```
 1 "C:\Program Files\Python311\python.exe" C:\Users\user
   \Desktop\Project\models\roberta_resnet\roberta_resnet
   .py
 2 Reading twitter - 1grams ...
 3 Reading twitter - 2grams ...
 4 Reading english - 1grams ...
 5 Reading twitter - 1grams ...
 6 Reading twitter - 2grams ...
 7 Reading english - 1grams ...
 8 Reading twitter - 1grams ...
 9 Reading twitter - 2grams ...
10 Reading english - 1grams ...
11 Some weights of RobertaModel were not initialized
   from the model checkpoint at roberta-base and are
   newly initialized: ['roberta.pooler.dense.bias', '
   roberta.pooler.dense.weight']
12 You should probably TRAIN this model on a down-stream
    task to be able to use it for predictions and
   inference.
13  Start Training on,  mami 1429 15 143
14 Epoch 0: 100%|          | 1429/1429 [2:21:45<00:00,
   5.95s/batch, loss=0.681]
15 loss  tensor(0.6920, device='cuda:0', grad_fn=<
   DivBackward0>)
16 Accuracy Score = 0.56
17 F1 Score (Macro) = 0.358974358974359
18 new best saving,  0.358974358974359
19 Best so far,  0.358974358974359
20 Epoch 1: 100%|          | 1429/1429 [2:22:15<00:00,
   5.97s/batch, loss=0.508]
21 loss  tensor(0.6482, device='cuda:0', grad_fn=<
   DivBackward0>)
22 Accuracy Score = 0.61
23 F1 Score (Macro) = 0.60996099609961
24 new best saving,  0.60996099609961
25 Best so far,  0.60996099609961
26 Epoch 2: 100%|          | 1429/1429 [3:07:16<00:00,
   7.86s/batch, loss=0.825]
27 loss  tensor(0.5854, device='cuda:0', grad_fn=<
   DivBackward0>)
28 Accuracy Score = 0.76
```

```
29 F1 Score (Macro) = 0.7519636213311286
30 new best saving,  0.7519636213311286
31 Best so far,  0.7519636213311286
32 Epoch 3: 100%|          | 1429/1429 [2:13:58<00:00,
   5.63s/batch, loss=0.254]
33 loss  tensor(0.5562, device='cuda:0', grad_fn=<
   DivBackward0>)
34 Accuracy Score = 0.76
35 F1 Score (Macro) = 0.75
36 Best so far,  0.7519636213311286
37 Epoch 4: 100%|          | 1429/1429 [2:11:50<00:00,
   5.54s/batch, loss=0.655]
38 loss  tensor(0.5255, device='cuda:0', grad_fn=<
   DivBackward0>)
39 Accuracy Score = 0.83
40 F1 Score (Macro) = 0.8270776116366595
41 new best saving,  0.8270776116366595
42 Best so far,  0.8270776116366595
43 Epoch 5: 100%|          | 1429/1429 [2:11:26<00:00,
   5.52s/batch, loss=0.668]
44 loss  tensor(0.4904, device='cuda:0', grad_fn=<
   DivBackward0>)
45 Accuracy Score = 0.85
46 F1 Score (Macro) = 0.8487750781328762
47 new best saving,  0.8487750781328762
48   0%|          | 0/1429 [00:00<?, ?batch/s]Best so
   far,  0.8487750781328762
49 Epoch 6: 100%|          | 1429/1429 [2:10:20<00:00,
   5.47s/batch, loss=0.494]
50 loss  tensor(0.4485, device='cuda:0', grad_fn=<
   DivBackward0>)
51 Accuracy Score = 0.88
52 F1 Score (Macro) = 0.8795664391810518
53 new best saving,  0.8795664391810518
54 Best so far,  0.8795664391810518
55 Epoch 7: 100%|          | 1429/1429 [2:08:23<00:00,
   5.39s/batch, loss=0.177]
56 loss  tensor(0.3944, device='cuda:0', grad_fn=<
   DivBackward0>)
57 Accuracy Score = 0.94
58 F1 Score (Macro) = 0.938423645320197
```

```
new best saving,  0.938423645320197
Best so far,  0.938423645320197
Epoch 8: 100%|          | 1429/1429 [2:10:36<00:00,
  5.48s/batch, loss=0.514]
loss  tensor(0.3449, device='cuda:0', grad_fn=<
  DivBackward0>)
Accuracy Score = 0.94
F1 Score (Macro) = 0.938423645320197
Best so far,  0.938423645320197
Epoch 9: 100%|          | 1429/1429 [2:12:29<00:00,
  5.56s/batch, loss=0.528]
loss  tensor(0.2745, device='cuda:0', grad_fn=<
  DivBackward0>)
Accuracy Score = 0.95
F1 Score (Macro) = 0.9493875898370281
new best saving,  0.9493875898370281
Best so far,  0.9493875898370281
Epoch 10: 100%|          | 1429/1429 [2:16:09<00:00
  ,  5.72s/batch, loss=0.232]
loss  tensor(0.2388, device='cuda:0', grad_fn=<
  DivBackward0>)
Accuracy Score = 0.95
F1 Score (Macro) = 0.9491404740107823
Best so far,  0.9493875898370281
Epoch 11: 100%|          | 1429/1429 [2:24:58<00:00
  ,  6.09s/batch, loss=0.0374]
loss  tensor(0.2139, device='cuda:0', grad_fn=<
  DivBackward0>)
Accuracy Score = 0.97
F1 Score (Macro) = 0.9694842844064693
new best saving,  0.9694842844064693
Best so far,  0.9694842844064693
Epoch 12: 100%|          | 1429/1429 [2:15:31<00:00
  ,  5.69s/batch, loss=0.0155]
loss  tensor(0.1846, device='cuda:0', grad_fn=<
  DivBackward0>)
Accuracy Score = 0.98
F1 Score (Macro) = 0.9797077922077921
new best saving,  0.9797077922077921
Best so far,  0.9797077922077921
Epoch 13: 100%|          | 1429/1429 [2:12:11<00:00
```

```
 89 ,  5.55s/batch, loss=0.348]
 90 loss  tensor(0.1620, device='cuda:0', grad_fn=<
    DivBackward0>)
 91   0%|          | 0/1429 [00:00<?, ?batch/s]Accuracy
    Score = 0.97
 92 F1 Score (Macro) = 0.9696325539022169
 93 Best so far,  0.9797077922077921
 94 Epoch 14: 100%|██████████| 1429/1429 [2:13:02<00:00
    ,  5.59s/batch, loss=0.00918]
 95 loss  tensor(0.1524, device='cuda:0', grad_fn=<
    DivBackward0>)
 96 Accuracy Score = 0.99
 97 F1 Score (Macro) = 0.9898775179674055
 98 new best saving,  0.9898775179674055
 99 Best so far,  0.9898775179674055
100 Epoch 15: 100%|██████████| 1429/1429 [2:15:20<00:00
    ,  5.68s/batch, loss=0.00887]
101 loss  tensor(0.1362, device='cuda:0', grad_fn=<
    DivBackward0>)
102   0%|          | 0/1429 [00:00<?, ?batch/s]Accuracy
    Score = 0.98
103 F1 Score (Macro) = 0.9797979797979798
104 Best so far,  0.9898775179674055
105 Epoch 16: 100%|██████████| 1429/1429 [6:06:33<00:00
    , 15.39s/batch, loss=0.445]
106 loss  tensor(0.1364, device='cuda:0', grad_fn=<
    DivBackward0>)
107   0%|          | 0/1429 [00:00<?, ?batch/s]Accuracy
    Score = 0.98
108 F1 Score (Macro) = 0.9797979797979798
109 Best so far,  0.9898775179674055
110 Epoch 17: 100%|██████████| 1429/1429 [2:11:13<00:00
    ,  5.51s/batch, loss=1.75]
111 loss  tensor(0.1184, device='cuda:0', grad_fn=<
    DivBackward0>)
112   0%|          | 0/1429 [00:00<?, ?batch/s]Accuracy
    Score = 0.98
113 F1 Score (Macro) = 0.9797979797979798
114 Best so far,  0.9898775179674055
115 Epoch 18: 100%|██████████| 1429/1429 [2:11:44<00:00
    ,  5.53s/batch, loss=0.0371]
```

```
116 loss  tensor(0.1085, device='cuda:0', grad_fn=<
    DivBackward0>)
117 Accuracy Score = 0.97
118 F1 Score (Macro) = 0.9697550156265753
119 Best so far,  0.9898775179674055
120 Epoch 19: 100%|          | 1429/1429 [2:09:03<00:00
    ,  5.42s/batch, loss=0.055]
121 loss  tensor(0.0970, device='cuda:0', grad_fn=<
    DivBackward0>)
122 Accuracy Score = 0.98
123 F1 Score (Macro) = 0.9797077922077921
124 Best so far,  0.9898775179674055
125 Final F1 score on test set:  mami 0.9898775179674055
126 Final Accuracy on test set:  mami 0.99
127 Reading twitter - 1grams ...
128 Reading twitter - 2grams ...
129 Reading english - 1grams ...
130 Reading twitter - 1grams ...
131 Reading twitter - 2grams ...
132 Reading english - 1grams ...
133 Reading twitter - 1grams ...
134 Reading twitter - 2grams ...
135 Reading english - 1grams ...
136 Some weights of RobertaModel were not initialized
    from the model checkpoint at roberta-base and are
    newly initialized: ['roberta.pooler.dense.bias', '
    roberta.pooler.dense.weight']
137 You should probably TRAIN this model on a down-
    stream task to be able to use it for predictions and
     inference.
138  Start Training on,  fb 1215 72 143
139 Epoch 0: 100%|          | 1215/1215 [2:05:16<00:00
    ,  6.19s/batch, loss=0.698]
140 loss  tensor(0.6681, device='cuda:0', grad_fn=<
    DivBackward0>)
141 Accuracy Score = 0.5
142 F1 Score (Macro) = 0.3333333333333333
143 new best saving,  0.3333333333333333
144 Best so far,  0.3333333333333333
145
```