alpha = 2

1 run......

Downloading: 100%

411/411 [00:00<00:00, 13.9kB/s]

Downloading: 100%

263M/263M [00:08<00:00, 30.5MB/s]

training 1 epoch...

116.92729687690735 seconds used......

1 training finished! average train loss: 2.048751645854541

total clf loss: 166.32282680273056 total adv loss: 88.93372517824173

evaluating...

Confusion matrix, without normalization

[[ 94 23 7]

[ 49 110 5]

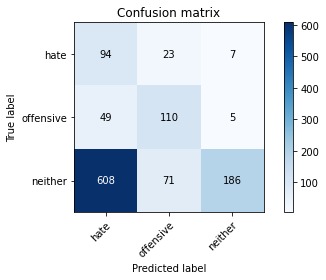
[608 71 186]]

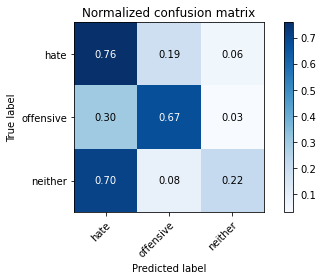
Normalized confusion matrix

[[0.75806452 0.18548387 0.05645161]

[0.29878049 0.67073171 0.0304878 ]

[0.70289017 0.08208092 0.2150289 ]]





precision recall f1-score support

hate 0.13 0.76 0.21 124

offensive 0.54 0.67 0.60 164

neither 0.94 0.22 0.35 865

accuracy 0.34 1153

macro avg 0.53 0.55 0.39 1153

weighted avg 0.79 0.34 0.37 1153

average eval\_loss: 0.956808445782497, accuracy: 0.3382480485689506

training 2 epoch...

116.06594586372375 seconds used......

2 training finished! average train loss: 1.7133670725992747

total clf loss: 143.82660365104675 total adv loss: 72.00953224301338

evaluating...

Confusion matrix, without normalization

[[102 16 6]

[ 55 102 7]

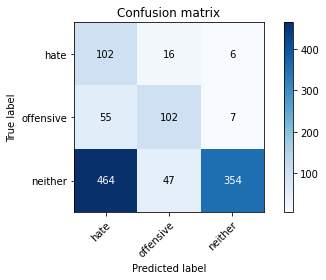
[464 47 354]]

Normalized confusion matrix

[[0.82258065 0.12903226 0.0483871 ]

[0.33536585 0.62195122 0.04268293]

[0.53641618 0.05433526 0.40924855]]





precision recall f1-score support

hate 0.16 0.82 0.27 124

offensive 0.62 0.62 0.62 164

neither 0.96 0.41 0.57 865

accuracy 0.48 1153

macro avg 0.58 0.62 0.49 1153

weighted avg 0.83 0.48 0.55 1153

average eval\_loss: 0.8778123526737608, accuracy: 0.4839549002601908

training 3 epoch...

116.17527508735657 seconds used......

3 training finished! average train loss: 1.3549143391705694

total clf loss: 120.90428766608238 total adv loss: 53.36066050082445

evaluating...

Confusion matrix, without normalization

[[ 97 16 11]

[ 49 104 11]

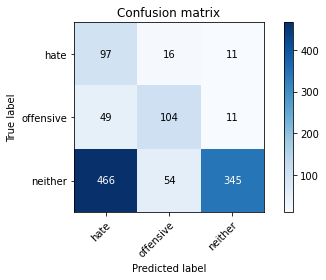
[466 54 345]]

Normalized confusion matrix

[[0.78225806 0.12903226 0.08870968]

[0.29878049 0.63414634 0.06707317]

[0.53872832 0.06242775 0.39884393]]





precision recall f1-score support

hate 0.16 0.78 0.26 124

offensive 0.60 0.63 0.62 164

neither 0.94 0.40 0.56 865

accuracy 0.47 1153

macro avg 0.57 0.61 0.48 1153

weighted avg 0.81 0.47 0.54 1153

average eval\_loss: 0.9652609023554571, accuracy: 0.4735472679965308

training 4 epoch...

116.09383392333984 seconds used......

4 training finished! average train loss: 0.9513354950717517

total clf loss: 94.26471583545208 total adv loss: 32.77982360124588

evaluating...

Confusion matrix, without normalization

[[ 89 27 8]

[ 45 108 11]

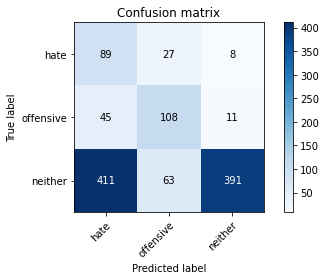
[411 63 391]]

Normalized confusion matrix

[[0.71774194 0.21774194 0.06451613]

[0.27439024 0.65853659 0.06707317]

[0.47514451 0.07283237 0.45202312]]





precision recall f1-score support

hate 0.16 0.72 0.27 124

offensive 0.55 0.66 0.60 164

neither 0.95 0.45 0.61 865

accuracy 0.51 1153

macro avg 0.55 0.61 0.49 1153

weighted avg 0.81 0.51 0.57 1153

average eval\_loss: 1.049618725118966, accuracy: 0.5099739809193409

training 5 epoch...

116.14886140823364 seconds used......

5 training finished! average train loss: 0.5813559436549743

total clf loss: 69.62494164705276 total adv loss: 14.021428364794701

evaluating...

Confusion matrix, without normalization

[[ 91 26 7]

[ 42 114 8]

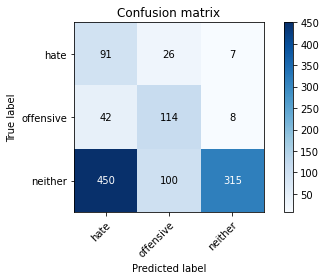
[450 100 315]]

Normalized confusion matrix

[[0.73387097 0.20967742 0.05645161]

[0.25609756 0.69512195 0.04878049]

[0.52023121 0.11560694 0.36416185]]





precision recall f1-score support

hate 0.16 0.73 0.26 124

offensive 0.47 0.70 0.56 164

neither 0.95 0.36 0.53 865

accuracy 0.45 1153

macro avg 0.53 0.60 0.45 1153

weighted avg 0.80 0.45 0.50 1153

average eval\_loss: 1.3413900352757553, accuracy: 0.45099739809193407

training 6 epoch...

116.11472225189209 seconds used......

6 training finished! average train loss: 0.4077856588576521

total clf loss: 50.8757682479918 total adv loss: 8.816111269276007

evaluating...

Confusion matrix, without normalization

[[101 19 4]

[ 45 111 8]

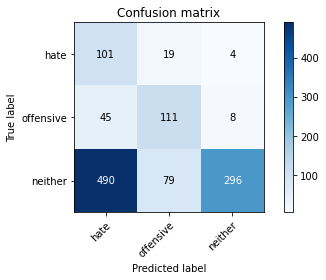
[490 79 296]]

Normalized confusion matrix

[[0.81451613 0.15322581 0.03225806]

[0.27439024 0.67682927 0.04878049]

[0.56647399 0.09132948 0.34219653]]





precision recall f1-score support

hate 0.16 0.81 0.27 124

offensive 0.53 0.68 0.60 164

neither 0.96 0.34 0.50 865

accuracy 0.44 1153

macro avg 0.55 0.61 0.46 1153

weighted avg 0.81 0.44 0.49 1153

average eval\_loss: 1.6456546290167446, accuracy: 0.4405897658282741

training 7 epoch...

116.17559552192688 seconds used......

7 training finished! average train loss: 0.2979960080147499

total clf loss: 35.7453198954463 total adv loss: 7.159004712011665

evaluating...

Confusion matrix, without normalization

[[ 98 22 4]

[ 44 112 8]

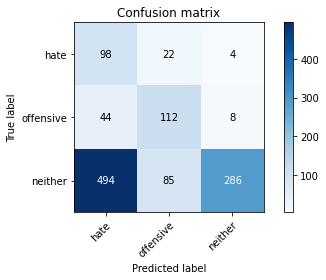
[494 85 286]]

Normalized confusion matrix

[[0.79032258 0.17741935 0.03225806]

[0.26829268 0.68292683 0.04878049]

[0.57109827 0.0982659 0.33063584]]





precision recall f1-score support

hate 0.15 0.79 0.26 124

offensive 0.51 0.68 0.58 164

neither 0.96 0.33 0.49 865

accuracy 0.43 1153

macro avg 0.54 0.60 0.44 1153

weighted avg 0.81 0.43 0.48 1153

average eval\_loss: 1.9614461660385132, accuracy: 0.43018213356461404

training 8 epoch...

116.20279431343079 seconds used......

8 training finished! average train loss: 0.210449840629562

total clf loss: 24.893989080563188 total adv loss: 5.230792101116094

evaluating...

Confusion matrix, without normalization

[[ 92 21 11]

[ 48 103 13]

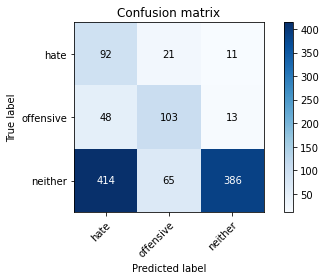
[414 65 386]]

Normalized confusion matrix

[[0.74193548 0.16935484 0.08870968]

[0.29268293 0.62804878 0.07926829]

[0.47861272 0.07514451 0.44624277]]





precision recall f1-score support

hate 0.17 0.74 0.27 124

offensive 0.54 0.63 0.58 164

neither 0.94 0.45 0.61 865

accuracy 0.50 1153

macro avg 0.55 0.61 0.49 1153

weighted avg 0.80 0.50 0.57 1153

average eval\_loss: 1.7538628711782653, accuracy: 0.5039028620988725

training 9 epoch...

116.27554178237915 seconds used......

9 training finished! average train loss: 0.12364781470525832

total clf loss: 15.68423028755933 total adv loss: 2.544301307782007

evaluating...

Confusion matrix, without normalization

[[109 10 5]

[ 65 91 8]

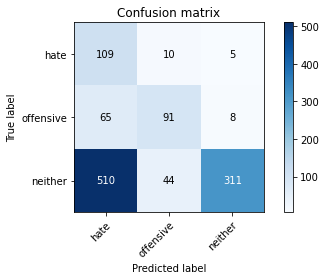
[510 44 311]]

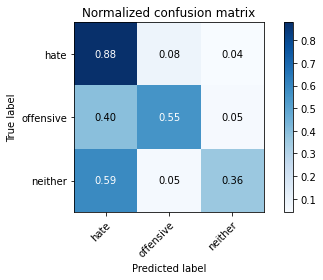
Normalized confusion matrix

[[0.87903226 0.08064516 0.04032258]

[0.39634146 0.55487805 0.04878049]

[0.58959538 0.05086705 0.35953757]]





precision recall f1-score support

hate 0.16 0.88 0.27 124

offensive 0.63 0.55 0.59 164

neither 0.96 0.36 0.52 865

accuracy 0.44 1153

macro avg 0.58 0.60 0.46 1153

weighted avg 0.83 0.44 0.51 1153

average eval\_loss: 2.4124159525180686, accuracy: 0.44319167389418906

training 10 epoch...

116.23623180389404 seconds used......

10 training finished! average train loss: 0.08131899891997732

total clf loss: 9.831605307292193 total adv loss: 1.9149932780992458

evaluating...

Confusion matrix, without normalization

[[104 16 4]

[ 49 103 12]

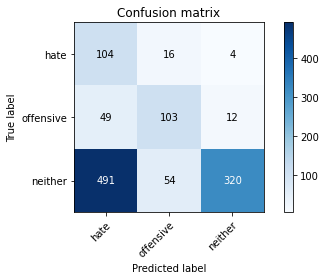
[491 54 320]]

Normalized confusion matrix

[[0.83870968 0.12903226 0.03225806]

[0.29878049 0.62804878 0.07317073]

[0.56763006 0.06242775 0.3699422 ]]





precision recall f1-score support

hate 0.16 0.84 0.27 124

offensive 0.60 0.63 0.61 164

neither 0.95 0.37 0.53 865

accuracy 0.46 1153

macro avg 0.57 0.61 0.47 1153

weighted avg 0.82 0.46 0.52 1153

average eval\_loss: 2.425491096644566, accuracy: 0.4570685169124024

2 run......

training 1 epoch...

116.23918151855469 seconds used......

1 training finished! average train loss: 2.0686380905764445

total clf loss: 166.9477556347847 total adv loss: 90.29172170162201

evaluating...

Confusion matrix, without normalization

[[ 86 23 15]

[ 48 106 10]

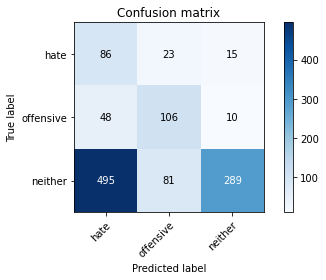
[495 81 289]]

Normalized confusion matrix

[[0.69354839 0.18548387 0.12096774]

[0.29268293 0.64634146 0.06097561]

[0.57225434 0.09364162 0.33410405]]





precision recall f1-score support

hate 0.14 0.69 0.23 124

offensive 0.50 0.65 0.57 164

neither 0.92 0.33 0.49 865

accuracy 0.42 1153

macro avg 0.52 0.56 0.43 1153

weighted avg 0.78 0.42 0.47 1153

average eval\_loss: 0.9342295249988293, accuracy: 0.417172593235039

training 2 epoch...

116.3741888999939 seconds used......

2 training finished! average train loss: 1.7084815913722629

total clf loss: 143.46588683128357 total adv loss: 71.77951078116894

evaluating...

Confusion matrix, without normalization

[[ 74 34 16]

[ 30 125 9]

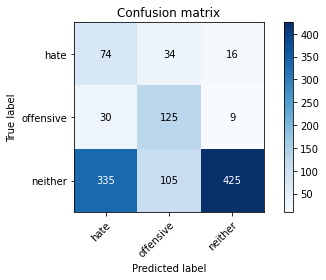
[335 105 425]]

Normalized confusion matrix

[[0.59677419 0.27419355 0.12903226]

[0.18292683 0.76219512 0.05487805]

[0.38728324 0.12138728 0.49132948]]





precision recall f1-score support

hate 0.17 0.60 0.26 124

offensive 0.47 0.76 0.58 164

neither 0.94 0.49 0.65 865

accuracy 0.54 1153

macro avg 0.53 0.62 0.50 1153

weighted avg 0.79 0.54 0.60 1153

average eval\_loss: 0.8839080035686493, accuracy: 0.5411968777103209

training 3 epoch...

116.2640061378479 seconds used......

3 training finished! average train loss: 1.3765982013373148

total clf loss: 122.07322055101395 total adv loss: 54.59763843566179

evaluating...

Confusion matrix, without normalization

[[ 79 23 22]

[ 49 97 18]

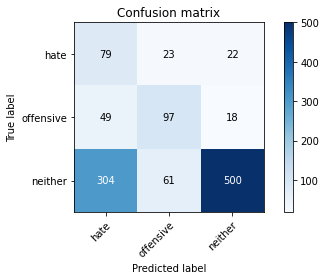
[304 61 500]]

Normalized confusion matrix

[[0.63709677 0.18548387 0.17741935]

[0.29878049 0.59146341 0.1097561 ]

[0.35144509 0.07052023 0.57803468]]





precision recall f1-score support

hate 0.18 0.64 0.28 124

offensive 0.54 0.59 0.56 164

neither 0.93 0.58 0.71 865

accuracy 0.59 1153

macro avg 0.55 0.60 0.52 1153

weighted avg 0.79 0.59 0.64 1153

average eval\_loss: 0.8478814599842861, accuracy: 0.5862966175195143

training 4 epoch...

116.36494016647339 seconds used......

4 training finished! average train loss: 0.9167461118527821

total clf loss: 93.53745231032372 total adv loss: 30.237947307527065

evaluating...

Confusion matrix, without normalization

[[ 64 43 17]

[ 29 123 12]

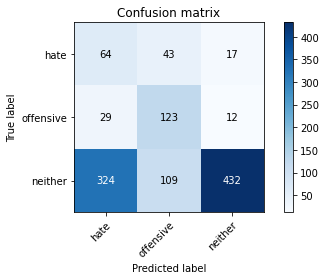
[324 109 432]]

Normalized confusion matrix

[[0.51612903 0.34677419 0.13709677]

[0.17682927 0.75 0.07317073]

[0.37456647 0.12601156 0.49942197]]





precision recall f1-score support

hate 0.15 0.52 0.24 124

offensive 0.45 0.75 0.56 164

neither 0.94 0.50 0.65 865

accuracy 0.54 1153

macro avg 0.51 0.59 0.48 1153

weighted avg 0.78 0.54 0.59 1153

average eval\_loss: 1.0651138753726566, accuracy: 0.5368603642671292

training 5 epoch...

116.2752947807312 seconds used......

5 training finished! average train loss: 0.6852621357178404

total clf loss: 74.41059166193008 total adv loss: 20.356723264441825

evaluating...

Confusion matrix, without normalization

[[ 90 12 22]

[ 46 106 12]

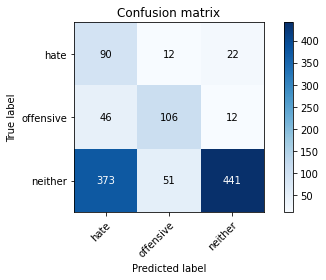
[373 51 441]]

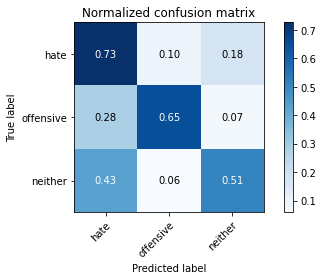
Normalized confusion matrix

[[0.72580645 0.09677419 0.17741935]

[0.2804878 0.64634146 0.07317073]

[0.43121387 0.05895954 0.50982659]]





precision recall f1-score support

hate 0.18 0.73 0.28 124

offensive 0.63 0.65 0.64 164

neither 0.93 0.51 0.66 865

accuracy 0.55 1153

macro avg 0.58 0.63 0.53 1153

weighted avg 0.80 0.55 0.61 1153

average eval\_loss: 1.1142125170806358, accuracy: 0.5524718126626192

training 6 epoch...

116.16670751571655 seconds used......

6 training finished! average train loss: 0.4073069223335811

total clf loss: 51.5881762355566 total adv loss: 8.419693331001326

evaluating...

Confusion matrix, without normalization

[[ 96 17 11]

[ 54 97 13]

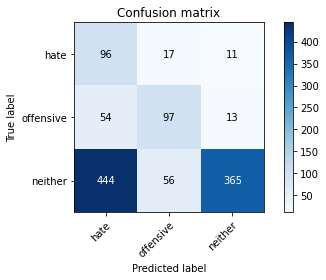
[444 56 365]]

Normalized confusion matrix

[[0.77419355 0.13709677 0.08870968]

[0.32926829 0.59146341 0.07926829]

[0.5132948 0.06473988 0.42196532]]





precision recall f1-score support

hate 0.16 0.77 0.27 124

offensive 0.57 0.59 0.58 164

neither 0.94 0.42 0.58 865

accuracy 0.48 1153

macro avg 0.56 0.60 0.48 1153

weighted avg 0.80 0.48 0.55 1153

average eval\_loss: 1.4452994229464695, accuracy: 0.4839549002601908

training 7 epoch...

116.23192954063416 seconds used......

7 training finished! average train loss: 0.28544432876099435

total clf loss: 35.90554064512253 total adv loss: 6.024553335720157

evaluating...

Confusion matrix, without normalization

[[ 99 21 4]

[ 47 109 8]

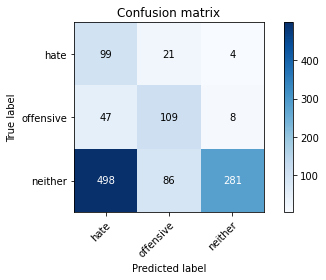
[498 86 281]]

Normalized confusion matrix

[[0.7983871 0.16935484 0.03225806]

[0.28658537 0.66463415 0.04878049]

[0.57572254 0.09942197 0.32485549]]





precision recall f1-score support

hate 0.15 0.80 0.26 124

offensive 0.50 0.66 0.57 164

neither 0.96 0.32 0.49 865

accuracy 0.42 1153

macro avg 0.54 0.60 0.44 1153

weighted avg 0.81 0.42 0.47 1153

average eval\_loss: 1.9644006100194207, accuracy: 0.4241110147441457

training 8 epoch...

116.22144675254822 seconds used......

8 training finished! average train loss: 0.2162904414082212

total clf loss: 25.583018140867352 total adv loss: 5.37688812522174

evaluating...

Confusion matrix, without normalization

[[ 88 22 14]

[ 40 111 13]

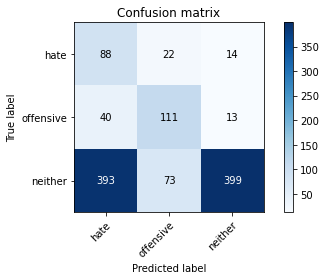
[393 73 399]]

Normalized confusion matrix

[[0.70967742 0.17741935 0.11290323]

[0.24390244 0.67682927 0.07926829]

[0.45433526 0.08439306 0.46127168]]





precision recall f1-score support

hate 0.17 0.71 0.27 124

offensive 0.54 0.68 0.60 164

neither 0.94 0.46 0.62 865

accuracy 0.52 1153

macro avg 0.55 0.62 0.50 1153

weighted avg 0.80 0.52 0.58 1153

average eval\_loss: 1.7747464416355923, accuracy: 0.5186470078057241

training 9 epoch...

116.24342465400696 seconds used......

9 training finished! average train loss: 0.19471233488903159

total clf loss: 18.485676111653447 total adv loss: 7.112998063163104

evaluating...

Confusion matrix, without normalization

[[ 98 19 7]

[ 47 102 15]

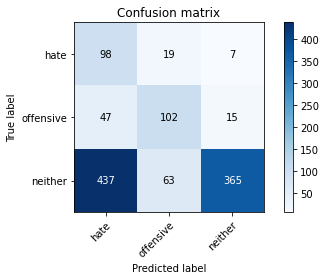
[437 63 365]]

Normalized confusion matrix

[[0.79032258 0.15322581 0.05645161]

[0.28658537 0.62195122 0.09146341]

[0.50520231 0.07283237 0.42196532]]





precision recall f1-score support

hate 0.17 0.79 0.28 124

offensive 0.55 0.62 0.59 164

neither 0.94 0.42 0.58 865

accuracy 0.49 1153

macro avg 0.56 0.61 0.48 1153

weighted avg 0.80 0.49 0.55 1153

average eval\_loss: 2.2230446400313544, accuracy: 0.4900260190806591

training 10 epoch...

116.33077931404114 seconds used......

10 training finished! average train loss: 0.11778788112098396

total clf loss: 13.202143365284428 total adv loss: 3.2931103445043846

evaluating...

Confusion matrix, without normalization

[[ 80 14 30]

[ 37 106 21]

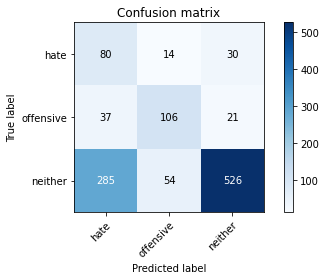
[285 54 526]]

Normalized confusion matrix

[[0.64516129 0.11290323 0.24193548]

[0.22560976 0.64634146 0.12804878]

[0.32947977 0.06242775 0.60809249]]





precision recall f1-score support

hate 0.20 0.65 0.30 124

offensive 0.61 0.65 0.63 164

neither 0.91 0.61 0.73 865

accuracy 0.62 1153

macro avg 0.57 0.63 0.55 1153

weighted avg 0.79 0.62 0.67 1153

average eval\_loss: 1.584141194820404, accuracy: 0.6175195143104943

3 run......

training 1 epoch...

116.19786357879639 seconds used......

1 training finished! average train loss: 2.0741279181979952

total clf loss: 166.9737672805786 total adv loss: 90.73986111581326

evaluating...

Confusion matrix, without normalization

[[107 17 0]

[ 54 108 2]

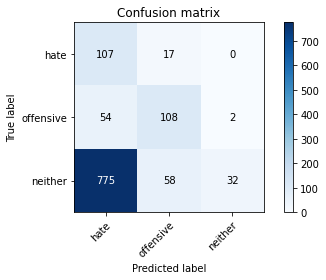
[775 58 32]]

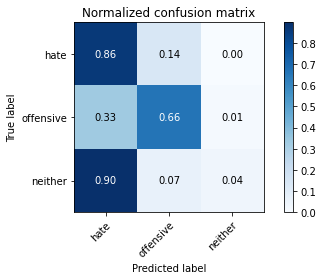
Normalized confusion matrix

[[0.86290323 0.13709677 0. ]

[0.32926829 0.65853659 0.01219512]

[0.89595376 0.06705202 0.03699422]]





precision recall f1-score support

hate 0.11 0.86 0.20 124

offensive 0.59 0.66 0.62 164

neither 0.94 0.04 0.07 865

accuracy 0.21 1153

macro avg 0.55 0.52 0.30 1153

weighted avg 0.80 0.21 0.16 1153

average eval\_loss: 1.0402508943245328, accuracy: 0.2142237640936687

training 2 epoch...

116.20381808280945 seconds used......

2 training finished! average train loss: 1.6823789101271402

total clf loss: 141.73535376787186 total adv loss: 70.45215149223804

evaluating...

Confusion matrix, without normalization

[[ 94 27 3]

[ 39 120 5]

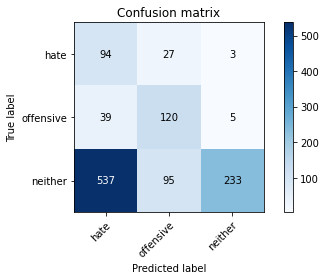
[537 95 233]]

Normalized confusion matrix

[[0.75806452 0.21774194 0.02419355]

[0.23780488 0.73170732 0.0304878 ]

[0.62080925 0.10982659 0.26936416]]





precision recall f1-score support

hate 0.14 0.76 0.24 124

offensive 0.50 0.73 0.59 164

neither 0.97 0.27 0.42 865

accuracy 0.39 1153

macro avg 0.53 0.59 0.42 1153

weighted avg 0.81 0.39 0.43 1153

average eval\_loss: 0.9550889072747066, accuracy: 0.3876843018213356

training 3 epoch...

116.18919491767883 seconds used......

3 training finished! average train loss: 1.3545072529287565

total clf loss: 120.27218395471573 total adv loss: 53.642517518252134

evaluating...

Confusion matrix, without normalization

[[ 89 30 5]

[ 40 117 7]

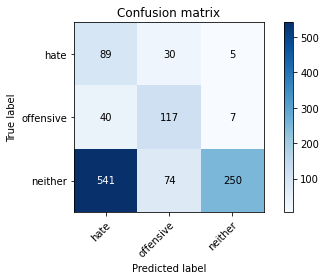
[541 74 250]]

Normalized confusion matrix

[[0.71774194 0.24193548 0.04032258]

[0.24390244 0.71341463 0.04268293]

[0.62543353 0.08554913 0.28901734]]





precision recall f1-score support

hate 0.13 0.72 0.22 124

offensive 0.53 0.71 0.61 164

neither 0.95 0.29 0.44 865

accuracy 0.40 1153

macro avg 0.54 0.57 0.43 1153

weighted avg 0.81 0.40 0.44 1153

average eval\_loss: 1.0346107154056943, accuracy: 0.3954900260190807

training 4 epoch...

116.20402431488037 seconds used......

4 training finished! average train loss: 0.9169943689235619

total clf loss: 91.50614909827709 total adv loss: 31.274451918900013

evaluating...

Confusion matrix, without normalization

[[ 79 34 11]

[ 31 123 10]

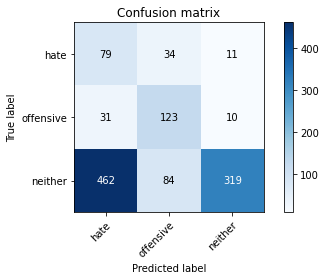
[462 84 319]]

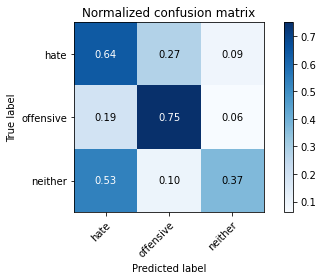
Normalized confusion matrix

[[0.63709677 0.27419355 0.08870968]

[0.18902439 0.75 0.06097561]

[0.53410405 0.09710983 0.36878613]]





precision recall f1-score support

hate 0.14 0.64 0.23 124

offensive 0.51 0.75 0.61 164

neither 0.94 0.37 0.53 865

accuracy 0.45 1153

macro avg 0.53 0.59 0.45 1153

weighted avg 0.79 0.45 0.51 1153

average eval\_loss: 1.1092092816171975, accuracy: 0.4518647007805724

training 5 epoch...

116.20147347450256 seconds used......

5 training finished! average train loss: 0.5606803447895107

total clf loss: 66.00696135312319 total adv loss: 14.093668093672022

evaluating...

Confusion matrix, without normalization

[[ 99 15 10]

[ 54 96 14]

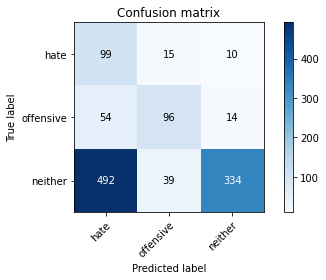
[492 39 334]]

Normalized confusion matrix

[[0.7983871 0.12096774 0.08064516]

[0.32926829 0.58536585 0.08536585]

[0.56878613 0.04508671 0.38612717]]





precision recall f1-score support

hate 0.15 0.80 0.26 124

offensive 0.64 0.59 0.61 164

neither 0.93 0.39 0.55 865

accuracy 0.46 1153

macro avg 0.58 0.59 0.47 1153

weighted avg 0.81 0.46 0.52 1153

average eval\_loss: 1.3064865788509106, accuracy: 0.4588031222896791

training 6 epoch...

116.23531723022461 seconds used......

6 training finished! average train loss: 0.41687386190252645

total clf loss: 49.79031276702881 total adv loss: 10.122248088089691

evaluating...

Confusion matrix, without normalization

[[ 96 19 9]

[ 46 105 13]

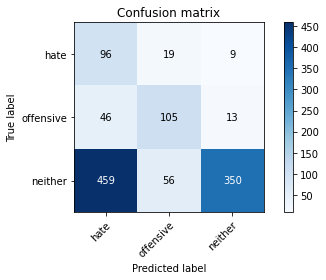
[459 56 350]]

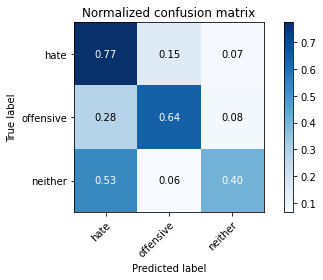
Normalized confusion matrix

[[0.77419355 0.15322581 0.07258065]

[0.2804878 0.6402439 0.07926829]

[0.53063584 0.06473988 0.40462428]]





precision recall f1-score support

hate 0.16 0.77 0.26 124

offensive 0.58 0.64 0.61 164

neither 0.94 0.40 0.57 865

accuracy 0.48 1153

macro avg 0.56 0.61 0.48 1153

weighted avg 0.81 0.48 0.54 1153

average eval\_loss: 1.43878118436912, accuracy: 0.4778837814397225

training 7 epoch...

116.26948833465576 seconds used......

7 training finished! average train loss: 0.26278564901578993

total clf loss: 31.715209253132343 total adv loss: 6.216389767439978

evaluating...

Confusion matrix, without normalization

[[ 99 17 8]

[ 52 104 8]

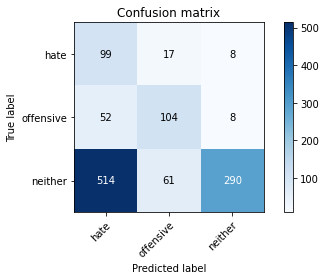
[514 61 290]]

Normalized confusion matrix

[[0.7983871 0.13709677 0.06451613]

[0.31707317 0.63414634 0.04878049]

[0.59421965 0.07052023 0.33526012]]





precision recall f1-score support

hate 0.15 0.80 0.25 124

offensive 0.57 0.63 0.60 164

neither 0.95 0.34 0.50 865

accuracy 0.43 1153

macro avg 0.56 0.59 0.45 1153

weighted avg 0.81 0.43 0.48 1153

average eval\_loss: 1.7533940035721352, accuracy: 0.42758022549869906

training 8 epoch...

116.27413702011108 seconds used......

8 training finished! average train loss: 0.1964714768608766

total clf loss: 23.344626562669873 total adv loss: 4.831290752572386

evaluating...

Confusion matrix, without normalization

[[ 90 24 10]

[ 38 111 15]

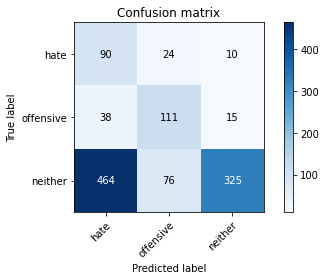
[464 76 325]]

Normalized confusion matrix

[[0.72580645 0.19354839 0.08064516]

[0.23170732 0.67682927 0.09146341]

[0.53641618 0.08786127 0.37572254]]





precision recall f1-score support

hate 0.15 0.73 0.25 124

offensive 0.53 0.68 0.59 164

neither 0.93 0.38 0.53 865

accuracy 0.46 1153

macro avg 0.54 0.59 0.46 1153

weighted avg 0.79 0.46 0.51 1153

average eval\_loss: 1.9443379918049122, accuracy: 0.4562012142237641

training 9 epoch...

116.29113054275513 seconds used......

9 training finished! average train loss: 0.12910537950838694

total clf loss: 14.676830248441547 total adv loss: 3.506436750290959

evaluating...

Confusion matrix, without normalization

[[ 88 22 14]

[ 43 107 14]

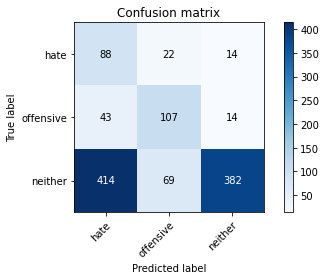
[414 69 382]]

Normalized confusion matrix

[[0.70967742 0.17741935 0.11290323]

[0.26219512 0.65243902 0.08536585]

[0.47861272 0.07976879 0.4416185 ]]





precision recall f1-score support

hate 0.16 0.71 0.26 124

offensive 0.54 0.65 0.59 164

neither 0.93 0.44 0.60 865

accuracy 0.50 1153

macro avg 0.54 0.60 0.48 1153

weighted avg 0.79 0.50 0.56 1153

average eval\_loss: 1.933607294641692, accuracy: 0.5004336513443192

training 10 epoch...

116.28842258453369 seconds used......

10 training finished! average train loss: 0.11961192213874754

total clf loss: 13.017373267095536 total adv loss: 3.538714857873856

evaluating...

Confusion matrix, without normalization

[[ 91 22 11]

[ 43 109 12]

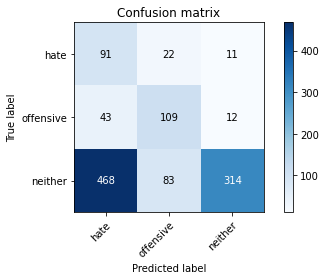
[468 83 314]]

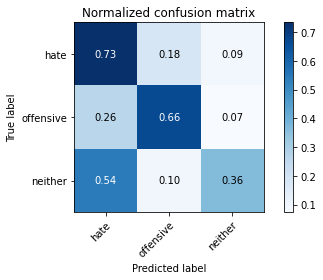
Normalized confusion matrix

[[0.73387097 0.17741935 0.08870968]

[0.26219512 0.66463415 0.07317073]

[0.54104046 0.09595376 0.36300578]]





precision recall f1-score support

hate 0.15 0.73 0.25 124

offensive 0.51 0.66 0.58 164

neither 0.93 0.36 0.52 865

accuracy 0.45 1153

macro avg 0.53 0.59 0.45 1153

weighted avg 0.79 0.45 0.50 1153

average eval\_loss: 2.4227566349095313, accuracy: 0.4457935819601041

4 run......

training 1 epoch...

116.33519268035889 seconds used......

1 training finished! average train loss: 2.0731182112580253

total clf loss: 166.89376533031464 total adv loss: 90.69504699110985

evaluating...

Confusion matrix, without normalization

[[ 93 29 2]

[ 46 115 3]

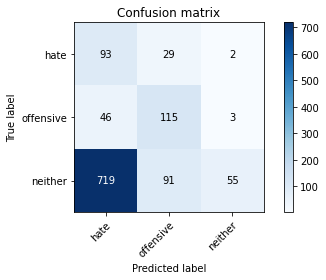
[719 91 55]]

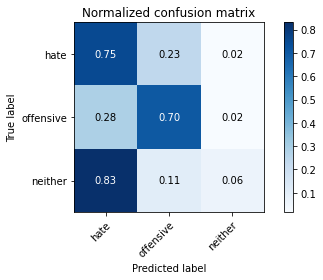
Normalized confusion matrix

[[0.75 0.23387097 0.01612903]

[0.2804878 0.70121951 0.01829268]

[0.83121387 0.10520231 0.06358382]]





precision recall f1-score support

hate 0.11 0.75 0.19 124

offensive 0.49 0.70 0.58 164

neither 0.92 0.06 0.12 865

accuracy 0.23 1153

macro avg 0.50 0.50 0.29 1153

weighted avg 0.77 0.23 0.19 1153

average eval\_loss: 1.0733452089901627, accuracy: 0.22810060711188204

training 2 epoch...

116.24849104881287 seconds used......

2 training finished! average train loss: 1.6976196084703719

total clf loss: 142.67900878190994 total adv loss: 71.26054288446903

evaluating...

Confusion matrix, without normalization

[[107 16 1]

[ 55 107 2]

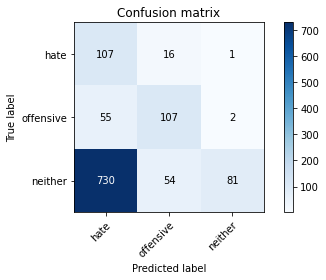
[730 54 81]]

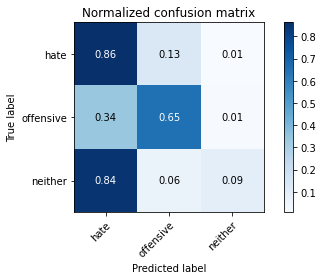
Normalized confusion matrix

[[0.86290323 0.12903226 0.00806452]

[0.33536585 0.65243902 0.01219512]

[0.84393064 0.06242775 0.09364162]]





precision recall f1-score support

hate 0.12 0.86 0.21 124

offensive 0.60 0.65 0.63 164

neither 0.96 0.09 0.17 865

accuracy 0.26 1153

macro avg 0.56 0.54 0.34 1153

weighted avg 0.82 0.26 0.24 1153

average eval\_loss: 1.0746820507378414, accuracy: 0.25585429314830876

training 3 epoch...

116.24441194534302 seconds used......

3 training finished! average train loss: 1.3675945237988518

total clf loss: 121.47428405284882 total adv loss: 54.14079800993204

evaluating...

Confusion matrix, without normalization

[[100 18 6]

[ 55 101 8]

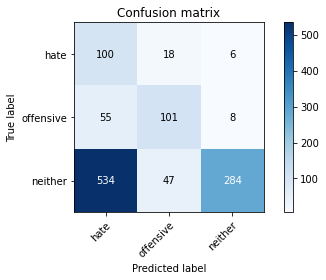
[534 47 284]]

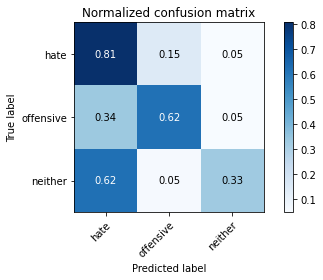
Normalized confusion matrix

[[0.80645161 0.14516129 0.0483871 ]

[0.33536585 0.61585366 0.04878049]

[0.61734104 0.05433526 0.3283237 ]]





precision recall f1-score support

hate 0.15 0.81 0.25 124

offensive 0.61 0.62 0.61 164

neither 0.95 0.33 0.49 865

accuracy 0.42 1153

macro avg 0.57 0.58 0.45 1153

weighted avg 0.82 0.42 0.48 1153

average eval\_loss: 0.9678122226534218, accuracy: 0.4206418039895924

training 4 epoch...

116.29190063476562 seconds used......

4 training finished! average train loss: 0.9487699307501316

total clf loss: 94.59123522043228 total adv loss: 32.40105628501624

evaluating...

Confusion matrix, without normalization

[[ 91 27 6]

[ 47 109 8]

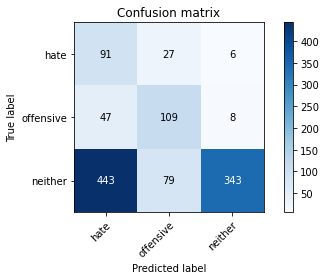
[443 79 343]]

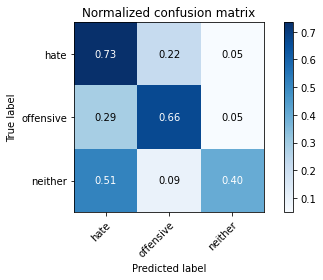
Normalized confusion matrix

[[0.73387097 0.21774194 0.0483871 ]

[0.28658537 0.66463415 0.04878049]

[0.51213873 0.09132948 0.39653179]]





precision recall f1-score support

hate 0.16 0.73 0.26 124

offensive 0.51 0.66 0.58 164

neither 0.96 0.40 0.56 865

accuracy 0.47 1153

macro avg 0.54 0.60 0.46 1153

weighted avg 0.81 0.47 0.53 1153

average eval\_loss: 1.0721149547346707, accuracy: 0.47094535993061576

training 5 epoch...

116.21817588806152 seconds used......

5 training finished! average train loss: 0.585797514145573

total clf loss: 69.05310901999474 total adv loss: 14.680436588125303

evaluating...

Confusion matrix, without normalization

[[ 79 36 9]

[ 41 114 9]

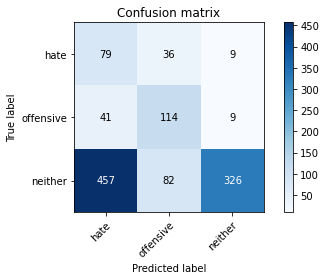
[457 82 326]]

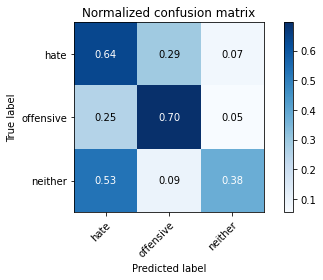
Normalized confusion matrix

[[0.63709677 0.29032258 0.07258065]

[0.25 0.69512195 0.05487805]

[0.5283237 0.09479769 0.37687861]]





precision recall f1-score support

hate 0.14 0.64 0.23 124

offensive 0.49 0.70 0.58 164

neither 0.95 0.38 0.54 865

accuracy 0.45 1153

macro avg 0.53 0.57 0.45 1153

weighted avg 0.80 0.45 0.51 1153

average eval\_loss: 1.2458948955453675, accuracy: 0.45013009540329574

training 6 epoch...

116.32187342643738 seconds used......

6 training finished! average train loss: 0.40600431178297314

total clf loss: 50.49225902557373 total adv loss: 8.858232598402537

evaluating...

Confusion matrix, without normalization

[[ 78 32 14]

[ 37 118 9]

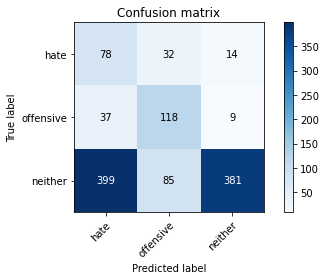
[399 85 381]]

Normalized confusion matrix

[[0.62903226 0.25806452 0.11290323]

[0.22560976 0.7195122 0.05487805]

[0.46127168 0.0982659 0.44046243]]





precision recall f1-score support

hate 0.15 0.63 0.24 124

offensive 0.50 0.72 0.59 164

neither 0.94 0.44 0.60 865

accuracy 0.50 1153

macro avg 0.53 0.60 0.48 1153

weighted avg 0.80 0.50 0.56 1153

average eval\_loss: 1.428349402444116, accuracy: 0.5004336513443192

training 7 epoch...

116.24252963066101 seconds used......

7 training finished! average train loss: 0.32377390940451906

total clf loss: 37.35599763691425 total adv loss: 8.519009623734746

evaluating...

Confusion matrix, without normalization

[[ 74 42 8]

[ 37 115 12]

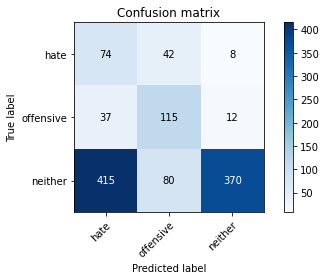
[415 80 370]]

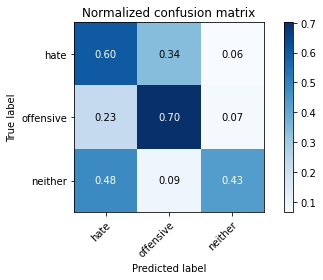
Normalized confusion matrix

[[0.59677419 0.33870968 0.06451613]

[0.22560976 0.70121951 0.07317073]

[0.47976879 0.09248555 0.42774566]]





precision recall f1-score support

hate 0.14 0.60 0.23 124

offensive 0.49 0.70 0.57 164

neither 0.95 0.43 0.59 865

accuracy 0.48 1153

macro avg 0.52 0.58 0.46 1153

weighted avg 0.80 0.48 0.55 1153

average eval\_loss: 1.568725753447105, accuracy: 0.48482220294882916

training 8 epoch...

116.2019464969635 seconds used......

8 training finished! average train loss: 0.20991840314430496

total clf loss: 23.902659635990858 total adv loss: 5.681815999225364

evaluating...

Confusion matrix, without normalization

[[ 80 25 19]

[ 40 105 19]

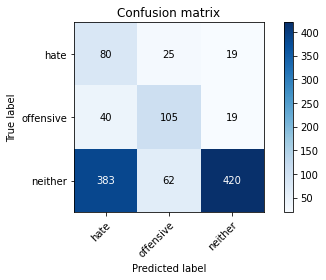
[383 62 420]]

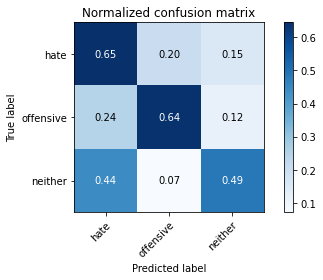
Normalized confusion matrix

[[0.64516129 0.2016129 0.15322581]

[0.24390244 0.6402439 0.11585366]

[0.44277457 0.0716763 0.48554913]]





precision recall f1-score support

hate 0.16 0.65 0.26 124

offensive 0.55 0.64 0.59 164

neither 0.92 0.49 0.63 865

accuracy 0.52 1153

macro avg 0.54 0.59 0.49 1153

weighted avg 0.78 0.52 0.59 1153

average eval\_loss: 1.6193981797530734, accuracy: 0.5247181266261925

training 9 epoch...

116.3926146030426 seconds used......

9 training finished! average train loss: 0.13825913531417472

total clf loss: 16.581547713838518 total adv loss: 3.3229935417475644

evaluating...

Confusion matrix, without normalization

[[ 75 37 12]

[ 31 118 15]

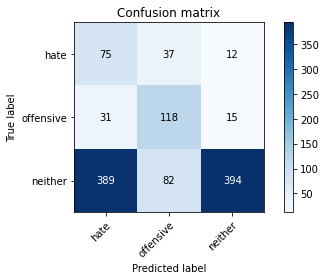
[389 82 394]]

Normalized confusion matrix

[[0.60483871 0.2983871 0.09677419]

[0.18902439 0.7195122 0.09146341]

[0.44971098 0.09479769 0.45549133]]





precision recall f1-score support

hate 0.15 0.60 0.24 124

offensive 0.50 0.72 0.59 164

neither 0.94 0.46 0.61 865

accuracy 0.51 1153

macro avg 0.53 0.59 0.48 1153

weighted avg 0.79 0.51 0.57 1153

average eval\_loss: 1.888856745999435, accuracy: 0.5091066782307025

training 10 epoch...

116.35905814170837 seconds used......

10 training finished! average train loss: 0.11736234599846371

total clf loss: 12.619498886168003 total adv loss: 3.54868762247861

evaluating...

Confusion matrix, without normalization

[[ 83 36 5]

[ 29 122 13]

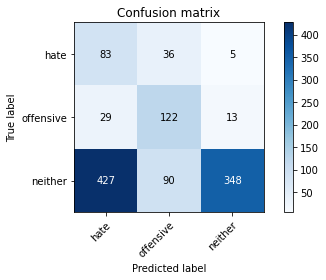
[427 90 348]]

Normalized confusion matrix

[[0.66935484 0.29032258 0.04032258]

[0.17682927 0.74390244 0.07926829]

[0.49364162 0.10404624 0.40231214]]





precision recall f1-score support

hate 0.15 0.67 0.25 124

offensive 0.49 0.74 0.59 164

neither 0.95 0.40 0.57 865

accuracy 0.48 1153

macro avg 0.53 0.61 0.47 1153

weighted avg 0.80 0.48 0.54 1153

average eval\_loss: 2.329729114113183, accuracy: 0.47961838681699914

5 run......

training 1 epoch...

116.35351300239563 seconds used......

1 training finished! average train loss: 2.0534131122486934

total clf loss: 165.72689163684845 total adv loss: 89.623255610466

evaluating...

Confusion matrix, without normalization

[[109 15 0]

[ 61 101 2]

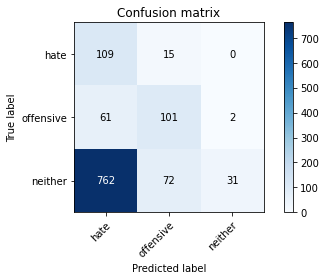
[762 72 31]]

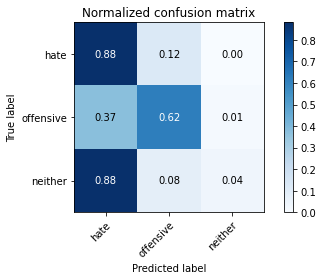
Normalized confusion matrix

[[0.87903226 0.12096774 0. ]

[0.37195122 0.61585366 0.01219512]

[0.88092486 0.08323699 0.03583815]]





precision recall f1-score support

hate 0.12 0.88 0.21 124

offensive 0.54 0.62 0.57 164

neither 0.94 0.04 0.07 865

accuracy 0.21 1153

macro avg 0.53 0.51 0.28 1153

weighted avg 0.79 0.21 0.16 1153

average eval\_loss: 1.0712298298704213, accuracy: 0.20901994796183868

training 2 epoch...

116.28941702842712 seconds used......

2 training finished! average train loss: 1.6955619362138568

total clf loss: 142.27962225675583 total adv loss: 71.28739188611507

evaluating...

Confusion matrix, without normalization

[[ 93 21 10]

[ 38 113 13]

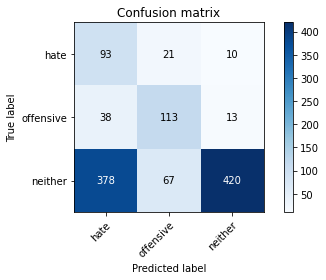
[378 67 420]]

Normalized confusion matrix

[[0.75 0.16935484 0.08064516]

[0.23170732 0.68902439 0.07926829]

[0.43699422 0.07745665 0.48554913]]





precision recall f1-score support

hate 0.18 0.75 0.29 124

offensive 0.56 0.69 0.62 164

neither 0.95 0.49 0.64 865

accuracy 0.54 1153

macro avg 0.56 0.64 0.52 1153

weighted avg 0.81 0.54 0.60 1153

average eval\_loss: 0.8694339600102655, accuracy: 0.5429314830875975

training 3 epoch...

116.42501211166382 seconds used......

3 training finished! average train loss: 1.3523063241016298

total clf loss: 120.77563980221748 total adv loss: 53.2059114202857

evaluating...

Confusion matrix, without normalization

[[ 92 25 7]

[ 46 108 10]

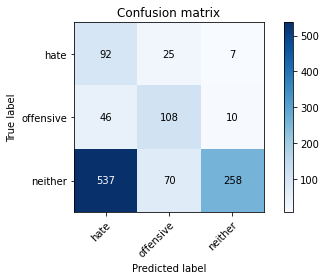
[537 70 258]]

Normalized confusion matrix

[[0.74193548 0.2016129 0.05645161]

[0.2804878 0.65853659 0.06097561]

[0.62080925 0.08092486 0.2982659 ]]





precision recall f1-score support

hate 0.14 0.74 0.23 124

offensive 0.53 0.66 0.59 164

neither 0.94 0.30 0.45 865

accuracy 0.40 1153

macro avg 0.54 0.57 0.42 1153

weighted avg 0.79 0.40 0.45 1153

average eval\_loss: 1.039973390513453, accuracy: 0.39722463139635733

training 4 epoch...

116.2628333568573 seconds used......

4 training finished! average train loss: 0.9244345156919389

total clf loss: 93.50735777616501 total adv loss: 30.898820557631552

evaluating...

Confusion matrix, without normalization

[[ 74 37 13]

[ 35 113 16]

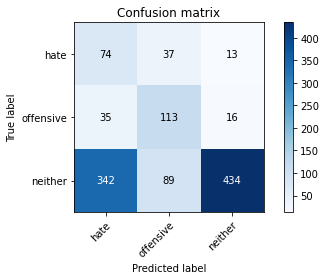
[342 89 434]]

Normalized confusion matrix

[[0.59677419 0.2983871 0.10483871]

[0.21341463 0.68902439 0.09756098]

[0.39537572 0.10289017 0.5017341 ]]





precision recall f1-score support

hate 0.16 0.60 0.26 124

offensive 0.47 0.69 0.56 164

neither 0.94 0.50 0.65 865

accuracy 0.54 1153

macro avg 0.52 0.60 0.49 1153

weighted avg 0.79 0.54 0.60 1153

average eval\_loss: 0.995436625234012, accuracy: 0.5385949696444059

training 5 epoch...

116.26485753059387 seconds used......

5 training finished! average train loss: 0.6054773790140947

total clf loss: 69.3612547069788 total adv loss: 16.179472405347042

evaluating...

Confusion matrix, without normalization

[[ 78 40 6]

[ 39 115 10]

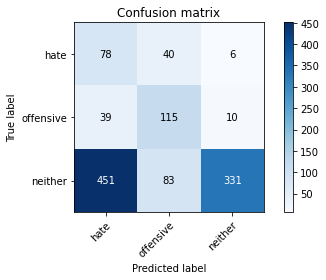
[451 83 331]]

Normalized confusion matrix

[[0.62903226 0.32258065 0.0483871 ]

[0.23780488 0.70121951 0.06097561]

[0.52138728 0.09595376 0.38265896]]





precision recall f1-score support

hate 0.14 0.63 0.23 124

offensive 0.48 0.70 0.57 164

neither 0.95 0.38 0.55 865

accuracy 0.45 1153

macro avg 0.52 0.57 0.45 1153

weighted avg 0.80 0.45 0.52 1153

average eval\_loss: 1.2835162251160062, accuracy: 0.45446660884648743

training 6 epoch...

116.26817226409912 seconds used......

6 training finished! average train loss: 0.4357503015725386

total clf loss: 54.36595295369625 total adv loss: 9.420049018983264

evaluating...

Confusion matrix, without normalization

[[ 89 27 8]

[ 37 113 14]

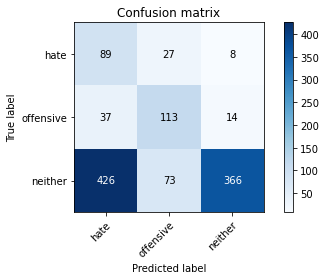
[426 73 366]]

Normalized confusion matrix

[[0.71774194 0.21774194 0.06451613]

[0.22560976 0.68902439 0.08536585]

[0.49248555 0.08439306 0.42312139]]





precision recall f1-score support

hate 0.16 0.72 0.26 124

offensive 0.53 0.69 0.60 164

neither 0.94 0.42 0.58 865

accuracy 0.49 1153

macro avg 0.55 0.61 0.48 1153

weighted avg 0.80 0.49 0.55 1153

average eval\_loss: 1.3647626258175949, accuracy: 0.49262792714657416

training 7 epoch...

116.19051718711853 seconds used......

7 training finished! average train loss: 0.3022861758779202

total clf loss: 38.30042069405317 total adv loss: 6.241828365426045

evaluating...

Confusion matrix, without normalization

[[ 90 28 6]

[ 39 114 11]

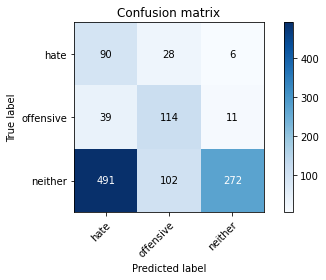
[491 102 272]]

Normalized confusion matrix

[[0.72580645 0.22580645 0.0483871 ]

[0.23780488 0.69512195 0.06707317]

[0.56763006 0.11791908 0.31445087]]





precision recall f1-score support

hate 0.15 0.73 0.24 124

offensive 0.47 0.70 0.56 164

neither 0.94 0.31 0.47 865

accuracy 0.41 1153

macro avg 0.52 0.58 0.42 1153

weighted avg 0.79 0.41 0.46 1153

average eval\_loss: 1.9616232473274757, accuracy: 0.41283607979184733

training 8 epoch...

116.3031313419342 seconds used......

8 training finished! average train loss: 0.2501811549688379

total clf loss: 29.10068354383111 total adv loss: 6.464875234312785

evaluating...

Confusion matrix, without normalization

[[ 94 23 7]

[ 40 114 10]

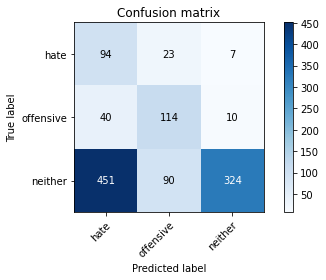
[451 90 324]]

Normalized confusion matrix

[[0.75806452 0.18548387 0.05645161]

[0.24390244 0.69512195 0.06097561]

[0.52138728 0.10404624 0.37456647]]





precision recall f1-score support

hate 0.16 0.76 0.27 124

offensive 0.50 0.70 0.58 164

neither 0.95 0.37 0.54 865

accuracy 0.46 1153

macro avg 0.54 0.61 0.46 1153

weighted avg 0.80 0.46 0.51 1153

average eval\_loss: 2.012853625519522, accuracy: 0.4614050303555941

training 9 epoch...

116.26242566108704 seconds used......

9 training finished! average train loss: 0.14615241008522434

total clf loss: 18.035313645377755 total adv loss: 3.259145649666607

evaluating...

Confusion matrix, without normalization

[[ 85 21 18]

[ 39 107 18]

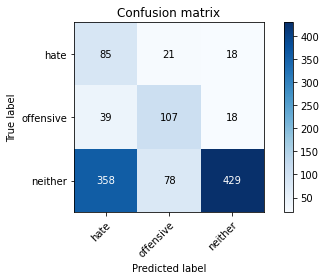
[358 78 429]]

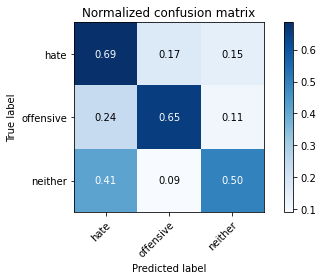
Normalized confusion matrix

[[0.68548387 0.16935484 0.14516129]

[0.23780488 0.65243902 0.1097561 ]

[0.41387283 0.09017341 0.49595376]]





precision recall f1-score support

hate 0.18 0.69 0.28 124

offensive 0.52 0.65 0.58 164

neither 0.92 0.50 0.65 865

accuracy 0.54 1153

macro avg 0.54 0.61 0.50 1153

weighted avg 0.78 0.54 0.60 1153

average eval\_loss: 1.7134061404343308, accuracy: 0.5385949696444059

training 10 epoch...

116.26373362541199 seconds used......

10 training finished! average train loss: 0.15719879879837945

total clf loss: 15.673081059008837 total adv loss: 5.368158613855485

evaluating...

Confusion matrix, without normalization

[[103 10 11]

[ 56 96 12]

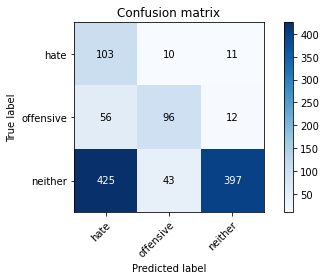
[425 43 397]]

Normalized confusion matrix

[[0.83064516 0.08064516 0.08870968]

[0.34146341 0.58536585 0.07317073]

[0.49132948 0.04971098 0.45895954]]





precision recall f1-score support

hate 0.18 0.83 0.29 124

offensive 0.64 0.59 0.61 164

neither 0.95 0.46 0.62 865

accuracy 0.52 1153

macro avg 0.59 0.62 0.51 1153

weighted avg 0.82 0.52 0.58 1153

average eval\_loss: 2.0572111616874564, accuracy: 0.5169124024284475