training 1 epoch...

820.9811835289001 seconds used......

1 training finished! average train loss: 1.3694700901925119

total clf loss: 787.2896695435047 total adv loss: 432.8303650394082

evaluating...

Confusion matrix, without normalization

[[ 663 123 160]

[ 947 4051 206]

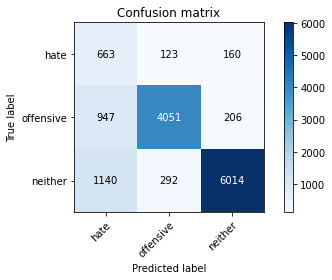
[1140 292 6014]]

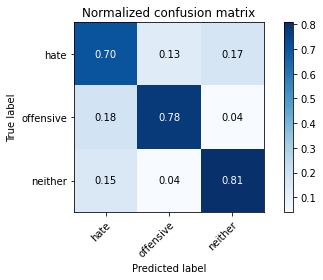
Normalized confusion matrix

[[0.70084567 0.13002114 0.16913319]

[0.1819754 0.77843966 0.03958493]

[0.15310234 0.03921569 0.80768198]]





precision recall f1-score support

hate 0.24 0.70 0.36 946

offensive 0.91 0.78 0.84 5204

neither 0.94 0.81 0.87 7446

accuracy 0.79 13596

macro avg 0.70 0.76 0.69 13596

weighted avg 0.88 0.79 0.82 13596

average eval\_loss: 0.5224491288556772, accuracy: 0.7890556045895851

training 2 epoch...

821.0218353271484 seconds used......

2 training finished! average train loss: 1.104982332370453

total clf loss: 631.205634303391 total adv loss: 351.2540201693773

evaluating...

Confusion matrix, without normalization

[[ 673 133 140]

[ 858 4149 197]

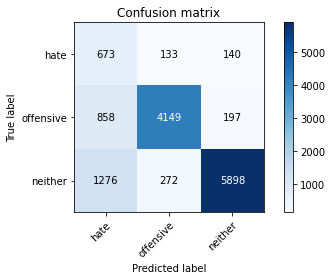
[1276 272 5898]]

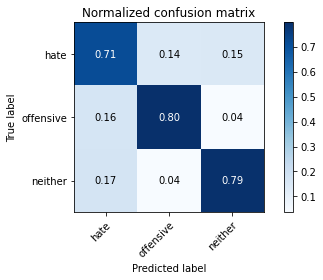
Normalized confusion matrix

[[0.71141649 0.14059197 0.14799154]

[0.16487317 0.79727133 0.0378555 ]

[0.17136718 0.03652968 0.79210314]]





precision recall f1-score support

hate 0.24 0.71 0.36 946

offensive 0.91 0.80 0.85 5204

neither 0.95 0.79 0.86 7446

accuracy 0.79 13596

macro avg 0.70 0.77 0.69 13596

weighted avg 0.88 0.79 0.82 13596

average eval\_loss: 0.5164110729580417, accuracy: 0.7884671962341865

training 3 epoch...

820.6553354263306 seconds used......

3 training finished! average train loss: 0.8779005713964555

total clf loss: 504.31018217653036 total adv loss: 277.657903123647

evaluating...

Confusion matrix, without normalization

[[ 643 136 167]

[ 787 4141 276]

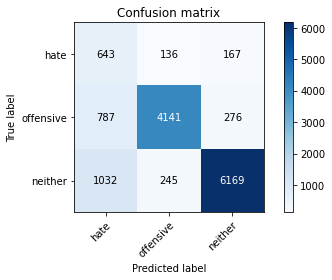
[1032 245 6169]]

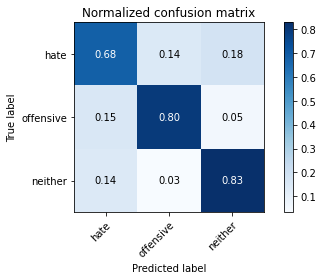
Normalized confusion matrix

[[0.67970402 0.14376321 0.17653277]

[0.15122982 0.79573405 0.05303613]

[0.1385979 0.03290357 0.82849852]]





precision recall f1-score support

hate 0.26 0.68 0.38 946

offensive 0.92 0.80 0.85 5204

neither 0.93 0.83 0.88 7446

accuracy 0.81 13596

macro avg 0.70 0.77 0.70 13596

weighted avg 0.88 0.81 0.83 13596

average eval\_loss: 0.5190276961883201, accuracy: 0.8056045895851721

Confusion matrix, without normalization

[[1449 166 220]

[1341 4272 312]

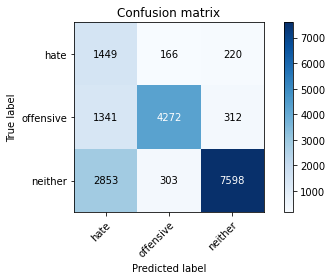
[2853 303 7598]]

Normalized confusion matrix

[[0.78964578 0.09046322 0.11989101]

[0.22632911 0.72101266 0.05265823]

[0.26529663 0.02817556 0.7065278 ]]





precision recall f1-score support

hate 0.26 0.79 0.39 1835

offensive 0.90 0.72 0.80 5925

neither 0.93 0.71 0.80 10754

accuracy 0.72 18514

macro avg 0.70 0.74 0.66 18514

weighted avg 0.86 0.72 0.76 18514

average eval\_loss: 0.6432414342112212, accuracy: 0.7194015339742897

[20]:

2 run

training 1 epoch...

809.5431501865387 seconds used......

1 training finished! average train loss: 1.3805274788564952

total clf loss: 793.1441584378481 total adv loss: 436.57625536993146

evaluating...

Confusion matrix, without normalization

[[ 726 116 104]

[1055 3978 171]

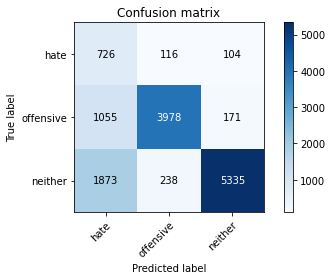
[1873 238 5335]]

Normalized confusion matrix

[[0.76744186 0.12262156 0.10993658]

[0.20272867 0.76441199 0.03285934]

[0.25154445 0.03196347 0.71649208]]





precision recall f1-score support

hate 0.20 0.77 0.32 946

offensive 0.92 0.76 0.83 5204

neither 0.95 0.72 0.82 7446

accuracy 0.74 13596

macro avg 0.69 0.75 0.66 13596

weighted avg 0.89 0.74 0.79 13596

average eval\_loss: 0.6107651100877454, accuracy: 0.7383789349808767

training 2 epoch...

809.4902048110962 seconds used......

2 training finished! average train loss: 1.09872059857391

total clf loss: 629.9882264584303 total adv loss: 348.0837684907019

evaluating...

Confusion matrix, without normalization

[[ 708 113 125]

[1007 3961 236]

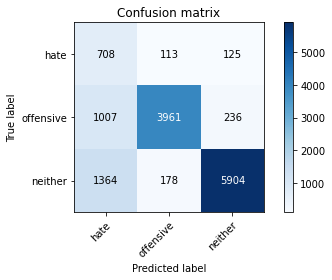
[1364 178 5904]]

Normalized confusion matrix

[[0.74841438 0.11945032 0.13213531]

[0.193505 0.76114527 0.04534973]

[0.1831856 0.02390545 0.79290894]]





precision recall f1-score support

hate 0.23 0.75 0.35 946

offensive 0.93 0.76 0.84 5204

neither 0.94 0.79 0.86 7446

accuracy 0.78 13596

macro avg 0.70 0.77 0.68 13596

weighted avg 0.89 0.78 0.82 13596

average eval\_loss: 0.54564766343683, accuracy: 0.7776551927037364

training 3 epoch...

809.5689709186554 seconds used......

3 training finished! average train loss: 0.8762056681990327

total clf loss: 504.15314220637083 total adv loss: 276.7135496586561

evaluating...

Confusion matrix, without normalization

[[ 570 174 202]

[ 665 4267 272]

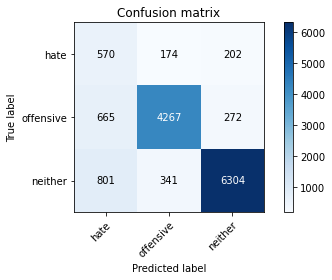
[ 801 341 6304]]

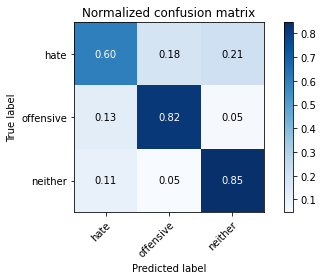
Normalized confusion matrix

[[0.602537 0.18393235 0.21353066]

[0.12778632 0.8199462 0.05226749]

[0.10757454 0.0457964 0.84662906]]





precision recall f1-score support

hate 0.28 0.60 0.38 946

offensive 0.89 0.82 0.85 5204

neither 0.93 0.85 0.89 7446

accuracy 0.82 13596

macro avg 0.70 0.76 0.71 13596

weighted avg 0.87 0.82 0.84 13596

average eval\_loss: 0.48369855968391195, accuracy: 0.8194321859370403

Confusion matrix, without normalization

[[1528 143 164]

[1614 4060 251]

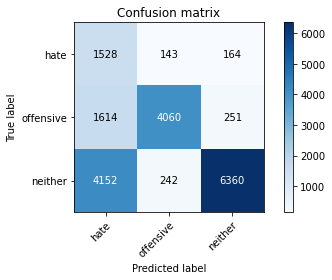
[4152 242 6360]]

Normalized confusion matrix

[[0.83269755 0.07792916 0.0893733 ]

[0.27240506 0.68523207 0.04236287]

[0.3860889 0.02250325 0.59140785]]





precision recall f1-score support

hate 0.21 0.83 0.33 1835

offensive 0.91 0.69 0.78 5925

neither 0.94 0.59 0.73 10754

accuracy 0.65 18514

macro avg 0.69 0.70 0.61 18514

weighted avg 0.86 0.65 0.71 18514

average eval\_loss: 0.7626169219967098, accuracy: 0.6453494652695257

0.6453494652695257

3 run

training 1 epoch...

808.9357578754425 seconds used......

1 training finished! average train loss: 1.3940498033089388

total clf loss: 799.8254721462727 total adv loss: 441.39631973206997

evaluating...

Confusion matrix, without normalization

[[ 649 148 149]

[ 866 4135 203]

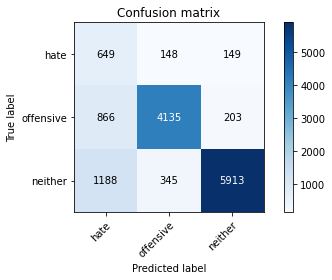
[1188 345 5913]]

Normalized confusion matrix

[[0.68604651 0.1564482 0.15750529]

[0.16641045 0.79458109 0.03900846]

[0.15954875 0.0463336 0.79411765]]





precision recall f1-score support

hate 0.24 0.69 0.36 946

offensive 0.89 0.79 0.84 5204

neither 0.94 0.79 0.86 7446

accuracy 0.79 13596

macro avg 0.69 0.76 0.69 13596

weighted avg 0.88 0.79 0.82 13596

average eval\_loss: 0.5333022186203914, accuracy: 0.7867755222124154

training 2 epoch...

808.76562333107 seconds used......

2 training finished! average train loss: 1.0999275105939785

total clf loss: 627.4801021665335 total adv loss: 350.0662025436759

evaluating...

Confusion matrix, without normalization

[[ 685 122 139]

[ 954 4002 248]

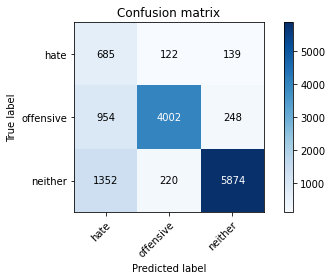
[1352 220 5874]]

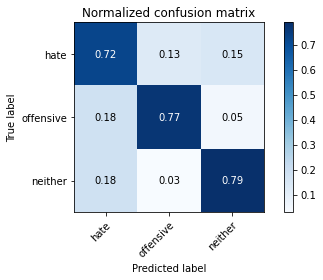
Normalized confusion matrix

[[0.72410148 0.12896406 0.14693446]

[0.18332052 0.76902383 0.04765565]

[0.181574 0.02954607 0.78887994]]





precision recall f1-score support

hate 0.23 0.72 0.35 946

offensive 0.92 0.77 0.84 5204

neither 0.94 0.79 0.86 7446

accuracy 0.78 13596

macro avg 0.70 0.76 0.68 13596

weighted avg 0.88 0.78 0.81 13596

average eval\_loss: 0.5280043711228406, accuracy: 0.7767725801706384

training 3 epoch...

Confusion matrix, without normalization

[[1412 151 272]

[1435 4126 364]

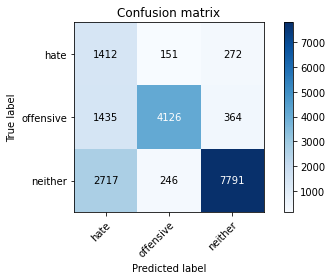
[2717 246 7791]]

Normalized confusion matrix

[[0.76948229 0.08228883 0.14822888]

[0.24219409 0.69637131 0.0614346 ]

[0.25265018 0.02287521 0.72447461]]





precision recall f1-score support

hate 0.25 0.77 0.38 1835

offensive 0.91 0.70 0.79 5925

neither 0.92 0.72 0.81 10754

accuracy 0.72 18514

macro avg 0.70 0.73 0.66 18514

weighted avg 0.85 0.72 0.76 18514

average eval\_loss: 0.6380681810703442, accuracy: 0.719941665766447

0.719941665766447

4 run

training 1 epoch...

810.7532229423523 seconds used......

1 training finished! average train loss: 1.3852340181010165

total clf loss: 797.2340463846922 total adv loss: 437.37170696258545

evaluating...

Confusion matrix, without normalization

[[ 631 158 157]

[ 797 4193 214]

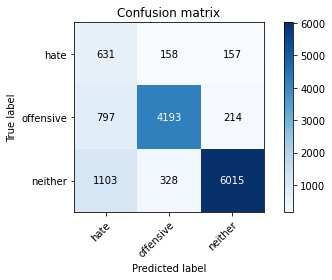
[1103 328 6015]]

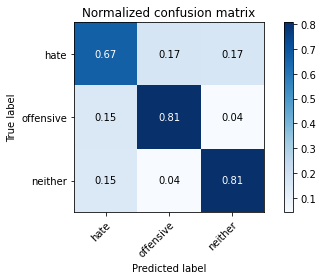
Normalized confusion matrix

[[0.66701903 0.16701903 0.16596195]

[0.15315142 0.80572636 0.04112221]

[0.14813323 0.0440505 0.80781628]]





precision recall f1-score support

hate 0.25 0.67 0.36 946

offensive 0.90 0.81 0.85 5204

neither 0.94 0.81 0.87 7446

accuracy 0.80 13596

macro avg 0.70 0.76 0.69 13596

weighted avg 0.88 0.80 0.83 13596

average eval\_loss: 0.504848613677656, accuracy: 0.7972197705207414

training 2 epoch...

810.2944974899292 seconds used......

2 training finished! average train loss: 1.1100409561035154

total clf loss: 634.9292114675045 total adv loss: 352.44511123746634

evaluating...

Confusion matrix, without normalization

[[ 674 144 128]

[ 861 4204 139]

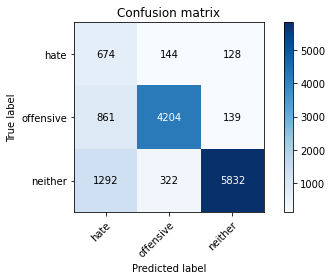
[1292 322 5832]]

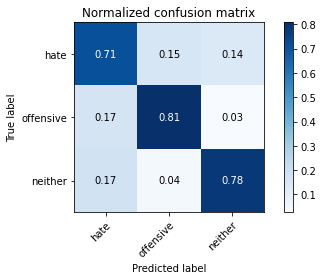
Normalized confusion matrix

[[0.71247357 0.15221987 0.13530655]

[0.16544965 0.80784012 0.02671022]

[0.17351598 0.0432447 0.78323932]]





precision recall f1-score support

hate 0.24 0.71 0.36 946

offensive 0.90 0.81 0.85 5204

neither 0.96 0.78 0.86 7446

accuracy 0.79 13596

macro avg 0.70 0.77 0.69 13596

weighted avg 0.88 0.79 0.82 13596

average eval\_loss: 0.5148591061067932, accuracy: 0.7877316857899382

training 3 epoch...

810.1551558971405 seconds used......

3 training finished! average train loss: 0.8886333239243953

total clf loss: 509.3418358154595 total adv loss: 281.6192927416414

evaluating...

Confusion matrix, without normalization

[[ 650 146 150]

[ 940 4053 211]

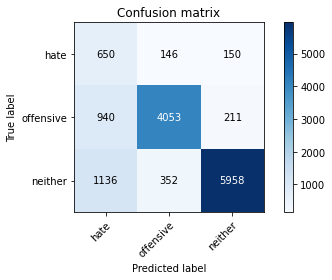
[1136 352 5958]]

Normalized confusion matrix

[[0.68710359 0.15433404 0.15856237]

[0.18063028 0.77882398 0.04054573]

[0.15256514 0.0472737 0.80016116]]





precision recall f1-score support

hate 0.24 0.69 0.35 946

offensive 0.89 0.78 0.83 5204

neither 0.94 0.80 0.87 7446

accuracy 0.78 13596

macro avg 0.69 0.76 0.68 13596

weighted avg 0.87 0.78 0.82 13596

average eval\_loss: 0.5553761613390902, accuracy: 0.7841276846131215

Confusion matrix, without normalization

[[1454 183 198]

[1300 4359 266]

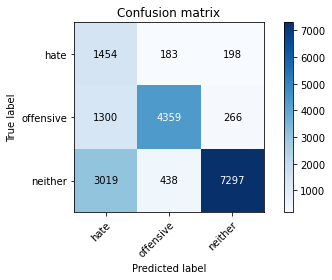
[3019 438 7297]]

Normalized confusion matrix

[[0.79237057 0.09972752 0.10790191]

[0.21940928 0.7356962 0.04489451]

[0.28073275 0.04072903 0.67853822]]





precision recall f1-score support

hate 0.25 0.79 0.38 1835

offensive 0.88 0.74 0.80 5925

neither 0.94 0.68 0.79 10754

accuracy 0.71 18514

macro avg 0.69 0.74 0.66 18514

weighted avg 0.85 0.71 0.75 18514

average eval\_loss: 0.6532344136781363, accuracy: 0.7081127795182024

[19]:

0.7081127795182024

5 run

training 1 epoch...

810.2652530670166 seconds used......

1 training finished! average train loss: 1.3823327936460483

total clf loss: 794.0817585736513 total adv loss: 437.1969612650573

evaluating...

Confusion matrix, without normalization

[[ 699 123 124]

[1061 3993 150]

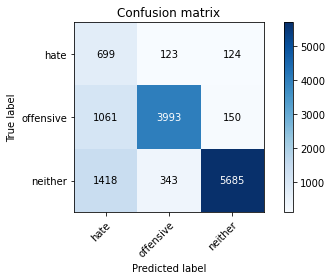
[1418 343 5685]]

Normalized confusion matrix

[[0.73890063 0.13002114 0.13107822]

[0.20388163 0.76729439 0.02882398]

[0.19043782 0.046065 0.76349718]]





precision recall f1-score support

hate 0.22 0.74 0.34 946

offensive 0.90 0.77 0.83 5204

neither 0.95 0.76 0.85 7446

accuracy 0.76 13596

macro avg 0.69 0.76 0.67 13596

weighted avg 0.88 0.76 0.80 13596

average eval\_loss: 0.5455694466610165, accuracy: 0.7632391879964695

training 2 epoch...

810.6149859428406 seconds used......

2 training finished! average train loss: 1.100736927645406

total clf loss: 631.5042068064213 total adv loss: 348.5426328703761

evaluating...

Confusion matrix, without normalization

[[ 713 141 92]

[ 877 4200 127]

[1639 405 5402]]

Normalized confusion matrix

[[0.75369979 0.14904863 0.09725159]

[0.16852421 0.80707148 0.0244043 ]

[0.22011818 0.05439162 0.7254902 ]]

810.61745262146 seconds used......

3 training finished! average train loss: 0.8752650315468234

total clf loss: 505.45853316038847 total adv loss: 275.4931797068566

evaluating...

Confusion matrix, without normalization

[[ 660 127 159]

[ 926 4057 221]

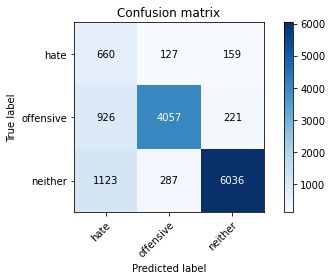
[1123 287 6036]]

Normalized confusion matrix

[[0.69767442 0.13424947 0.16807611]

[0.17794005 0.77959262 0.04246733]

[0.15081923 0.03854418 0.81063658]]





precision recall f1-score support

hate 0.24 0.70 0.36 946

offensive 0.91 0.78 0.84 5204

neither 0.94 0.81 0.87 7446

accuracy 0.79 13596

macro avg 0.70 0.76 0.69 13596

weighted avg 0.88 0.79 0.82 13596

average eval\_loss: 0.554430223738446, accuracy: 0.7908943807002059

Confusion matrix, without normalization

[[1471 202 162]

[1331 4398 196]

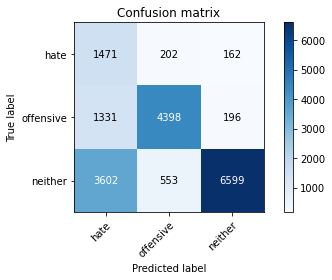
[3602 553 6599]]

Normalized confusion matrix

[[0.80163488 0.11008174 0.08828338]

[0.22464135 0.74227848 0.03308017]

[0.33494514 0.05142273 0.61363214]]





precision recall f1-score support

hate 0.23 0.80 0.36 1835

offensive 0.85 0.74 0.79 5925

neither 0.95 0.61 0.75 10754

accuracy 0.67 18514

macro avg 0.68 0.72 0.63 18514

weighted avg 0.85 0.67 0.72 18514

average eval\_loss: 0.7291162181203102, accuracy: 0.6734363184617047