training 1 epoch...

1580.595807313919 seconds used......

1 training finished! average train loss: 1.4164890121504998

total clf loss: 815.8664034157991 total adv loss: 446.91791632398963

evaluating...

Confusion matrix, without normalization

[[ 766 109 71]

[1213 3901 90]

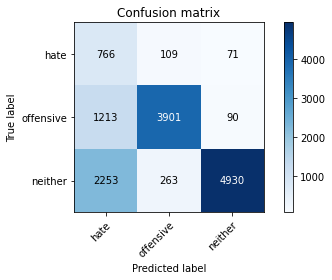
[2253 263 4930]]

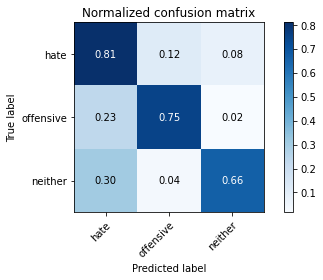
Normalized confusion matrix

[[0.80972516 0.11522199 0.07505285]

[0.23308993 0.74961568 0.01729439]

[0.30257857 0.03532098 0.66210046]]





precision recall f1-score support

hate 0.18 0.81 0.30 946

offensive 0.91 0.75 0.82 5204

neither 0.97 0.66 0.79 7446

accuracy 0.71 13596

macro avg 0.69 0.74 0.64 13596

weighted avg 0.89 0.71 0.77 13596

average eval\_loss: 0.6822710173752378, accuracy: 0.7058693733451015

training 2 epoch...

1580.699589252472 seconds used......

2 training finished! average train loss: 1.1549671628846545

total clf loss: 660.5815692096949 total adv loss: 366.7318984698504

evaluating...

Confusion matrix, without normalization

[[ 657 135 154]

[ 860 4106 238]

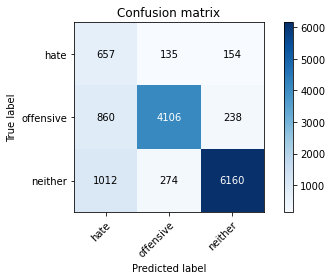
[1012 274 6160]]

Normalized confusion matrix

[[0.69450317 0.14270613 0.1627907 ]

[0.16525749 0.78900846 0.04573405]

[0.1359119 0.03679828 0.82728982]]





precision recall f1-score support

hate 0.26 0.69 0.38 946

offensive 0.91 0.79 0.84 5204

neither 0.94 0.83 0.88 7446

accuracy 0.80 13596

macro avg 0.70 0.77 0.70 13596

weighted avg 0.88 0.80 0.83 13596

average eval\_loss: 0.4976462957832743, accuracy: 0.8033980582524272

training 3 epoch...

1581.2898864746094 seconds used......

3 training finished! average train loss: 0.9820470814627859

total clf loss: 563.7267547696829 total adv loss: 310.80203691124916

evaluating...

Confusion matrix, without normalization

[[ 563 246 137]

[ 458 4595 151]

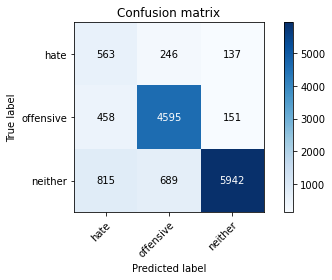
[ 815 689 5942]]

Normalized confusion matrix

[[0.59513742 0.26004228 0.1448203 ]

[0.08800922 0.88297463 0.02901614]

[0.10945474 0.0925329 0.79801236]]





precision recall f1-score support

hate 0.31 0.60 0.40 946

offensive 0.83 0.88 0.86 5204

neither 0.95 0.80 0.87 7446

accuracy 0.82 13596

macro avg 0.70 0.76 0.71 13596

weighted avg 0.86 0.82 0.83 13596

average eval\_loss: 0.4932992054259076, accuracy: 0.8164165931156222

Confusion matrix, without normalization

[[1584 120 131]

[1816 3959 150]

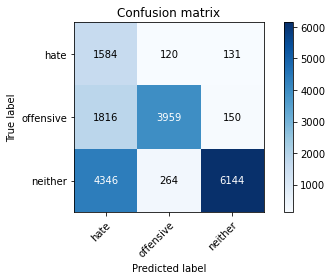
[4346 264 6144]]

Normalized confusion matrix

[[0.86321526 0.0653951 0.07138965]

[0.30649789 0.66818565 0.02531646]

[0.4041287 0.02454901 0.5713223 ]]





precision recall f1-score support

hate 0.20 0.86 0.33 1835

offensive 0.91 0.67 0.77 5925

neither 0.96 0.57 0.72 10754

accuracy 0.63 18514

macro avg 0.69 0.70 0.61 18514

weighted avg 0.87 0.63 0.70 18514

average eval\_loss: 0.8039760288679831, accuracy: 0.6312520254942205

[21]:

0.6312520254942205