Movies Reviews Classification Using BERT

ID	NAME
6338	Ahmed Mohamed Ahmed Ali
6672	Mostafa Mohamed tawfiq
6160	Mazen medhat Farid
6670	Mohamed zayton

❖ **Running model (use Bert directly)**

With preprocessing

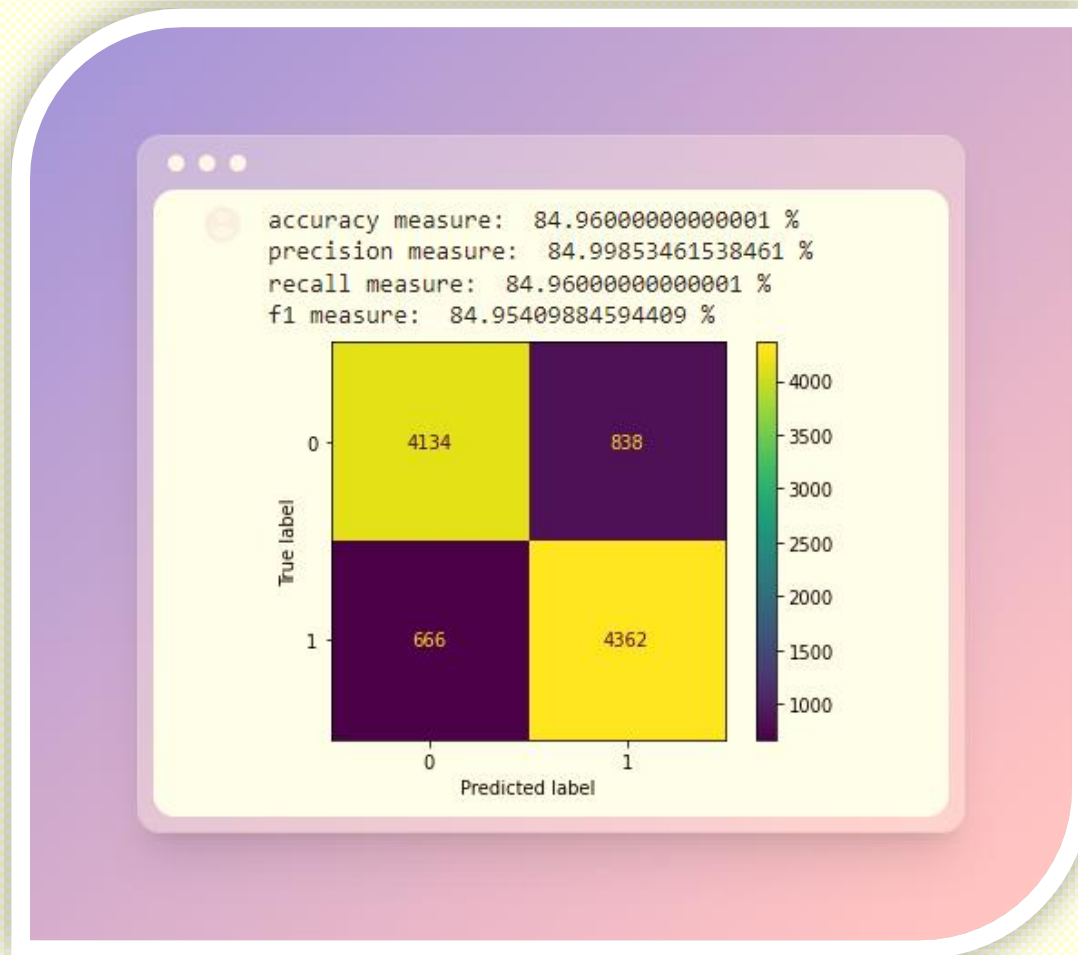**LR=1e-6**

```
model = BertClassifier()
# 512, 256, 128, 64

Some weights of the model checkpoint at bert-base-cased were not used when initializing BertModel: ['cls.pr
- This IS expected if you are initializing BertModel from the checkpoint of a model trained on another task
- This IS NOT expected if you are initializing BertModel from the checkpoint of a model that you expect to
```

```
[ ] run_model(model, train_dataloader, val_dataloader, LR)

100%|████████| 563/563 [17:19<00:00,  1.85s/it]
100%|████████| 63/63 [01:47<00:00,  1.70s/it]
Epochs: 1 | Train Loss:  0.010 | Train Accuracy:  0.583 | Val Loss:  0.008 | Val Accuracy:  0.779
100%|████████| 563/563 [16:49<00:00,  1.79s/it]
100%|████████| 63/63 [01:47<00:00,  1.71s/it]
Epochs: 2 | Train Loss:  0.007 | Train Accuracy:  0.784 | Val Loss:  0.007 | Val Accuracy:  0.811
100%|████████| 563/563 [16:58<00:00,  1.81s/it]
100%|████████| 63/63 [01:48<00:00,  1.72s/it]
Epochs: 3 | Train Loss:  0.007 | Train Accuracy:  0.805 | Val Loss:  0.006 | Val Accuracy:  0.822
100%|████████| 563/563 [17:03<00:00,  1.82s/it]
100%|████████| 63/63 [01:45<00:00,  1.68s/it]
Epochs: 4 | Train Loss:  0.006 | Train Accuracy:  0.820 | Val Loss:  0.006 | Val Accuracy:  0.831
100%|████████| 563/563 [17:16<00:00,  1.84s/it]
100%|████████| 63/63 [02:32<00:00,  2.42s/it]
Epochs: 5 | Train Loss:  0.006 | Train Accuracy:  0.829 | Val Loss:  0.006 | Val Accuracy:  0.843
(29843, 217.27602770924568, 3371, 22.142377495765686)
```

Show confusion matrix

(Best Accuracy)


```
accuracy measure:   84.96000000000001 %
precision measure:   84.99853461538461 %
recall measure:   84.96000000000001 %
f1 measure:   84.95409884594409 %
```

❖ **Running model (using 4 hidden layers)**

With preprocessing

LR=1e-6

```
model2 = BertClassifier2(512, 256, 128, 64)

Some weights of the model checkpoint at bert-base-cased were not used when initializing BertModel: ['
- This IS expected if you are initializing BertModel from the checkpoint of a model trained on anothe
- This IS NOT expected if you are initializing BertModel from the checkpoint of a model that you expe
◄

[ ]  run_model(model2, train_dataloader, val_dataloader, LR)

  0%|           | 0/563 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:2
100%|██████████| 563/563 [06:06<00:00,  1.54it/s]
100%|██████████| 63/63 [00:25<00:00,  2.51it/s]
Epochs: 1 | Train Loss:  0.011 | Train Accuracy:  0.531 | Val Loss:  0.011 | Val Accuracy:  0.664
100%|██████████| 563/563 [06:06<00:00,  1.54it/s]
100%|██████████| 63/63 [00:23<00:00,  2.68it/s]
Epochs: 2 | Train Loss:  0.010 | Train Accuracy:  0.740 | Val Loss:  0.010 | Val Accuracy:  0.798
(26627, 368.1936562657356, 3192, 38.11210232973099)
```
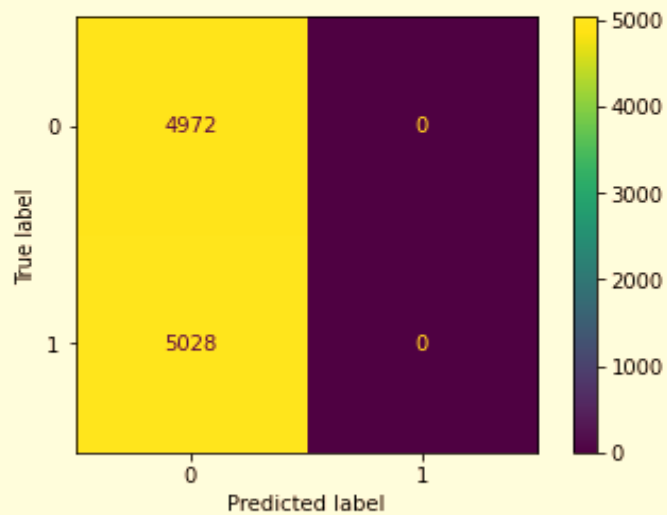
Show confusion matrix


```
accuracy measure:   49.72 %
precision measure:   24.720784 %
recall measure:   49.72 %
f1 measure:   33.022687683676196 %
```
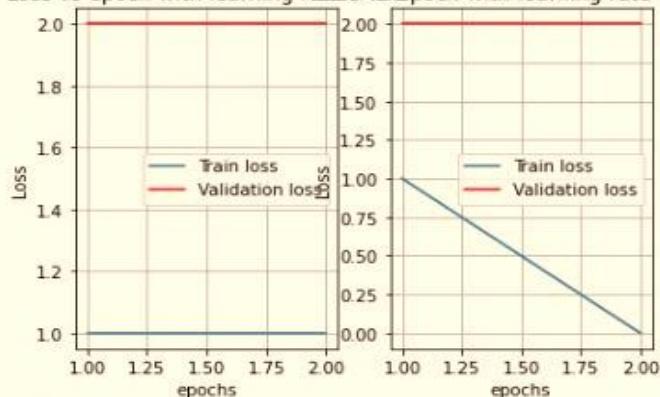
❖ **Running model**

Without preprocessing

○ LR=1e-2

```
100%|████████████| 563/563 [03:58<00:00,  2.36it/s]
100%|████████████| 63/63 [00:22<00:00,  2.80it/s]
Epochs: 1 | Train Loss:  0.000 | Train Accuracy:  0.500 | Val Loss:  0.001 | Val Accuracy:  0.493
100%|████████████| 563/563 [04:00<00:00,  2.34it/s]
100%|████████████| 63/63 [00:21<00:00,  2.96it/s]
Epochs: 2 | Train Loss:  0.000 | Train Accuracy:  0.500 | Val Loss:  0.001 | Val Accuracy:  0.493
547851
100%|████████████| 563/563 [04:00<00:00,  2.35it/s]
100%|████████████| 63/63 [00:23<00:00,  2.70it/s]
Epochs: 1 | Train Loss:  0.000 | Train Accuracy:  0.500 | Val Loss:  0.001 | Val Accuracy:  0.493
100%|████████████| 563/563 [04:01<00:00,  2.33it/s]
100%|████████████| 63/63 [00:21<00:00,  2.96it/s]
Epochs: 2 | Train Loss:  0.000 | Train Accuracy:  0.500 | Val Loss:  0.001 | Val Accuracy:  0.493
547851
100%|████████████| 563/563 [03:59<00:00,  2.35it/s]
100%|████████████| 63/63 [00:21<00:00,  2.98it/s]
Epochs: 1 | Train Loss:  0.000 | Train Accuracy:  0.500 | Val Loss:  0.001 | Val Accuracy:  0.493
100%|████████████| 563/563 [04:03<00:00,  2.31it/s]
100%|████████████| 63/63 [00:21<00:00,  2.95it/s]
Epochs: 2 | Train Loss:  0.000 | Train Accuracy:  0.500 | Val Loss:  0.001 | Val Accuracy:  0.493
547851
100%|████████████| 563/563 [04:01<00:00,  2.33it/s]
100%|████████████| 63/63 [00:21<00:00,  2.93it/s]
Epochs: 1 | Train Loss:  0.000 | Train Accuracy:  0.500 | Val Loss:  0.001 | Val Accuracy:  0.493
100%|████████████| 563/563 [04:01<00:00,  2.33it/s]
100%|████████████| 63/63 [00:23<00:00,  2.72it/s]
Epochs: 2 | Train Loss:  0.000 | Train Accuracy:  0.500 | Val Loss:  0.001 | Val Accuracy:  0.493
```
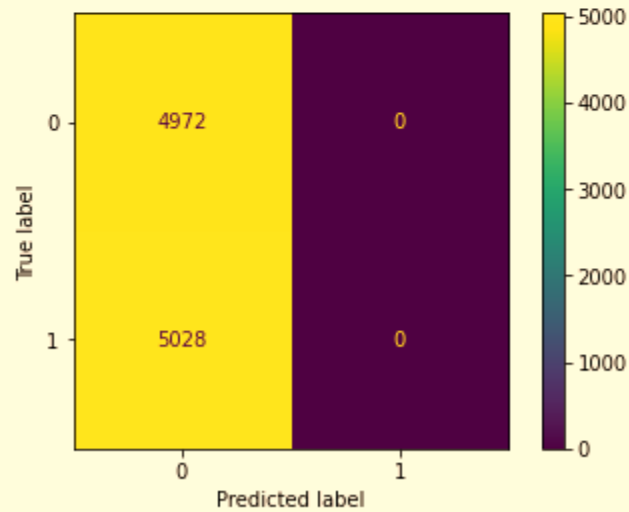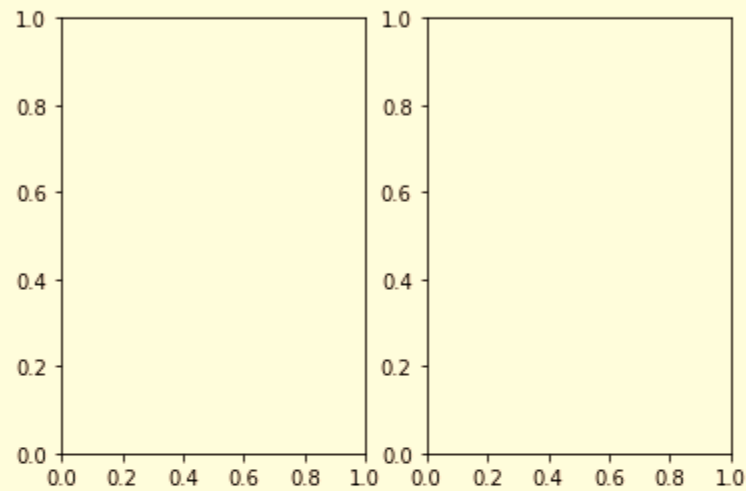

Loss vs epoch with learning rate = 0.01

```
Best loss: 1
Best accuracy: 2
Best learning rate: 1
```

Show confusion matrix

```
accuracy measure:   49.72 %
precision measure:   24.720784 %
recall measure:   49.72 %
f1 measure:   33.022687683676196 %
```
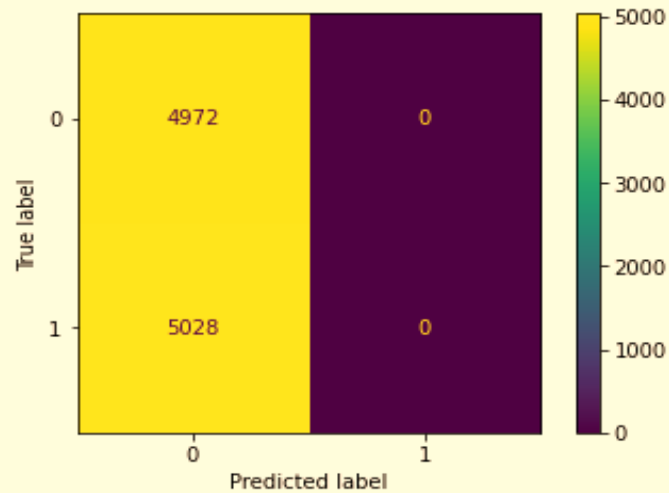
○   **LR=1e-6**

```
  0%|          | 0/563 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py
100%|██████████| 563/563 [05:32<00:00,  1.69it/s]
100%|██████████| 63/63 [00:23<00:00,  2.66it/s]
Epochs: 1 | Train Loss:  0.013 | Train Accuracy:  0.500 | Val Loss:  0.013 | Val Accuracy:  0.507
100%|██████████| 563/563 [05:33<00:00,  1.69it/s]
100%|██████████| 63/63 [00:23<00:00,  2.71it/s]
Epochs: 2 | Train Loss:  0.013 | Train Accuracy:  0.500 | Val Loss:  0.013 | Val Accuracy:  0.507
100%|██████████| 563/563 [05:33<00:00,  1.69it/s]
100%|██████████| 63/63 [00:23<00:00,  2.68it/s]
Epochs: 1 | Train Loss:  0.013 | Train Accuracy:  0.500 | Val Loss:  0.013 | Val Accuracy:  0.507
100%|██████████| 563/563 [05:33<00:00,  1.69it/s]
100%|██████████| 63/63 [00:23<00:00,  2.72it/s]
Epochs: 2 | Train Loss:  0.013 | Train Accuracy:  0.500 | Val Loss:  0.013 | Val Accuracy:  0.507
accuracy measure:  49.72 %
precision measure:  24.720784 %
recall measure:  49.72 %
f1 measure:  33.022687683676196 %
```
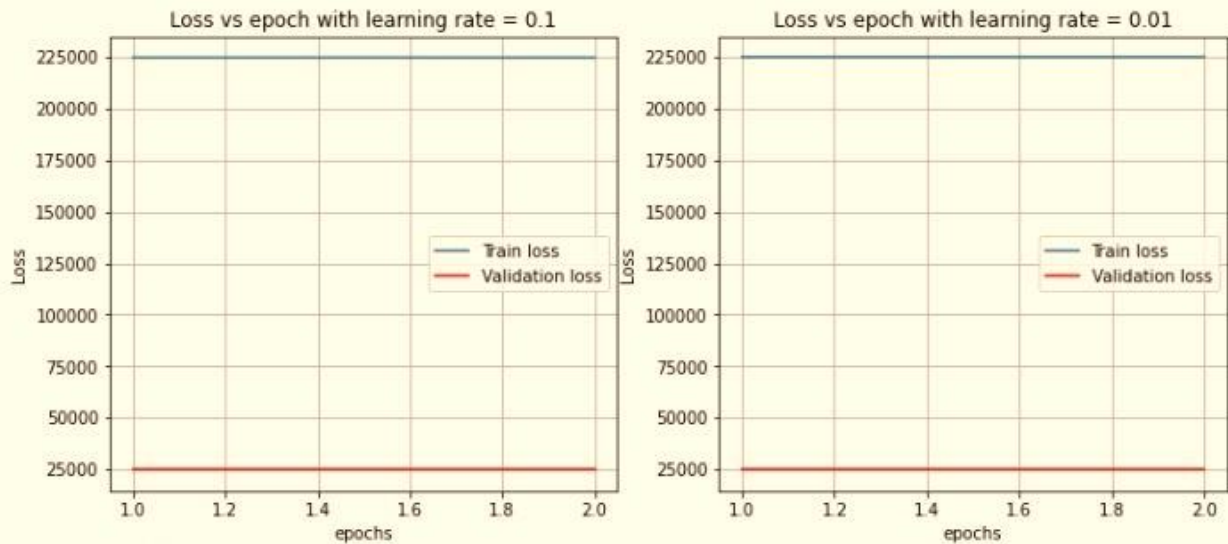
Show confusion matrix

```
accuracy measure:   49.72 %
precision measure:   24.720784 %
recall measure:   49.72 %
f1 measure:   33.022687683676196 %
/usr/local/lib/python3.7/dist-packages/sklea
  _warn_prf(average, modifier, msg_start, le
```



```
Best loss: 1
Best accuracy: 2
Best learning rate: 1
```

comparing between the best validation accuracies for the different values of learning rate:

Loss vs epoch with learning rate = 0.1 — Loss vs epoch with learning rate = 0.01

```
Best loss: 24662.5
Best accuracy: 1973
Best learning rate: 0.1
Test-acc = 5028
```

❖ Bonus part hidden layers:

LR = 1e-6

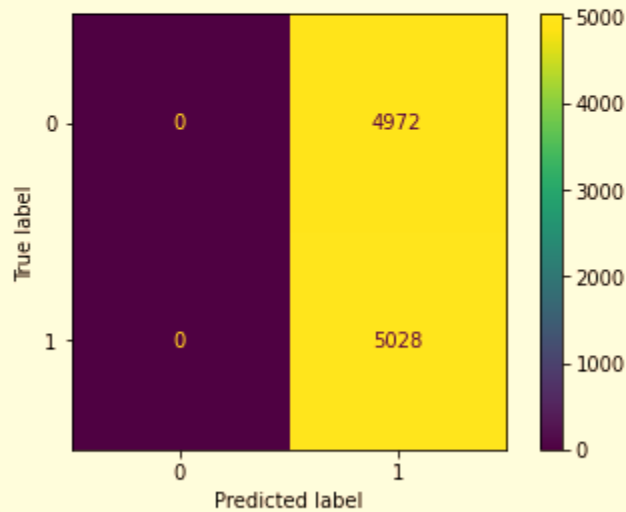**model = BertClassifier3(256, 128, 64, 32, 16)**

```
[ ]  run_model(model, train_dataloader, val_dataloader, LR)

       0%|           | 0/563 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:24:
     100%|███████████| 563/563 [06:07<00:00,  1.53it/s]
     100%|███████████| 63/63 [00:22<00:00,  2.74it/s]
     Epochs: 1 | Train Loss:  0.011 | Train Accuracy:  0.500 | Val Loss:  0.011 | Val Accuracy:  0.493
     100%|███████████| 563/563 [06:06<00:00,  1.54it/s]
     100%|███████████| 63/63 [00:23<00:00,  2.68it/s]
     Epochs: 2 | Train Loss:  0.011 | Train Accuracy:  0.500 | Val Loss:  0.011 | Val Accuracy:  0.493
     (17991, 388.54497665166855, 1973, 42.99576020240784)
```

Show confusion matrix

```
accuracy measure:   50.28 %
precision measure:   25.280784 %
recall measure:   50.28 %
f1 measure:   33.64490817141336 %
/usr/local/lib/python3.7/dist-packages/sklearn/metrics/
  _warn_prf(average, modifier, msg_start, len(result))
```

LR = 1e-6

model = BertClassifier4(64, 32, 16)

```
[ ]  run_model(model, train_dataloader, val_dataloader, LR)

      0%|            | 0/563 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:2
    100%|██████████| 563/563 [06:06<00:00,  1.53it/s]
    100%|██████████| 63/63 [00:23<00:00,  2.70it/s]
    Epochs: 1 | Train Loss:  0.011 | Train Accuracy:  0.517 | Val Loss:  0.011 | Val Accuracy:  0.630
    100%|██████████| 563/563 [06:06<00:00,  1.54it/s]
    100%|██████████| 63/63 [00:23<00:00,  2.71it/s]
    Epochs: 2 | Train Loss:  0.010 | Train Accuracy:  0.691 | Val Loss:  0.010 | Val Accuracy:  0.785
    (24885, 374.5296521782875, 3139, 39.54024076461792)
```

Show confusion matrix

(Second Best accuracy)

accuracy measure: 77.92999999999999 %
precision measure: 77.97828209900118 %
recall measure: 77.92999999999999 %
f1 measure: 77.92386839707602 %