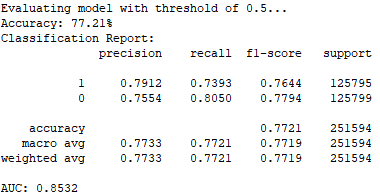
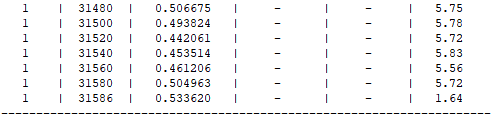
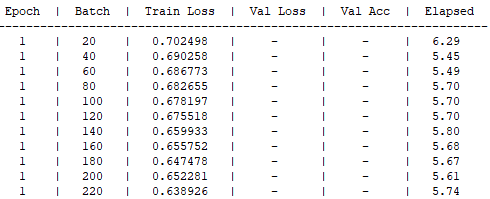
****

**Threshold = 0.5:**

**Threshold = 0.53 ->** 

**Best accuracy** with BERT base uncased with 1 hidden FF layer, 2 epoch, LR = 5e-5, batch size = 32, threshold = 0.53 **=> 78.03%**. **AUC = 0.8619**

Training time: 5hr 15min >.<

**Training on sarcasm\_train\_valid.csv(combined in the code) and testing on sarcasm\_test.csv for 1 epoch, LR = 5e-6:**

BERT without any FF layer

**Training on sarcasm\_train.csv and testing on sarcasm\_valid.csv:**

| Model | Num of Epoch | Learning Rate | Batch Size |  |
| --- | --- | --- | --- | --- |
| BERT base uncased | 1 | 2e-5 | 16 |  |
| BERT base uncased | 1 | 2e-4 | 16 |  |
| BERT base uncased | 1 | 2e-5 | 8 |  |
| BERT base uncased | 1 | 2e-5 | 32 |  |
| BERT base uncased | 1 | 2e-5 | 32 |  |
| BERT base uncased | 2 | 2e-5 | 32 |  |
| BERT base uncased | 1 | 2e-6 | 16 |  |
| BERT base uncased | 3 | 2e-6 | 16 |  |
| BERT **large cased** | 1 | 2e-5 | 16 |  |
| BERT **large uncased** | 1 | 2e-5 | 16 |  |

Best accuracy with BERT base uncased, 3 epoch, LR = 2e-6, batch size = 16 => 76.89%

Might return to this and try more different hyperparameters again.