1 run

alpha =0.05

training 1 epoch...

45.11322736740112 seconds used......

1 training finished! average train loss: 0.6640162399114921

total clf loss: 116.76753947138786 total adv loss: 64.415249183774

evaluating...

Confusion matrix, without normalization

[[ 217 45 32]

[ 694 2961 166]

[ 30 9 803]]

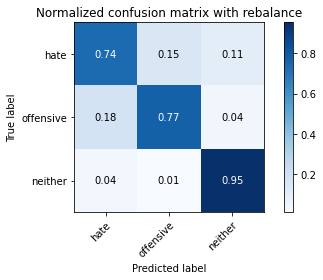
Normalized confusion matrix

[[0.73809524 0.15306122 0.10884354]

[0.18162785 0.77492803 0.04344412]

[0.03562945 0.01068884 0.95368171]]





precision recall f1-score support

hate 0.23 0.74 0.35 294

offensive 0.98 0.77 0.87 3821

neither 0.80 0.95 0.87 842

accuracy 0.80 4957

macro avg 0.67 0.82 0.70 4957

weighted avg 0.91 0.80 0.84 4957

average eval\_loss: 0.5054485385696734, accuracy: 0.8031067177728465

training 2 epoch...

44.531978130340576 seconds used......

2 training finished! average train loss: 0.38087432822811673

total clf loss: 67.49638611078262 total adv loss: 47.337521214038134

evaluating...

Confusion matrix, without normalization

[[ 226 46 22]

[ 680 3022 119]

[ 35 21 786]]

Normalized confusion matrix

[[0.76870748 0.15646259 0.07482993]

[0.17796388 0.79089244 0.03114368]

[0.0415677 0.02494062 0.93349169]]





precision recall f1-score support

hate 0.24 0.77 0.37 294

offensive 0.98 0.79 0.87 3821

neither 0.85 0.93 0.89 842

accuracy 0.81 4957

macro avg 0.69 0.83 0.71 4957

weighted avg 0.91 0.81 0.85 4957

average eval\_loss: 0.4321393030545404, accuracy: 0.8137986685495259

training 3 epoch...

44.585997343063354 seconds used......

3 training finished! average train loss: 0.2523659046151136

total clf loss: 45.49461840838194 total adv loss: 46.80097181722522

evaluating...

Confusion matrix, without normalization

[[ 203 67 24]

[ 511 3190 120]

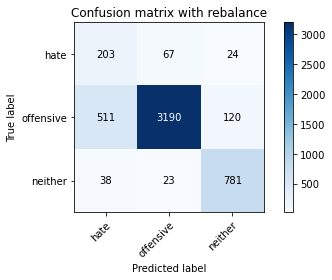
[ 38 23 781]]

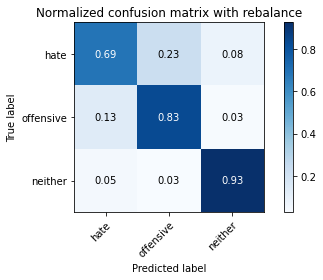
Normalized confusion matrix

[[0.69047619 0.22789116 0.08163265]

[0.13373462 0.83485998 0.03140539]

[0.04513064 0.02731591 0.92755344]]





precision recall f1-score support

hate 0.27 0.69 0.39 294

offensive 0.97 0.83 0.90 3821

neither 0.84 0.93 0.88 842

accuracy 0.84 4957

macro avg 0.70 0.82 0.72 4957

weighted avg 0.91 0.84 0.87 4957

average eval\_loss: 0.3930319487718084, accuracy: 0.8420415573935849

training 4 epoch...

44.56182909011841 seconds used......

4 training finished! average train loss: 0.13746030447862997

total clf loss: 25.923252683132887 total adv loss: 48.35081183165312

evaluating...

Confusion matrix, without normalization

[[ 194 79 21]

[ 452 3264 105]

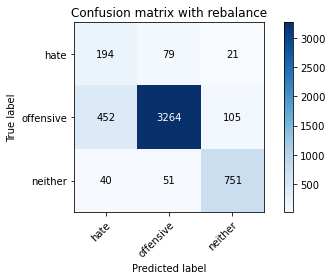
[ 40 51 751]]

Normalized confusion matrix

[[0.65986395 0.26870748 0.07142857]

[0.11829364 0.85422664 0.02747972]

[0.04750594 0.06057007 0.89192399]]





precision recall f1-score support

hate 0.28 0.66 0.40 294

offensive 0.96 0.85 0.90 3821

neither 0.86 0.89 0.87 842

accuracy 0.85 4957

macro avg 0.70 0.80 0.72 4957

weighted avg 0.90 0.85 0.87 4957

average eval\_loss: 0.4657518386074732, accuracy: 0.8491022796045996

training 5 epoch...

44.54462432861328 seconds used......

5 training finished! average train loss: 0.06918080948968554

total clf loss: 14.241284872405231 total adv loss: 48.22732912003994

evaluating...

Confusion matrix, without normalization

[[ 211 45 38]

[ 819 2832 170]

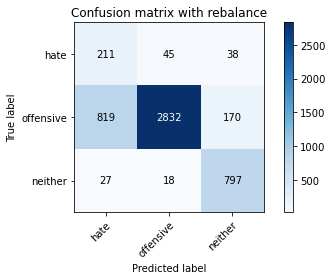
[ 27 18 797]]

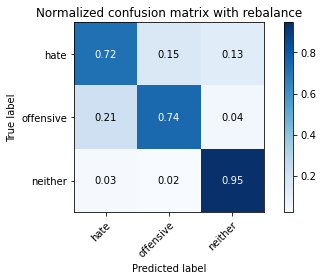
Normalized confusion matrix

[[0.71768707 0.15306122 0.1292517 ]

[0.2143418 0.74116723 0.04449097]

[0.03206651 0.02137767 0.94655582]]





precision recall f1-score support

hate 0.20 0.72 0.31 294

offensive 0.98 0.74 0.84 3821

neither 0.79 0.95 0.86 842

accuracy 0.77 4957

macro avg 0.66 0.80 0.67 4957

weighted avg 0.90 0.77 0.82 4957

average eval\_loss: 0.8647092720343461, accuracy: 0.7746620940084729

training 6 epoch...

44.52416253089905 seconds used......

6 training finished! average train loss: 0.03851997426286209

total clf loss: 9.047913938062266 total adv loss: 49.21996545791626

evaluating...

Confusion matrix, without normalization

[[ 177 92 25]

[ 452 3237 132]

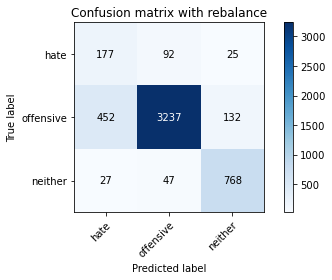
[ 27 47 768]]

Normalized confusion matrix

[[0.60204082 0.31292517 0.08503401]

[0.11829364 0.84716043 0.03454593]

[0.03206651 0.05581948 0.91211401]]





precision recall f1-score support

hate 0.27 0.60 0.37 294

offensive 0.96 0.85 0.90 3821

neither 0.83 0.91 0.87 842

accuracy 0.84 4957

macro avg 0.69 0.79 0.71 4957

weighted avg 0.90 0.84 0.86 4957

average eval\_loss: 0.5415283073351959, accuracy: 0.8436554367561024

training 7 epoch...

44.563153982162476 seconds used......

7 training finished! average train loss: 0.023226251006017478

total clf loss: 6.476115472149104 total adv loss: 50.08853031694889

evaluating...

Confusion matrix, without normalization

[[ 241 35 18]

[ 926 2773 122]

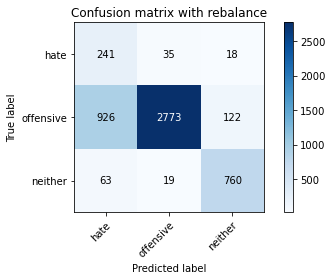
[ 63 19 760]]

Normalized confusion matrix

[[0.81972789 0.11904762 0.06122449]

[0.24234494 0.72572625 0.03192881]

[0.07482185 0.02256532 0.90261283]]





precision recall f1-score support

hate 0.20 0.82 0.32 294

offensive 0.98 0.73 0.83 3821

neither 0.84 0.90 0.87 842

accuracy 0.76 4957

macro avg 0.67 0.82 0.67 4957

weighted avg 0.91 0.76 0.81 4957

average eval\_loss: 1.141617839245291, accuracy: 0.7613475892677022

training 8 epoch...

44.54596781730652 seconds used......

8 training finished! average train loss: 0.00363198927554645

total clf loss: 3.165116558666341 total adv loss: 50.880926840007305

evaluating...

Confusion matrix, without normalization

[[ 199 78 17]

[ 536 3181 104]

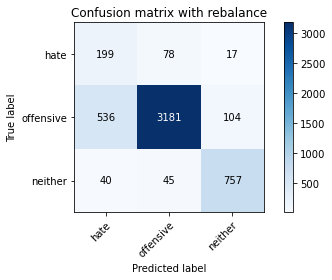
[ 40 45 757]]

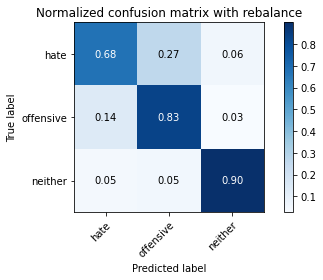
Normalized confusion matrix

[[0.67687075 0.26530612 0.05782313]

[0.14027741 0.83250458 0.02721801]

[0.04750594 0.05344418 0.89904988]]





precision recall f1-score support

hate 0.26 0.68 0.37 294

offensive 0.96 0.83 0.89 3821

neither 0.86 0.90 0.88 842

accuracy 0.83 4957

macro avg 0.69 0.80 0.72 4957

weighted avg 0.90 0.83 0.86 4957

average eval\_loss: 0.7792697569092261, accuracy: 0.8345773653419407

training 9 epoch...

44.537811279296875 seconds used......

9 training finished! average train loss: 0.0005272047128528357

total clf loss: 2.7200650150480215 total adv loss: 52.59825924783945

evaluating...

Confusion matrix, without normalization

[[ 227 56 11]

[ 750 2967 104]

[ 89 27 726]]

Normalized confusion matrix

[[0.77210884 0.19047619 0.03741497]

[0.1962837 0.7764983 0.02721801]

[0.10570071 0.03206651 0.86223278]]





precision recall f1-score support

hate 0.21 0.77 0.33 294

offensive 0.97 0.78 0.86 3821

neither 0.86 0.86 0.86 842

accuracy 0.79 4957

macro avg 0.68 0.80 0.69 4957

weighted avg 0.91 0.79 0.83 4957

average eval\_loss: 1.0781237150799678, accuracy: 0.7908008876336494

training 10 epoch...

44.41501808166504 seconds used......

10 training finished! average train loss: 0.001469938280551057

total clf loss: 2.9542617112165317 total adv loss: 54.05804510042071

evaluating...

Confusion matrix, without normalization

[[ 212 63 19]

[ 648 3053 120]

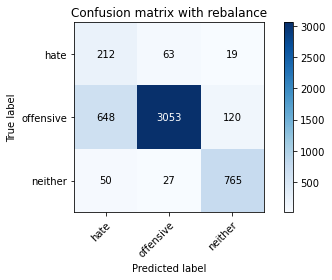
[ 50 27 765]]

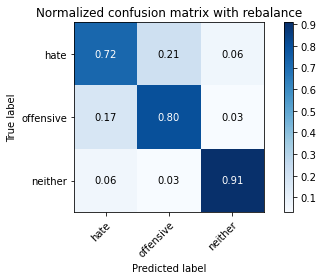
Normalized confusion matrix

[[0.72108844 0.21428571 0.06462585]

[0.16958911 0.7990055 0.03140539]

[0.05938242 0.03206651 0.90855107]]





precision recall f1-score support

hate 0.23 0.72 0.35 294

offensive 0.97 0.80 0.88 3821

neither 0.85 0.91 0.88 842

accuracy 0.81 4957

macro avg 0.68 0.81 0.70 4957

weighted avg 0.91 0.81 0.85 4957

average eval\_loss: 0.9700720950072976, accuracy: 0.8129917288682671

2 run

training 1 epoch...

44.424925327301025 seconds used......

1 training finished! average train loss: 0.6453680063200276

total clf loss: 113.42034758627415 total adv loss: 61.24836687743664

evaluating...

Confusion matrix, without normalization

[[ 161 82 51]

[ 227 3353 241]

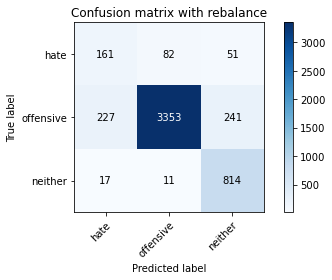
[ 17 11 814]]

Normalized confusion matrix

[[0.54761905 0.27891156 0.17346939]

[0.05940853 0.87751897 0.06307249]

[0.02019002 0.01306413 0.96674584]]





precision recall f1-score support

hate 0.40 0.55 0.46 294

offensive 0.97 0.88 0.92 3821

neither 0.74 0.97 0.84 842

accuracy 0.87 4957

macro avg 0.70 0.80 0.74 4957

weighted avg 0.90 0.87 0.88 4957

average eval\_loss: 0.4003336899763634, accuracy: 0.8731087351220497

training 2 epoch...

44.37682914733887 seconds used......

2 training finished! average train loss: 0.3823384334120834

total clf loss: 67.68777361512184 total adv loss: 46.158030323684216

evaluating...

Confusion matrix, without normalization

[[ 236 42 16]

[ 783 2938 100]

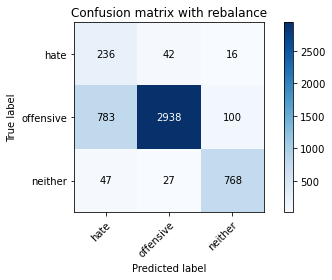
[ 47 27 768]]

Normalized confusion matrix

[[0.80272109 0.14285714 0.05442177]

[0.20492018 0.76890866 0.02617116]

[0.05581948 0.03206651 0.91211401]]





precision recall f1-score support

hate 0.22 0.80 0.35 294

offensive 0.98 0.77 0.86 3821

neither 0.87 0.91 0.89 842

accuracy 0.80 4957

macro avg 0.69 0.83 0.70 4957

weighted avg 0.91 0.80 0.84 4957

average eval\_loss: 0.465120138029658, accuracy: 0.795239055880573

training 3 epoch...

44.23863935470581 seconds used......

3 training finished! average train loss: 0.2585156051692558

total clf loss: 46.46849972754717 total adv loss: 45.24662365112454

evaluating...

Confusion matrix, without normalization

[[ 233 54 7]

[ 740 3026 55]

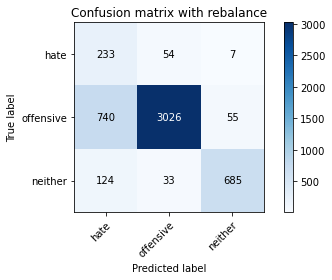
[ 124 33 685]]

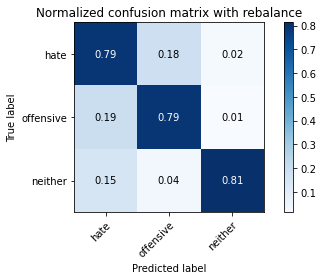
Normalized confusion matrix

[[0.79251701 0.18367347 0.02380952]

[0.19366658 0.79193928 0.01439414]

[0.14726841 0.0391924 0.81353919]]





precision recall f1-score support

hate 0.21 0.79 0.34 294

offensive 0.97 0.79 0.87 3821

neither 0.92 0.81 0.86 842

accuracy 0.80 4957

macro avg 0.70 0.80 0.69 4957

weighted avg 0.92 0.80 0.84 4957

average eval\_loss: 0.4958640183472345, accuracy: 0.7956425257212023

training 4 epoch...

44.295955181121826 seconds used......

4 training finished! average train loss: 0.14280336477102068

total clf loss: 26.699719823896885 total adv loss: 45.606889724731445

evaluating...

Confusion matrix, without normalization

[[ 233 50 11]

[ 779 2964 78]

[ 75 50 717]]

Normalized confusion matrix

[[0.79251701 0.17006803 0.03741497]

[0.20387333 0.77571316 0.0204135 ]

[0.08907363 0.05938242 0.85154394]]





precision recall f1-score support

hate 0.21 0.79 0.34 294

offensive 0.97 0.78 0.86 3821

neither 0.89 0.85 0.87 842

accuracy 0.79 4957

macro avg 0.69 0.81 0.69 4957

weighted avg 0.91 0.79 0.83 4957

average eval\_loss: 0.5959717974502353, accuracy: 0.7895904781117612

training 5 epoch...

44.2553596496582 seconds used......

5 training finished! average train loss: 0.07309974905931288

total clf loss: 14.92740893503651 total adv loss: 48.54703634977341

evaluating...

Confusion matrix, without normalization

[[ 260 27 7]

[1536 2227 58]

[ 131 37 674]]

Normalized confusion matrix

[[0.88435374 0.09183673 0.02380952]

[0.40198901 0.58283172 0.01517927]

[0.15558195 0.04394299 0.80047506]]





precision recall f1-score support

hate 0.13 0.88 0.23 294

offensive 0.97 0.58 0.73 3821

neither 0.91 0.80 0.85 842

accuracy 0.64 4957

macro avg 0.67 0.76 0.61 4957

weighted avg 0.91 0.64 0.72 4957

average eval\_loss: 1.2345365014888587, accuracy: 0.6376840831147872

training 6 epoch...

44.29470753669739 seconds used......

6 training finished! average train loss: 0.03216283635524978

total clf loss: 7.946513341506943 total adv loss: 48.933365389704704

evaluating...

Confusion matrix, without normalization

[[ 225 47 22]

[ 763 2956 102]

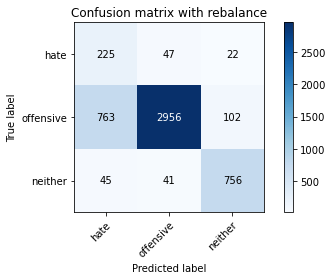
[ 45 41 756]]

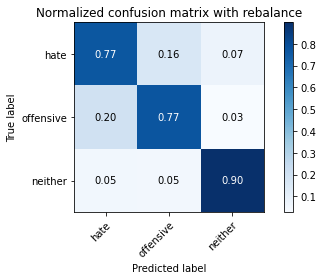
Normalized confusion matrix

[[0.76530612 0.15986395 0.07482993]

[0.19968595 0.77361947 0.02669458]

[0.05344418 0.04869359 0.89786223]]





precision recall f1-score support

hate 0.22 0.77 0.34 294

offensive 0.97 0.77 0.86 3821

neither 0.86 0.90 0.88 842

accuracy 0.79 4957

macro avg 0.68 0.81 0.69 4957

weighted avg 0.91 0.79 0.83 4957

average eval\_loss: 0.8471094027993404, accuracy: 0.7942303812789994

training 7 epoch...

44.289661169052124 seconds used......

7 training finished! average train loss: 0.008372588423543681

total clf loss: 3.984422770969104 total adv loss: 51.05420175939798

evaluating...

Confusion matrix, without normalization

[[ 213 61 20]

[ 733 2985 103]

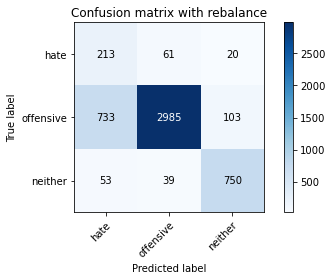
[ 53 39 750]]

Normalized confusion matrix

[[0.7244898 0.20748299 0.06802721]

[0.1918346 0.78120911 0.02695629]

[0.06294537 0.04631829 0.89073634]]





precision recall f1-score support

hate 0.21 0.72 0.33 294

offensive 0.97 0.78 0.86 3821

neither 0.86 0.89 0.87 842

accuracy 0.80 4957

macro avg 0.68 0.80 0.69 4957

weighted avg 0.90 0.80 0.83 4957

average eval\_loss: 0.9243201337052479, accuracy: 0.7964494654024612

training 8 epoch...

44.26032614707947 seconds used......

8 training finished! average train loss: 0.006266317375233038

total clf loss: 3.6503789992129896 total adv loss: 51.57677390426397

evaluating...

Confusion matrix, without normalization

[[ 218 58 18]

[ 758 2962 101]

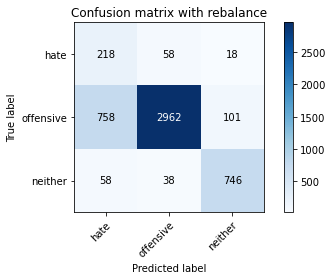
[ 58 38 746]]

Normalized confusion matrix

[[0.7414966 0.19727891 0.06122449]

[0.19837739 0.77518974 0.02643287]

[0.06888361 0.04513064 0.88598575]]





precision recall f1-score support

hate 0.21 0.74 0.33 294

offensive 0.97 0.78 0.86 3821

neither 0.86 0.89 0.87 842

accuracy 0.79 4957

macro avg 0.68 0.80 0.69 4957

weighted avg 0.91 0.79 0.83 4957

average eval\_loss: 0.9902433046072181, accuracy: 0.7920112971555376

training 9 epoch...

44.286858797073364 seconds used......

9 training finished! average train loss: -0.0011815018652344656

total clf loss: 2.485060703882482 total adv loss: 53.741949770599604

evaluating...

Confusion matrix, without normalization

[[ 231 42 21]

[1042 2655 124]

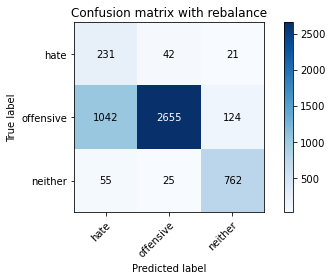
[ 55 25 762]]

Normalized confusion matrix

[[0.78571429 0.14285714 0.07142857]

[0.27270348 0.69484428 0.03245224]

[0.06532067 0.02969121 0.90498812]]





precision recall f1-score support

hate 0.17 0.79 0.28 294

offensive 0.98 0.69 0.81 3821

neither 0.84 0.90 0.87 842

accuracy 0.74 4957

macro avg 0.66 0.80 0.66 4957

weighted avg 0.90 0.74 0.79 4957

average eval\_loss: 1.3360332016534577, accuracy: 0.7359289893080492

training 10 epoch...

44.37171792984009 seconds used......

10 training finished! average train loss: 0.0062306382126931905

total clf loss: 3.7841510555590503 total adv loss: 54.37423697113991

evaluating...

Confusion matrix, without normalization

[[ 224 51 19]

[ 922 2757 142]

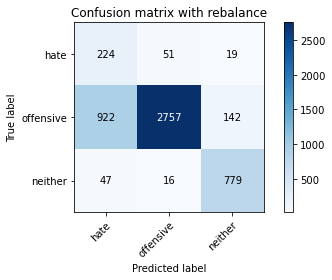
[ 47 16 779]]

Normalized confusion matrix

[[0.76190476 0.17346939 0.06462585]

[0.24129809 0.72153886 0.03716305]

[0.05581948 0.01900238 0.92517815]]





precision recall f1-score support

hate 0.19 0.76 0.30 294

offensive 0.98 0.72 0.83 3821

neither 0.83 0.93 0.87 842

accuracy 0.76 4957

macro avg 0.66 0.80 0.67 4957

weighted avg 0.90 0.76 0.81 4957

average eval\_loss: 1.3147659855744531, accuracy: 0.7585233003832963

3 run

training 1 epoch...

44.36198091506958 seconds used......

1 training finished! average train loss: 0.6676262225846799

total clf loss: 117.35342165827751 total adv loss: 63.78675416857004

evaluating...

Confusion matrix, without normalization

[[ 228 38 28]

[ 847 2799 175]

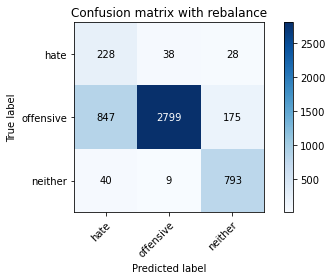
[ 40 9 793]]

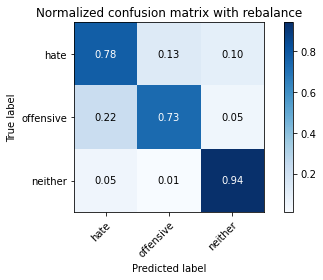
Normalized confusion matrix

[[0.7755102 0.1292517 0.0952381 ]

[0.22166972 0.73253075 0.04579953]

[0.04750594 0.01068884 0.94180523]]





precision recall f1-score support

hate 0.20 0.78 0.32 294

offensive 0.98 0.73 0.84 3821

neither 0.80 0.94 0.86 842

accuracy 0.77 4957

macro avg 0.66 0.82 0.68 4957

weighted avg 0.91 0.77 0.81 4957

average eval\_loss: 0.5678057254201943, accuracy: 0.7706273956021787

training 2 epoch...

44.279398918151855 seconds used......

2 training finished! average train loss: 0.3792306999539771

total clf loss: 67.28945955634117 total adv loss: 48.82019792497158

evaluating...

Confusion matrix, without normalization

[[ 245 34 15]

[1055 2669 97]

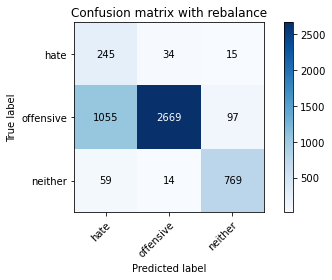
[ 59 14 769]]

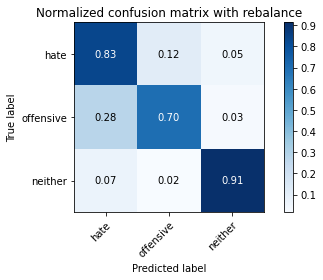
Normalized confusion matrix

[[0.83333333 0.11564626 0.05102041]

[0.27610573 0.69850824 0.02538602]

[0.07007126 0.01662708 0.91330166]]





precision recall f1-score support

hate 0.18 0.83 0.30 294

offensive 0.98 0.70 0.82 3821

neither 0.87 0.91 0.89 842

accuracy 0.74 4957

macro avg 0.68 0.82 0.67 4957

weighted avg 0.92 0.74 0.80 4957

average eval\_loss: 0.6165135061788943, accuracy: 0.7429897115190639

training 3 epoch...

44.32008504867554 seconds used......

3 training finished! average train loss: 0.2573265330266883

total clf loss: 46.33845100924373 total adv loss: 46.71227442845702

evaluating...

Confusion matrix, without normalization

[[ 210 64 20]

[ 672 3031 118]

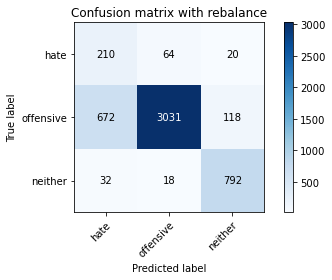
[ 32 18 792]]

Normalized confusion matrix

[[0.71428571 0.21768707 0.06802721]

[0.17587019 0.79324784 0.03088197]

[0.03800475 0.02137767 0.94061758]]





precision recall f1-score support

hate 0.23 0.71 0.35 294

offensive 0.97 0.79 0.87 3821

neither 0.85 0.94 0.89 842

accuracy 0.81 4957

macro avg 0.69 0.82 0.71 4957

weighted avg 0.91 0.81 0.85 4957

average eval\_loss: 0.45906819787717634, accuracy: 0.8135969336292113

training 4 epoch...

44.32939386367798 seconds used......

4 training finished! average train loss: 0.14387905350423347

total clf loss: 27.036606768146157 total adv loss: 48.6657734811306

evaluating...

Confusion matrix, without normalization

[[ 248 35 11]

[1220 2531 70]

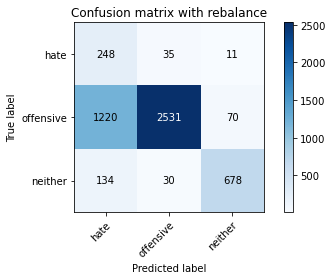
[ 134 30 678]]

Normalized confusion matrix

[[0.84353741 0.11904762 0.03741497]

[0.31928814 0.66239204 0.01831981]

[0.15914489 0.03562945 0.80522565]]





precision recall f1-score support

hate 0.15 0.84 0.26 294

offensive 0.97 0.66 0.79 3821

neither 0.89 0.81 0.85 842

accuracy 0.70 4957

macro avg 0.67 0.77 0.63 4957

weighted avg 0.91 0.70 0.77 4957

average eval\_loss: 0.8375389585421691, accuracy: 0.6973976195279403

training 5 epoch...

44.443055629730225 seconds used......

5 training finished! average train loss: 0.07466153477520099

total clf loss: 15.163591824937612 total adv loss: 47.929387509822845

evaluating...

Confusion matrix, without normalization

[[ 232 44 18]

[ 885 2808 128]

[ 58 12 772]]

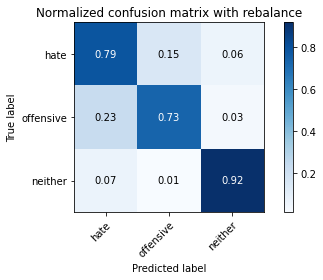
Normalized confusion matrix

[[0.78911565 0.14965986 0.06122449]

[0.23161476 0.73488616 0.03349908]

[0.06888361 0.01425178 0.91686461]]





precision recall f1-score support

hate 0.20 0.79 0.32 294

offensive 0.98 0.73 0.84 3821

neither 0.84 0.92 0.88 842

accuracy 0.77 4957

macro avg 0.67 0.81 0.68 4957

weighted avg 0.91 0.77 0.82 4957

average eval\_loss: 0.9185603289196508, accuracy: 0.7690135162396611

training 6 epoch...

44.3116934299469 seconds used......

6 training finished! average train loss: 0.04105439250036116

total clf loss: 9.547349698608741 total adv loss: 50.54097055643797

evaluating...

Confusion matrix, without normalization

[[ 206 68 20]

[ 554 3153 114]

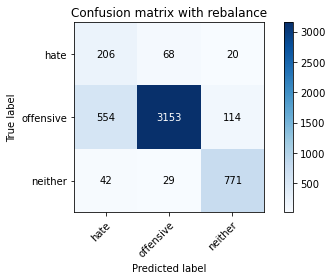
[ 42 29 771]]

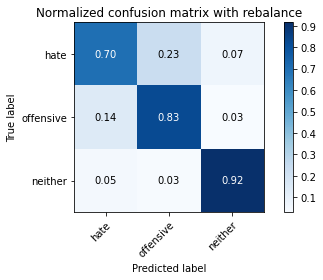
Normalized confusion matrix

[[0.70068027 0.23129252 0.06802721]

[0.14498822 0.82517666 0.02983512]

[0.04988124 0.03444181 0.91567696]]





precision recall f1-score support

hate 0.26 0.70 0.38 294

offensive 0.97 0.83 0.89 3821

neither 0.85 0.92 0.88 842

accuracy 0.83 4957

macro avg 0.69 0.81 0.72 4957

weighted avg 0.91 0.83 0.86 4957

average eval\_loss: 0.6354880621506562, accuracy: 0.8331652208997378

training 7 epoch...

44.33254790306091 seconds used......

7 training finished! average train loss: 0.021110398098499628

total clf loss: 6.188290699326899 total adv loss: 51.56825217604637

evaluating...

Confusion matrix, without normalization

[[ 229 47 18]

[ 937 2773 111]

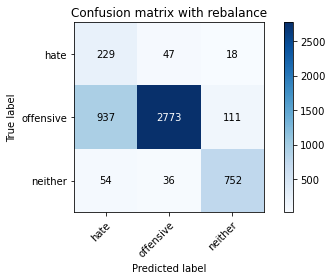
[ 54 36 752]]

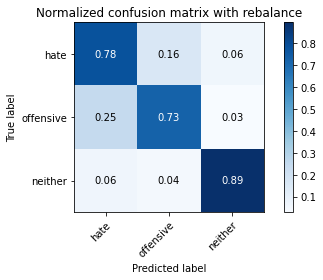
Normalized confusion matrix

[[0.77891156 0.15986395 0.06122449]

[0.24522376 0.72572625 0.02904999]

[0.06413302 0.04275534 0.89311164]]





precision recall f1-score support

hate 0.19 0.78 0.30 294

offensive 0.97 0.73 0.83 3821

neither 0.85 0.89 0.87 842

accuracy 0.76 4957

macro avg 0.67 0.80 0.67 4957

weighted avg 0.90 0.76 0.81 4957

average eval\_loss: 1.019998052195015, accuracy: 0.7573128908614081

training 8 epoch...

44.318124294281006 seconds used......

8 training finished! average train loss: 0.0018136091584786337

total clf loss: 3.1611648452817462 total adv loss: 57.02075282484293

evaluating...

Confusion matrix, without normalization

[[ 236 46 12]

[1092 2639 90]

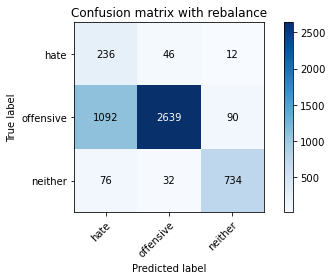
[ 76 32 734]]

Normalized confusion matrix

[[0.80272109 0.15646259 0.04081633]

[0.28578906 0.6906569 0.02355404]

[0.09026128 0.03800475 0.87173397]]





precision recall f1-score support

hate 0.17 0.80 0.28 294

offensive 0.97 0.69 0.81 3821

neither 0.88 0.87 0.87 842

accuracy 0.73 4957

macro avg 0.67 0.79 0.65 4957

weighted avg 0.91 0.73 0.79 4957

average eval\_loss: 1.3022372537112283, accuracy: 0.7280613274157757

training 9 epoch...

44.384297370910645 seconds used......

9 training finished! average train loss: 0.00461453372307601

total clf loss: 3.612701938371174 total adv loss: 56.47233162820339

evaluating...

Confusion matrix, without normalization

[[ 230 46 18]

[ 961 2761 99]

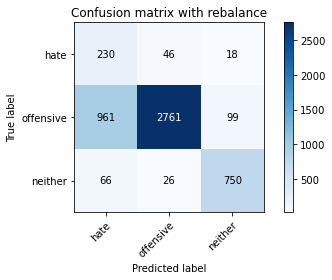
[ 66 26 750]]

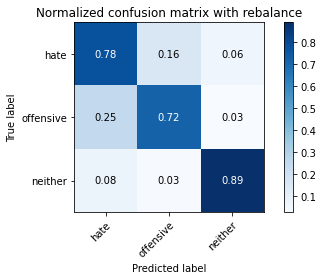
Normalized confusion matrix

[[0.78231293 0.15646259 0.06122449]

[0.25150484 0.72258571 0.02590945]

[0.0783848 0.03087886 0.89073634]]





precision recall f1-score support

hate 0.18 0.78 0.30 294

offensive 0.97 0.72 0.83 3821

neither 0.87 0.89 0.88 842

accuracy 0.75 4957

macro avg 0.67 0.80 0.67 4957

weighted avg 0.91 0.75 0.81 4957

average eval\_loss: 1.2376853671195287, accuracy: 0.7546903368973169

training 10 epoch...

44.38400435447693 seconds used......

10 training finished! average train loss: -0.002746551574006328

total clf loss: 2.3244452728395117 total adv loss: 55.882111079990864

evaluating...

Confusion matrix, without normalization

[[ 222 47 25]

[ 783 2892 146]

[ 41 12 789]]

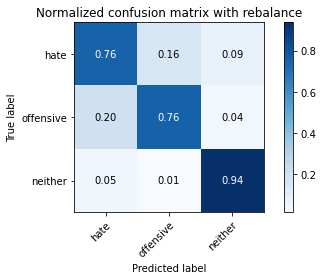
Normalized confusion matrix

[[0.75510204 0.15986395 0.08503401]

[0.20492018 0.75686993 0.03820989]

[0.04869359 0.01425178 0.93705463]]





precision recall f1-score support

hate 0.21 0.76 0.33 294

offensive 0.98 0.76 0.85 3821

neither 0.82 0.94 0.88 842

accuracy 0.79 4957

macro avg 0.67 0.82 0.69 4957

weighted avg 0.91 0.79 0.83 4957

average eval\_loss: 1.2325917801398676, accuracy: 0.7873713939882994

4 run

training 1 epoch...

44.506208419799805 seconds used......

1 training finished! average train loss: 0.6643238236681063

total clf loss: 116.75179915130138 total adv loss: 63.04850798845291

evaluating...

Confusion matrix, without normalization

[[ 239 33 22]

[ 943 2739 139]

[ 63 10 769]]

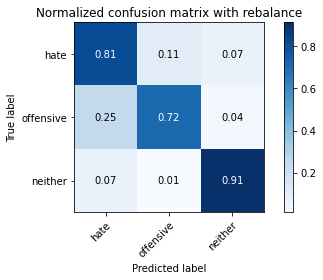
Normalized confusion matrix

[[0.81292517 0.1122449 0.07482993]

[0.24679403 0.71682806 0.03637791]

[0.07482185 0.01187648 0.91330166]]





precision recall f1-score support

hate 0.19 0.81 0.31 294

offensive 0.98 0.72 0.83 3821

neither 0.83 0.91 0.87 842

accuracy 0.76 4957

macro avg 0.67 0.81 0.67 4957

weighted avg 0.91 0.76 0.81 4957

average eval\_loss: 0.6287983358507195, accuracy: 0.7559007464192051

training 2 epoch...

44.50821328163147 seconds used......

2 training finished! average train loss: 0.38690339333829826

total clf loss: 68.47975242137909 total adv loss: 46.38544664159417

evaluating...

Confusion matrix, without normalization

[[ 206 57 31]

[ 573 3097 151]

[ 29 15 798]]

Normalized confusion matrix

[[0.70068027 0.19387755 0.10544218]

[0.14996074 0.81052081 0.03951845]

[0.03444181 0.01781473 0.94774347]]





precision recall f1-score support

hate 0.25 0.70 0.37 294

offensive 0.98 0.81 0.89 3821

neither 0.81 0.95 0.88 842

accuracy 0.83 4957

macro avg 0.68 0.82 0.71 4957

weighted avg 0.91 0.83 0.85 4957

average eval\_loss: 0.4407438601277048, accuracy: 0.8273149082106113

training 3 epoch...

44.28643274307251 seconds used......

3 training finished! average train loss: 0.2610753324884943

total clf loss: 46.96842301636934 total adv loss: 46.49082352221012

evaluating...

Confusion matrix, without normalization

[[ 218 59 17]

[ 560 3142 119]

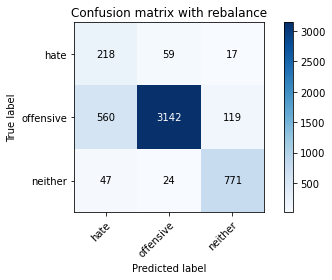
[ 47 24 771]]

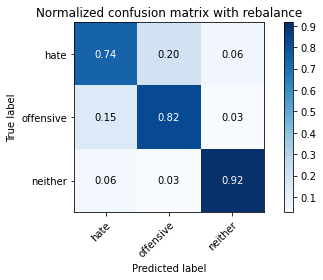
Normalized confusion matrix

[[0.7414966 0.20068027 0.05782313]

[0.14655849 0.82229783 0.03114368]

[0.05581948 0.02850356 0.91567696]]





precision recall f1-score support

hate 0.26 0.74 0.39 294

offensive 0.97 0.82 0.89 3821

neither 0.85 0.92 0.88 842

accuracy 0.83 4957

macro avg 0.70 0.83 0.72 4957

weighted avg 0.91 0.83 0.86 4957

average eval\_loss: 0.45397874776033625, accuracy: 0.8333669558200525

training 4 epoch...

44.334688901901245 seconds used......

4 training finished! average train loss: 0.14346446529335793

total clf loss: 26.873647732660174 total adv loss: 46.824483927339315

evaluating...

Confusion matrix, without normalization

[[ 206 75 13]

[ 490 3220 111]

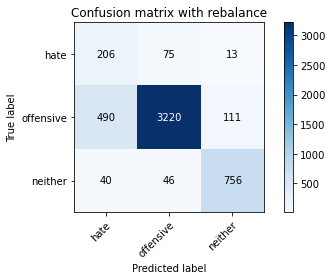
[ 40 46 756]]

Normalized confusion matrix

[[0.70068027 0.25510204 0.04421769]

[0.12823868 0.84271133 0.02904999]

[0.04750594 0.05463183 0.89786223]]





precision recall f1-score support

hate 0.28 0.70 0.40 294

offensive 0.96 0.84 0.90 3821

neither 0.86 0.90 0.88 842

accuracy 0.84 4957

macro avg 0.70 0.81 0.73 4957

weighted avg 0.91 0.84 0.87 4957

average eval\_loss: 0.49122365120603073, accuracy: 0.8436554367561024

training 5 epoch...

44.26951456069946 seconds used......

5 training finished! average train loss: 0.07396291303094367

total clf loss: 15.078385851345956 total adv loss: 48.6145536005497

evaluating...

Confusion matrix, without normalization

[[ 207 68 19]

[ 541 3148 132]

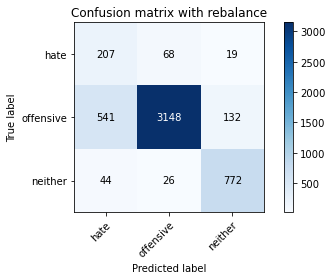
[ 44 26 772]]

Normalized confusion matrix

[[0.70408163 0.23129252 0.06462585]

[0.14158597 0.8238681 0.03454593]

[0.05225653 0.03087886 0.91686461]]





precision recall f1-score support

hate 0.26 0.70 0.38 294

offensive 0.97 0.82 0.89 3821

neither 0.84 0.92 0.87 842

accuracy 0.83 4957

macro avg 0.69 0.81 0.72 4957

weighted avg 0.91 0.83 0.86 4957

average eval\_loss: 0.5829287789134879, accuracy: 0.8325600161387936

training 6 epoch...

44.30049538612366 seconds used......

6 training finished! average train loss: 0.037916472165338826

total clf loss: 8.999451102688909 total adv loss: 50.31468719244003

evaluating...

Confusion matrix, without normalization

[[ 233 45 16]

[ 825 2900 96]

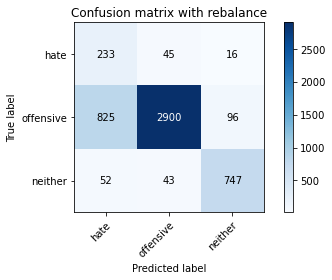
[ 52 43 747]]

Normalized confusion matrix

[[0.79251701 0.15306122 0.05442177]

[0.21591206 0.75896362 0.02512431]

[0.06175772 0.05106888 0.8871734 ]]





precision recall f1-score support

hate 0.21 0.79 0.33 294

offensive 0.97 0.76 0.85 3821

neither 0.87 0.89 0.88 842

accuracy 0.78 4957

macro avg 0.68 0.81 0.69 4957

weighted avg 0.91 0.78 0.83 4957

average eval\_loss: 0.797431011177269, accuracy: 0.7827314908210611

training 7 epoch...

44.31272840499878 seconds used......

7 training finished! average train loss: 0.0165428106362621

total clf loss: 5.464915244141594 total adv loss: 52.72189174592495

evaluating...

Confusion matrix, without normalization

[[ 235 48 11]

[ 843 2876 102]

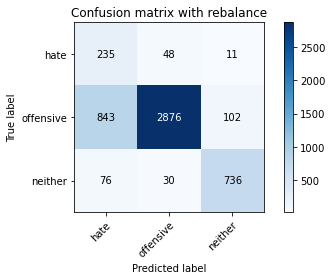
[ 76 30 736]]

Normalized confusion matrix

[[0.79931973 0.16326531 0.03741497]

[0.22062287 0.75268254 0.02669458]

[0.09026128 0.03562945 0.87410926]]





precision recall f1-score support

hate 0.20 0.80 0.32 294

offensive 0.97 0.75 0.85 3821

neither 0.87 0.87 0.87 842

accuracy 0.78 4957

macro avg 0.68 0.81 0.68 4957

weighted avg 0.91 0.78 0.82 4957

average eval\_loss: 0.9731637031579721, accuracy: 0.7760742384506758

training 8 epoch...

44.35116386413574 seconds used......

8 training finished! average train loss: 0.010194985205525456

total clf loss: 4.550601345836185 total adv loss: 56.14517657458782

evaluating...

Confusion matrix, without normalization

[[ 214 57 23]

[ 751 2931 139]

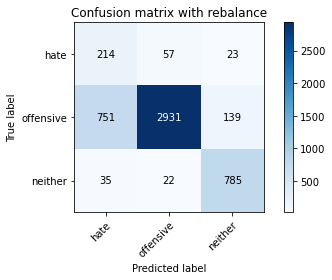
[ 35 22 785]]

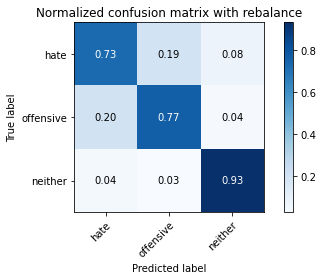
Normalized confusion matrix

[[0.72789116 0.19387755 0.07823129]

[0.19654541 0.76707668 0.03637791]

[0.0415677 0.02612827 0.93230404]]





precision recall f1-score support

hate 0.21 0.73 0.33 294

offensive 0.97 0.77 0.86 3821

neither 0.83 0.93 0.88 842

accuracy 0.79 4957

macro avg 0.67 0.81 0.69 4957

weighted avg 0.90 0.79 0.83 4957

average eval\_loss: 0.9678674819653914, accuracy: 0.7928182368367964

training 9 epoch...

44.34931015968323 seconds used......

9 training finished! average train loss: 0.007835719780724126

total clf loss: 4.028942354692845 total adv loss: 53.78068459033966

evaluating...

Confusion matrix, without normalization

[[ 214 54 26]

[ 686 2993 142]

[ 32 30 780]]

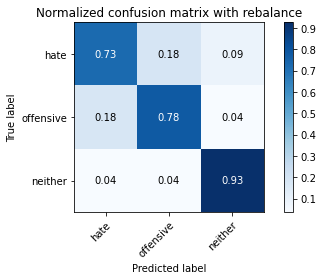
Normalized confusion matrix

[[0.72789116 0.18367347 0.08843537]

[0.17953415 0.7833028 0.03716305]

[0.03800475 0.03562945 0.9263658 ]]





precision recall f1-score support

hate 0.23 0.73 0.35 294

offensive 0.97 0.78 0.87 3821

neither 0.82 0.93 0.87 842

accuracy 0.80 4957

macro avg 0.68 0.81 0.70 4957

weighted avg 0.90 0.80 0.84 4957

average eval\_loss: 1.0553324398567765, accuracy: 0.8043171272947347

training 10 epoch...

44.40094709396362 seconds used......

10 training finished! average train loss: -0.006015015025323594

total clf loss: 1.6350450564059429 total adv loss: 53.27225162088871

evaluating...

Confusion matrix, without normalization

[[ 191 78 25]

[ 482 3202 137]

[ 32 33 777]]

Normalized confusion matrix

[[0.64965986 0.26530612 0.08503401]

[0.12614499 0.83800052 0.03585449]

[0.03800475 0.0391924 0.92280285]]





precision recall f1-score support

hate 0.27 0.65 0.38 294

offensive 0.97 0.84 0.90 3821

neither 0.83 0.92 0.87 842

accuracy 0.84 4957

macro avg 0.69 0.80 0.72 4957

weighted avg 0.90 0.84 0.86 4957

average eval\_loss: 0.8785603687607481, accuracy: 0.841234617712326

5 run

training 1 epoch...

45.39558172225952 seconds used......

1 training finished! average train loss: 0.638527939455551

total clf loss: 112.32127569615841 total adv loss: 62.65996987372637

evaluating...

Confusion matrix, without normalization

[[ 169 96 29]

[ 255 3445 121]

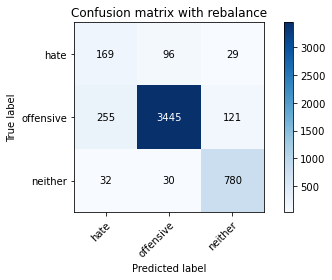
[ 32 30 780]]

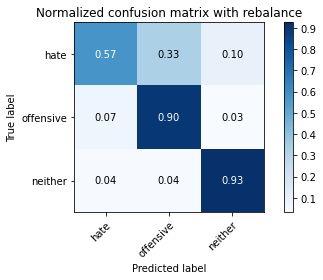
Normalized confusion matrix

[[0.57482993 0.32653061 0.09863946]

[0.06673646 0.90159644 0.0316671 ]

[0.03800475 0.03562945 0.9263658 ]]





precision recall f1-score support

hate 0.37 0.57 0.45 294

offensive 0.96 0.90 0.93 3821

neither 0.84 0.93 0.88 842

accuracy 0.89 4957

macro avg 0.72 0.80 0.75 4957

weighted avg 0.91 0.89 0.89 4957

average eval\_loss: 0.34021259859324465, accuracy: 0.8864232398628202

training 2 epoch...

44.736350536346436 seconds used......

2 training finished! average train loss: 0.37744018400621693

total clf loss: 67.03018271178007 total adv loss: 49.75822115689516

evaluating...

Confusion matrix, without normalization

[[ 199 60 35]

[ 504 3153 164]

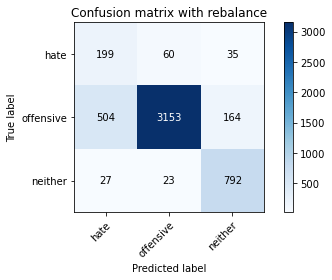
[ 27 23 792]]

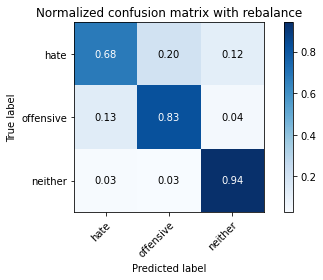
Normalized confusion matrix

[[0.67687075 0.20408163 0.11904762]

[0.13190264 0.82517666 0.0429207 ]

[0.03206651 0.02731591 0.94061758]]





precision recall f1-score support

hate 0.27 0.68 0.39 294

offensive 0.97 0.83 0.89 3821

neither 0.80 0.94 0.86 842

accuracy 0.84 4957

macro avg 0.68 0.81 0.72 4957

weighted avg 0.90 0.84 0.86 4957

average eval\_loss: 0.3889080216327021, accuracy: 0.8359895097841437

training 3 epoch...

44.73344826698303 seconds used......

3 training finished! average train loss: 0.26126080827667697

total clf loss: 47.10826414823532 total adv loss: 48.65331960096955

evaluating...

Confusion matrix, without normalization

[[ 221 54 19]

[ 693 3025 103]

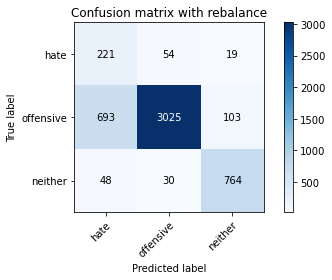
[ 48 30 764]]

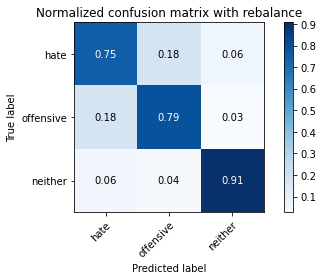
Normalized confusion matrix

[[0.75170068 0.18367347 0.06462585]

[0.18136613 0.79167757 0.02695629]

[0.05700713 0.03562945 0.90736342]]





precision recall f1-score support

hate 0.23 0.75 0.35 294

offensive 0.97 0.79 0.87 3821

neither 0.86 0.91 0.88 842

accuracy 0.81 4957

macro avg 0.69 0.82 0.70 4957

weighted avg 0.91 0.81 0.84 4957

average eval\_loss: 0.46656534021660206, accuracy: 0.808957030461973

training 4 epoch...

44.651755571365356 seconds used......

4 training finished! average train loss: 0.14651669427338573

total clf loss: 27.514713475480676 total adv loss: 49.207175083458424

evaluating...

Confusion matrix, without normalization

[[ 206 61 27]

[ 638 3036 147]

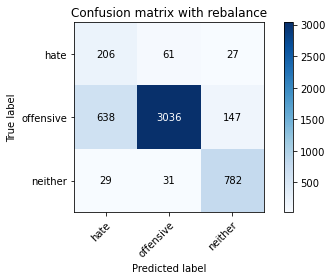
[ 29 31 782]]

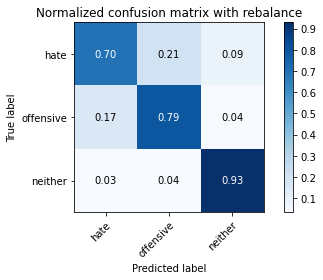
Normalized confusion matrix

[[0.70068027 0.20748299 0.09183673]

[0.166972 0.7945564 0.0384716 ]

[0.03444181 0.0368171 0.92874109]]





precision recall f1-score support

hate 0.24 0.70 0.35 294

offensive 0.97 0.79 0.87 3821

neither 0.82 0.93 0.87 842

accuracy 0.81 4957

macro avg 0.67 0.81 0.70 4957

weighted avg 0.90 0.81 0.84 4957

average eval\_loss: 0.5165339161964855, accuracy: 0.8117813193463789

training 5 epoch...

44.67612099647522 seconds used......

5 training finished! average train loss: 0.07573437359325631

total clf loss: 15.40578690636903 total adv loss: 49.104180946946144

evaluating...

Confusion matrix, without normalization

[[ 215 59 20]

[ 672 3029 120]

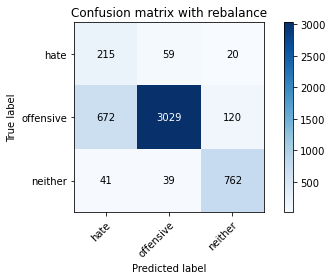
[ 41 39 762]]

Normalized confusion matrix

[[0.73129252 0.20068027 0.06802721]

[0.17587019 0.79272442 0.03140539]

[0.04869359 0.04631829 0.90498812]]





precision recall f1-score support

hate 0.23 0.73 0.35 294

offensive 0.97 0.79 0.87 3821

neither 0.84 0.90 0.87 842

accuracy 0.81 4957

macro avg 0.68 0.81 0.70 4957

weighted avg 0.90 0.81 0.84 4957

average eval\_loss: 0.6539906635504937, accuracy: 0.8081500907807141

training 6 epoch...

44.750460386276245 seconds used......

6 training finished! average train loss: 0.03044057951706369

total clf loss: 7.719881418161094 total adv loss: 50.29084566608071

evaluating...

Confusion matrix, without normalization

[[ 239 33 22]

[1083 2580 158]

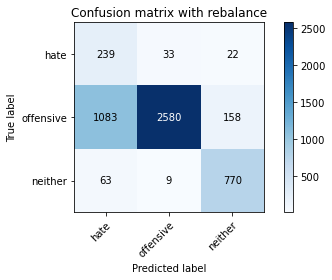
[ 63 9 770]]

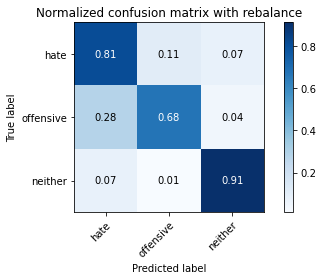
Normalized confusion matrix

[[0.81292517 0.1122449 0.07482993]

[0.28343366 0.67521591 0.04135043]

[0.07482185 0.01068884 0.91448931]]





precision recall f1-score support

hate 0.17 0.81 0.28 294

offensive 0.98 0.68 0.80 3821

neither 0.81 0.91 0.86 842

accuracy 0.72 4957

macro avg 0.66 0.80 0.65 4957

weighted avg 0.91 0.72 0.78 4957

average eval\_loss: 1.26541902849661, accuracy: 0.7240266290094816

training 7 epoch...

44.78090524673462 seconds used......

7 training finished! average train loss: 0.013538283734359065

total clf loss: 4.8484639813541435 total adv loss: 50.668348237872124

evaluating...

Confusion matrix, without normalization

[[ 242 31 21]

[1133 2550 138]

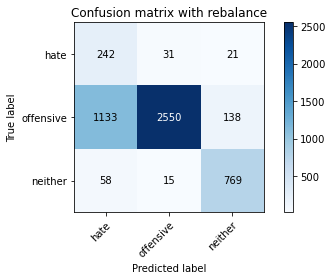
[ 58 15 769]]

Normalized confusion matrix

[[0.82312925 0.10544218 0.07142857]

[0.29651924 0.66736456 0.0361162 ]

[0.06888361 0.01781473 0.91330166]]





precision recall f1-score support

hate 0.17 0.82 0.28 294

offensive 0.98 0.67 0.79 3821

neither 0.83 0.91 0.87 842

accuracy 0.72 4957

macro avg 0.66 0.80 0.65 4957

weighted avg 0.91 0.72 0.78 4957

average eval\_loss: 1.384833384275196, accuracy: 0.7183780512406698

training 8 epoch...

44.792163133621216 seconds used......

8 training finished! average train loss: 0.003427063866534777

total clf loss: 3.2590912616287824 total adv loss: 53.46126524358988

evaluating...

Confusion matrix, without normalization

[[ 208 60 26]

[ 584 3103 134]

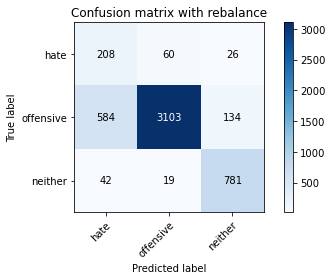
[ 42 19 781]]

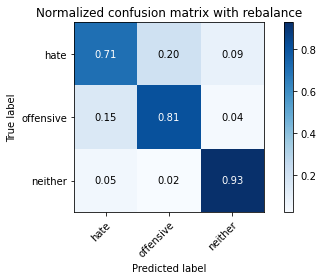
Normalized confusion matrix

[[0.70748299 0.20408163 0.08843537]

[0.15283957 0.81209108 0.03506935]

[0.04988124 0.02256532 0.92755344]]





precision recall f1-score support

hate 0.25 0.71 0.37 294

offensive 0.98 0.81 0.89 3821

neither 0.83 0.93 0.88 842

accuracy 0.83 4957

macro avg 0.68 0.82 0.71 4957

weighted avg 0.91 0.83 0.85 4957

average eval\_loss: 0.9615536329201272, accuracy: 0.8254992939277789

training 9 epoch...

44.656962156295776 seconds used......

9 training finished! average train loss: 0.0031465408923267795

total clf loss: 3.404191538138548 total adv loss: 57.32265993952751

evaluating...

Confusion matrix, without normalization

[[ 248 34 12]

[1265 2451 105]

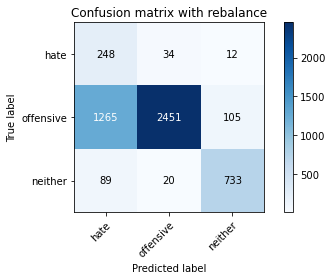
[ 89 20 733]]

Normalized confusion matrix

[[0.84353741 0.11564626 0.04081633]

[0.33106517 0.64145512 0.02747972]

[0.10570071 0.02375297 0.87054632]]





precision recall f1-score support

hate 0.15 0.84 0.26 294

offensive 0.98 0.64 0.77 3821

neither 0.86 0.87 0.87 842

accuracy 0.69 4957

macro avg 0.67 0.79 0.63 4957

weighted avg 0.91 0.69 0.76 4957

average eval\_loss: 1.619932395368514, accuracy: 0.6923542465200726

training 10 epoch...

44.683024168014526 seconds used......

10 training finished! average train loss: -0.002143601763704846

total clf loss: 2.4807417939591687 total adv loss: 56.94595246762037

evaluating...

Confusion matrix, without normalization

[[ 236 43 15]

[ 915 2793 113]

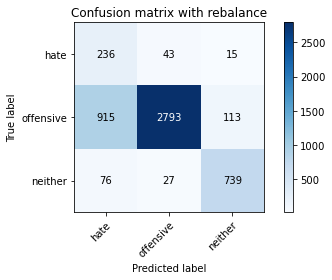
[ 76 27 739]]

Normalized confusion matrix

[[0.80272109 0.1462585 0.05102041]

[0.23946611 0.73096048 0.02957341]

[0.09026128 0.03206651 0.87767221]]





precision recall f1-score support

hate 0.19 0.80 0.31 294

offensive 0.98 0.73 0.84 3821

neither 0.85 0.88 0.86 842

accuracy 0.76 4957

macro avg 0.67 0.80 0.67 4957

weighted avg 0.91 0.76 0.81 4957

average eval\_loss: 1.3416179321150505, accuracy: 0.760137179745814