alpha=2

1 run......

Downloading: 100%

411/411 [00:00<00:00, 14.9kB/s]

Downloading: 100%

263M/263M [00:06<00:00, 45.2MB/s]

training 1 epoch...

115.40177178382874 seconds used......

1 training finished! average train loss: 2.049259607990583

total clf loss: 165.7420430779457 total adv loss: 89.26678550243378

evaluating...

Confusion matrix, without normalization

[[109 9 6]

[ 74 81 9]

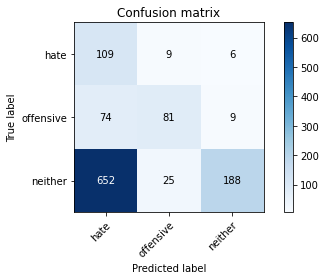
[652 25 188]]

Normalized confusion matrix

[[0.87903226 0.07258065 0.0483871 ]

[0.45121951 0.49390244 0.05487805]

[0.75375723 0.02890173 0.21734104]]





precision recall f1-score support

hate 0.13 0.88 0.23 124

offensive 0.70 0.49 0.58 164

neither 0.93 0.22 0.35 865

accuracy 0.33 1153

macro avg 0.59 0.53 0.39 1153

weighted avg 0.81 0.33 0.37 1153

average eval\_loss: 0.943305222124889, accuracy: 0.32784041630529054

training 2 epoch...

114.53626155853271 seconds used......

2 training finished! average train loss: 1.699177702977544

total clf loss: 143.09421080350876 total adv loss: 71.18382108211517

evaluating...

Confusion matrix, without normalization

[[ 83 31 10]

[ 37 117 10]

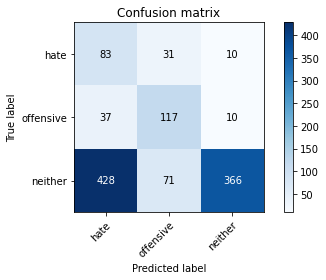
[428 71 366]]

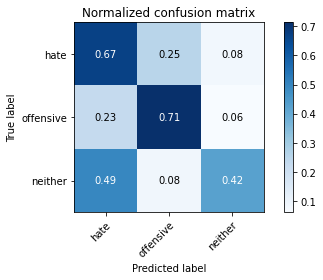
Normalized confusion matrix

[[0.66935484 0.25 0.08064516]

[0.22560976 0.71341463 0.06097561]

[0.49479769 0.08208092 0.42312139]]





precision recall f1-score support

hate 0.15 0.67 0.25 124

offensive 0.53 0.71 0.61 164

neither 0.95 0.42 0.59 865

accuracy 0.49 1153

macro avg 0.54 0.60 0.48 1153

weighted avg 0.80 0.49 0.55 1153

average eval\_loss: 0.888242571518339, accuracy: 0.4908933217692975

training 3 epoch...

114.57750296592712 seconds used......

3 training finished! average train loss: 1.3269087750287283

total clf loss: 119.83322381973267 total adv loss: 51.543724939227104

evaluating...

Confusion matrix, without normalization

[[ 93 27 4]

[ 45 113 6]

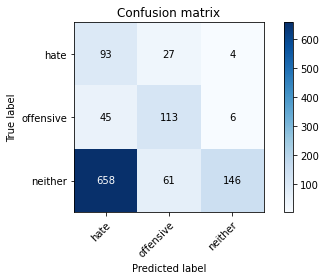
[658 61 146]]

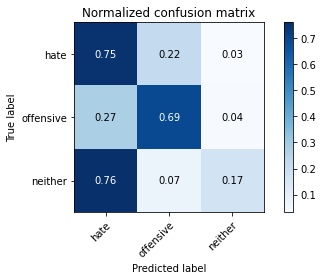
Normalized confusion matrix

[[0.75 0.21774194 0.03225806]

[0.27439024 0.68902439 0.03658537]

[0.76069364 0.07052023 0.16878613]]





precision recall f1-score support

hate 0.12 0.75 0.20 124

offensive 0.56 0.69 0.62 164

neither 0.94 0.17 0.29 865

accuracy 0.31 1153

macro avg 0.54 0.54 0.37 1153

weighted avg 0.79 0.31 0.32 1153

average eval\_loss: 1.204610928379256, accuracy: 0.30529054640069386

training 4 epoch...

114.64821696281433 seconds used......

4 training finished! average train loss: 0.9181741600235304

total clf loss: 92.40078213810921 total adv loss: 30.926238475833088

evaluating...

Confusion matrix, without normalization

[[ 67 36 21]

[ 33 116 15]

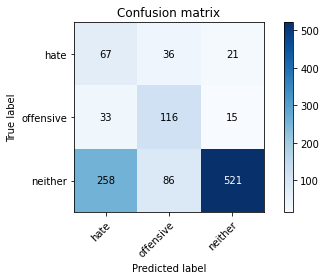
[258 86 521]]

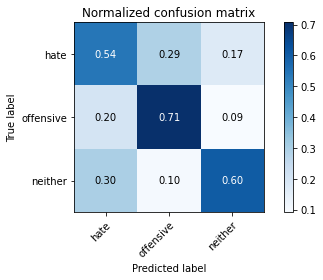
Normalized confusion matrix

[[0.54032258 0.29032258 0.16935484]

[0.20121951 0.70731707 0.09146341]

[0.2982659 0.09942197 0.60231214]]





precision recall f1-score support

hate 0.19 0.54 0.28 124

offensive 0.49 0.71 0.58 164

neither 0.94 0.60 0.73 865

accuracy 0.61 1153

macro avg 0.54 0.62 0.53 1153

weighted avg 0.79 0.61 0.66 1153

average eval\_loss: 0.9368120442176687, accuracy: 0.6105810928013877

training 5 epoch...

114.62341284751892 seconds used......

5 training finished! average train loss: 0.5939868954675538

total clf loss: 69.41663040220737 total adv loss: 15.186583883594722

evaluating...

Confusion matrix, without normalization

[[105 10 9]

[ 59 94 11]

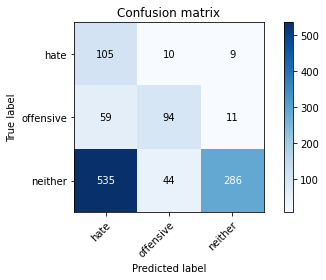
[535 44 286]]

Normalized confusion matrix

[[0.84677419 0.08064516 0.07258065]

[0.3597561 0.57317073 0.06707317]

[0.61849711 0.05086705 0.33063584]]





precision recall f1-score support

hate 0.15 0.85 0.26 124

offensive 0.64 0.57 0.60 164

neither 0.93 0.33 0.49 865

accuracy 0.42 1153

macro avg 0.57 0.58 0.45 1153

weighted avg 0.81 0.42 0.48 1153

average eval\_loss: 1.5584404273279782, accuracy: 0.4206418039895924

training 6 epoch...

114.53963971138 seconds used......

6 training finished! average train loss: 0.3954117605906157

total clf loss: 50.89842917025089 total adv loss: 7.7653733079641825

evaluating...

Confusion matrix, without normalization

[[ 95 17 12]

[ 48 101 15]

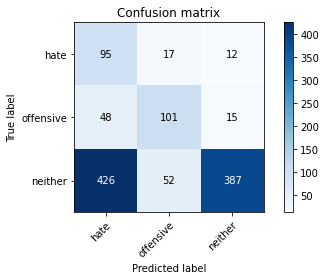
[426 52 387]]

Normalized confusion matrix

[[0.76612903 0.13709677 0.09677419]

[0.29268293 0.61585366 0.09146341]

[0.49248555 0.06011561 0.44739884]]





precision recall f1-score support

hate 0.17 0.77 0.27 124

offensive 0.59 0.62 0.60 164

neither 0.93 0.45 0.61 865

accuracy 0.51 1153

macro avg 0.57 0.61 0.49 1153

weighted avg 0.80 0.51 0.57 1153

average eval\_loss: 1.5583054649418797, accuracy: 0.5056374674761491

training 7 epoch...

114.57377886772156 seconds used......

7 training finished! average train loss: 0.20747437603062108

total clf loss: 29.731307700276375 total adv loss: 2.5621936828902108

evaluating...

Confusion matrix, without normalization

[[ 93 22 9]

[ 45 108 11]

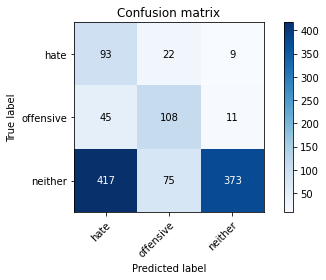
[417 75 373]]

Normalized confusion matrix

[[0.75 0.17741935 0.07258065]

[0.27439024 0.65853659 0.06707317]

[0.48208092 0.0867052 0.43121387]]





precision recall f1-score support

hate 0.17 0.75 0.27 124

offensive 0.53 0.66 0.59 164

neither 0.95 0.43 0.59 865

accuracy 0.50 1153

macro avg 0.55 0.61 0.48 1153

weighted avg 0.80 0.50 0.56 1153

average eval\_loss: 2.0159286599734734, accuracy: 0.4978317432784042

training 8 epoch...

114.60559511184692 seconds used......

8 training finished! average train loss: 0.18694650895139647

total clf loss: 21.791321363300085 total adv loss: 4.807846070456435

evaluating...

Confusion matrix, without normalization

[[ 99 15 10]

[ 44 105 15]

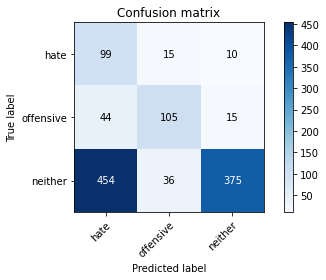
[454 36 375]]

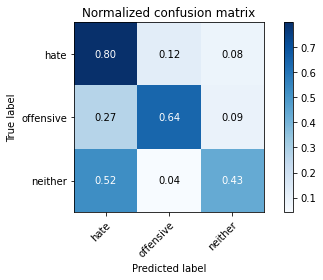
Normalized confusion matrix

[[0.7983871 0.12096774 0.08064516]

[0.26829268 0.6402439 0.09146341]

[0.52485549 0.0416185 0.43352601]]





precision recall f1-score support

hate 0.17 0.80 0.27 124

offensive 0.67 0.64 0.66 164

neither 0.94 0.43 0.59 865

accuracy 0.50 1153

macro avg 0.59 0.62 0.51 1153

weighted avg 0.82 0.50 0.57 1153

average eval\_loss: 2.185472103028462, accuracy: 0.5021682567215958

training 9 epoch...

114.64403343200684 seconds used......

9 training finished! average train loss: 0.13714035461834145

total clf loss: 15.122235467657447 total adv loss: 3.9586721493687946

evaluating...

Confusion matrix, without normalization

[[ 91 23 10]

[ 45 105 14]

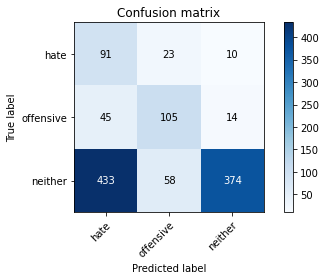
[433 58 374]]

Normalized confusion matrix

[[0.73387097 0.18548387 0.08064516]

[0.27439024 0.6402439 0.08536585]

[0.50057803 0.06705202 0.43236994]]





precision recall f1-score support

hate 0.16 0.73 0.26 124

offensive 0.56 0.64 0.60 164

neither 0.94 0.43 0.59 865

accuracy 0.49 1153

macro avg 0.55 0.60 0.48 1153

weighted avg 0.80 0.49 0.56 1153

average eval\_loss: 2.5752912800887535, accuracy: 0.4943625325238508

training 10 epoch...

114.6185610294342 seconds used......

10 training finished! average train loss: 0.15579396508456694

total clf loss: 15.067454038187861 total adv loss: 5.5529660894098924

evaluating...

Confusion matrix, without normalization

[[ 87 29 8]

[ 38 116 10]

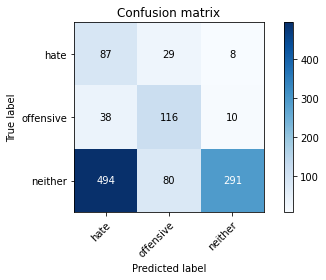
[494 80 291]]

Normalized confusion matrix

[[0.7016129 0.23387097 0.06451613]

[0.23170732 0.70731707 0.06097561]

[0.57109827 0.09248555 0.33641618]]





precision recall f1-score support

hate 0.14 0.70 0.23 124

offensive 0.52 0.71 0.60 164

neither 0.94 0.34 0.50 865

accuracy 0.43 1153

macro avg 0.53 0.58 0.44 1153

weighted avg 0.79 0.43 0.48 1153

average eval\_loss: 2.8851250368973305, accuracy: 0.4284475281873374

2 run......

training 1 epoch...

114.72112035751343 seconds used......

1 training finished! average train loss: 2.032946839928627

total clf loss: 164.79608017206192 total adv loss: 88.369493663311

evaluating...

Confusion matrix, without normalization

[[ 95 26 3]

[ 51 112 1]

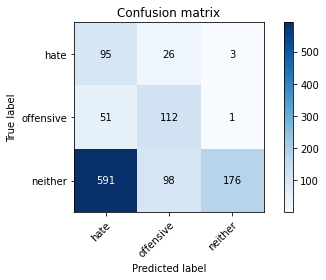
[591 98 176]]

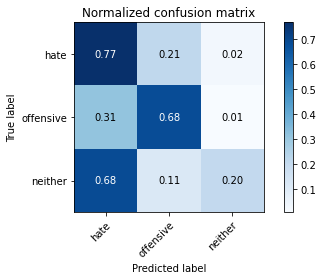
Normalized confusion matrix

[[0.76612903 0.20967742 0.02419355]

[0.31097561 0.68292683 0.00609756]

[0.68323699 0.1132948 0.20346821]]





precision recall f1-score support

hate 0.13 0.77 0.22 124

offensive 0.47 0.68 0.56 164

neither 0.98 0.20 0.34 865

accuracy 0.33 1153

macro avg 0.53 0.55 0.37 1153

weighted avg 0.81 0.33 0.36 1153

average eval\_loss: 0.9891979314129928, accuracy: 0.33217692974848223

training 2 epoch...

114.55437803268433 seconds used......

2 training finished! average train loss: 1.699517880876859

total clf loss: 142.39140355587006 total adv loss: 71.56379996240139

evaluating...

Confusion matrix, without normalization

[[ 97 24 3]

[ 41 122 1]

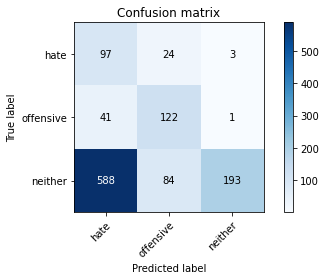
[588 84 193]]

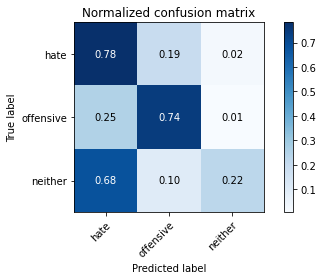
Normalized confusion matrix

[[0.78225806 0.19354839 0.02419355]

[0.25 0.74390244 0.00609756]

[0.67976879 0.09710983 0.22312139]]





precision recall f1-score support

hate 0.13 0.78 0.23 124

offensive 0.53 0.74 0.62 164

neither 0.98 0.22 0.36 865

accuracy 0.36 1153

macro avg 0.55 0.58 0.40 1153

weighted avg 0.82 0.36 0.39 1153

average eval\_loss: 1.0381282968767758, accuracy: 0.35732870771899394

training 3 epoch...

114.44392967224121 seconds used......

3 training finished! average train loss: 1.290484189632393

total clf loss: 116.894201785326 total adv loss: 49.95357118546963

evaluating...

Confusion matrix, without normalization

[[ 81 19 24]

[ 44 107 13]

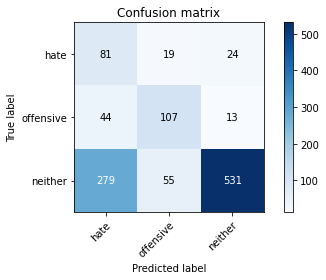
[279 55 531]]

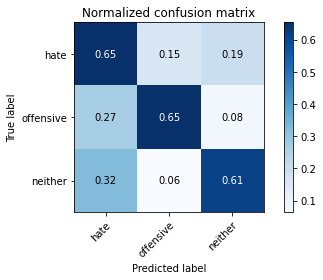
Normalized confusion matrix

[[0.65322581 0.15322581 0.19354839]

[0.26829268 0.65243902 0.07926829]

[0.32254335 0.06358382 0.61387283]]





precision recall f1-score support

hate 0.20 0.65 0.31 124

offensive 0.59 0.65 0.62 164

neither 0.93 0.61 0.74 865

accuracy 0.62 1153

macro avg 0.58 0.64 0.56 1153

weighted avg 0.81 0.62 0.68 1153

average eval\_loss: 0.8347561657428741, accuracy: 0.6235906331309627

training 4 epoch...

114.51472210884094 seconds used......

4 training finished! average train loss: 0.8715090533452374

total clf loss: 90.5372421592474 total adv loss: 27.93813903350383

evaluating...

Confusion matrix, without normalization

[[ 95 19 10]

[ 40 117 7]

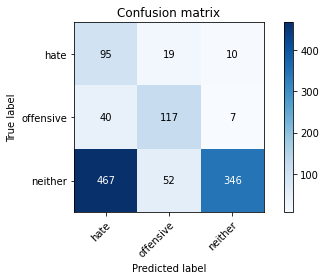
[467 52 346]]

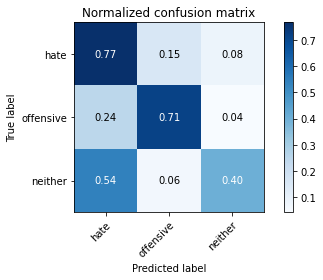
Normalized confusion matrix

[[0.76612903 0.15322581 0.08064516]

[0.24390244 0.71341463 0.04268293]

[0.53988439 0.06011561 0.4 ]]





precision recall f1-score support

hate 0.16 0.77 0.26 124

offensive 0.62 0.71 0.66 164

neither 0.95 0.40 0.56 865

accuracy 0.48 1153

macro avg 0.58 0.63 0.50 1153

weighted avg 0.82 0.48 0.55 1153

average eval\_loss: 1.1473097256545364, accuracy: 0.4839549002601908

training 5 epoch...

114.4771044254303 seconds used......

5 training finished! average train loss: 0.6055389179715088

total clf loss: 69.01156596839428 total adv loss: 16.359486249275506

evaluating...

Confusion matrix, without normalization

[[ 97 17 10]

[ 57 98 9]

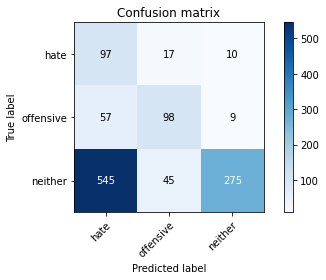
[545 45 275]]

Normalized confusion matrix

[[0.78225806 0.13709677 0.08064516]

[0.34756098 0.59756098 0.05487805]

[0.6300578 0.05202312 0.31791908]]





precision recall f1-score support

hate 0.14 0.78 0.24 124

offensive 0.61 0.60 0.60 164

neither 0.94 0.32 0.47 865

accuracy 0.41 1153

macro avg 0.56 0.57 0.44 1153

weighted avg 0.80 0.41 0.47 1153

average eval\_loss: 1.4038378051642715, accuracy: 0.40763226366001737

training 6 epoch...

114.45639681816101 seconds used......

6 training finished! average train loss: 0.37488656632957007

total clf loss: 47.95242000371218 total adv loss: 7.5142615704244236

evaluating...

Confusion matrix, without normalization

[[ 84 34 6]

[ 38 121 5]

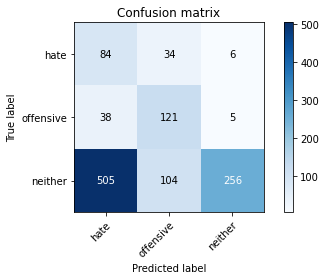
[505 104 256]]

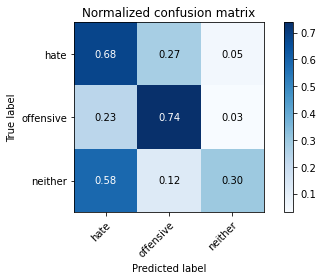
Normalized confusion matrix

[[0.67741935 0.27419355 0.0483871 ]

[0.23170732 0.73780488 0.0304878 ]

[0.58381503 0.12023121 0.29595376]]





precision recall f1-score support

hate 0.13 0.68 0.22 124

offensive 0.47 0.74 0.57 164

neither 0.96 0.30 0.45 865

accuracy 0.40 1153

macro avg 0.52 0.57 0.42 1153

weighted avg 0.80 0.40 0.44 1153

average eval\_loss: 1.9251970155485745, accuracy: 0.3998265394622723

training 7 epoch...

114.5150101184845 seconds used......

7 training finished! average train loss: 0.27622093650556745

total clf loss: 33.958788465708494 total adv loss: 6.223164447321324

evaluating...

Confusion matrix, without normalization

[[ 86 28 10]

[ 40 117 7]

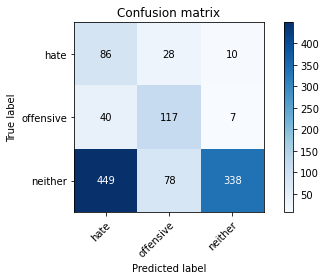
[449 78 338]]

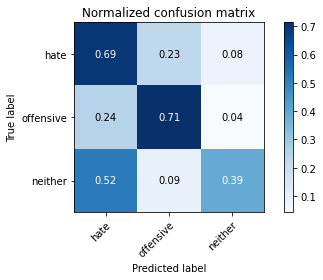
Normalized confusion matrix

[[0.69354839 0.22580645 0.08064516]

[0.24390244 0.71341463 0.04268293]

[0.51907514 0.09017341 0.39075145]]





precision recall f1-score support

hate 0.15 0.69 0.25 124

offensive 0.52 0.71 0.60 164

neither 0.95 0.39 0.55 865

accuracy 0.47 1153

macro avg 0.54 0.60 0.47 1153

weighted avg 0.81 0.47 0.53 1153

average eval\_loss: 1.9089251797774742, accuracy: 0.4692107545533391

training 8 epoch...

114.59400200843811 seconds used......

8 training finished! average train loss: 0.156256272391017

total clf loss: 20.42914334964007 total adv loss: 2.9109552445661393

evaluating...

Confusion matrix, without normalization

[[ 89 20 15]

[ 44 113 7]

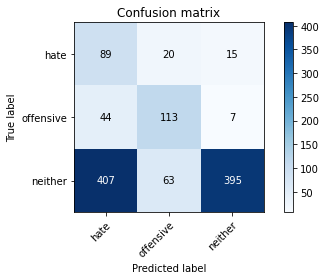
[407 63 395]]

Normalized confusion matrix

[[0.71774194 0.16129032 0.12096774]

[0.26829268 0.68902439 0.04268293]

[0.47052023 0.07283237 0.4566474 ]]





precision recall f1-score support

hate 0.16 0.72 0.27 124

offensive 0.58 0.69 0.63 164

neither 0.95 0.46 0.62 865

accuracy 0.52 1153

macro avg 0.56 0.62 0.50 1153

weighted avg 0.81 0.52 0.58 1153

average eval\_loss: 1.8985302951829186, accuracy: 0.5177797051170858

training 9 epoch...

114.54265570640564 seconds used......

9 training finished! average train loss: 0.11438803729556855

total clf loss: 13.791463830508292 total adv loss: 2.712863270444359

evaluating...

Confusion matrix, without normalization

[[ 85 23 16]

[ 42 115 7]

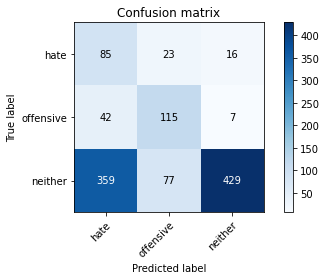
[359 77 429]]

Normalized confusion matrix

[[0.68548387 0.18548387 0.12903226]

[0.25609756 0.70121951 0.04268293]

[0.4150289 0.08901734 0.49595376]]





precision recall f1-score support

hate 0.17 0.69 0.28 124

offensive 0.53 0.70 0.61 164

neither 0.95 0.50 0.65 865

accuracy 0.55 1153

macro avg 0.55 0.63 0.51 1153

weighted avg 0.81 0.55 0.61 1153

average eval\_loss: 1.9552522353057205, accuracy: 0.5455333911535126

training 10 epoch...

114.61794543266296 seconds used......

10 training finished! average train loss: 0.11840104428397137

total clf loss: 12.245950459502637 total adv loss: 3.822712526694886

evaluating...

Confusion matrix, without normalization

[[ 84 24 16]

[ 38 118 8]

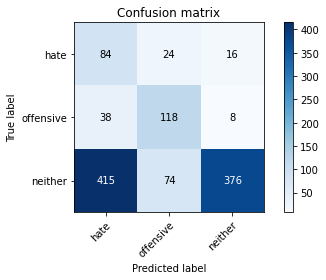
[415 74 376]]

Normalized confusion matrix

[[0.67741935 0.19354839 0.12903226]

[0.23170732 0.7195122 0.04878049]

[0.47976879 0.08554913 0.43468208]]





precision recall f1-score support

hate 0.16 0.68 0.25 124

offensive 0.55 0.72 0.62 164

neither 0.94 0.43 0.59 865

accuracy 0.50 1153

macro avg 0.55 0.61 0.49 1153

weighted avg 0.80 0.50 0.56 1153

average eval\_loss: 2.17541029227191, accuracy: 0.5013009540329575

3 run......

training 1 epoch...

114.5509090423584 seconds used......

1 training finished! average train loss: 2.040094188991047

total clf loss: 164.94796866178513 total adv loss: 88.89392775297165

evaluating...

Confusion matrix, without normalization

[[ 82 26 16]

[ 40 117 7]

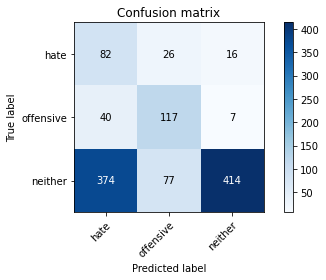
[374 77 414]]

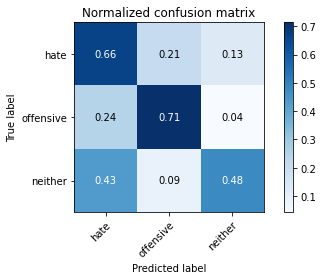
Normalized confusion matrix

[[0.66129032 0.20967742 0.12903226]

[0.24390244 0.71341463 0.04268293]

[0.43236994 0.08901734 0.47861272]]





precision recall f1-score support

hate 0.17 0.66 0.26 124

offensive 0.53 0.71 0.61 164

neither 0.95 0.48 0.64 865

accuracy 0.53 1153

macro avg 0.55 0.62 0.50 1153

weighted avg 0.80 0.53 0.59 1153

average eval\_loss: 0.8926727812865685, accuracy: 0.5316565481352992

training 2 epoch...

114.53317713737488 seconds used......

2 training finished! average train loss: 1.66597218811512

total clf loss: 139.97096887230873 total adv loss: 69.95617914199829

evaluating...

Confusion matrix, without normalization

[[ 96 21 7]

[ 51 110 3]

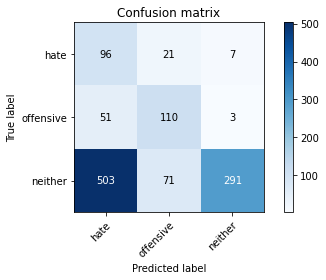
[503 71 291]]

Normalized confusion matrix

[[0.77419355 0.16935484 0.05645161]

[0.31097561 0.67073171 0.01829268]

[0.58150289 0.08208092 0.33641618]]





precision recall f1-score support

hate 0.15 0.77 0.25 124

offensive 0.54 0.67 0.60 164

neither 0.97 0.34 0.50 865

accuracy 0.43 1153

macro avg 0.55 0.59 0.45 1153

weighted avg 0.82 0.43 0.49 1153

average eval\_loss: 0.9557325058969958, accuracy: 0.43104943625325237

training 3 epoch...

114.55253267288208 seconds used......

3 training finished! average train loss: 1.3051618859171867

total clf loss: 117.6008768081665 total adv loss: 50.833159901201725

evaluating...

Confusion matrix, without normalization

[[ 91 18 15]

[ 45 105 14]

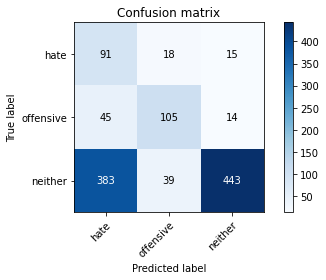
[383 39 443]]

Normalized confusion matrix

[[0.73387097 0.14516129 0.12096774]

[0.27439024 0.6402439 0.08536585]

[0.44277457 0.04508671 0.51213873]]





precision recall f1-score support

hate 0.18 0.73 0.28 124

offensive 0.65 0.64 0.64 164

neither 0.94 0.51 0.66 865

accuracy 0.55 1153

macro avg 0.59 0.63 0.53 1153

weighted avg 0.82 0.55 0.62 1153

average eval\_loss: 0.8672805654591528, accuracy: 0.5542064180398959

training 4 epoch...

114.50244069099426 seconds used......

4 training finished! average train loss: 0.8816092518113908

total clf loss: 88.87074336409569 total adv loss: 29.61980563402176

evaluating...

Confusion matrix, without normalization

[[ 85 24 15]

[ 37 115 12]

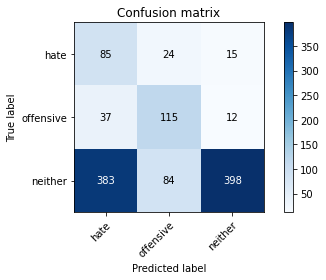
[383 84 398]]

Normalized confusion matrix

[[0.68548387 0.19354839 0.12096774]

[0.22560976 0.70121951 0.07317073]

[0.44277457 0.09710983 0.46011561]]





precision recall f1-score support

hate 0.17 0.69 0.27 124

offensive 0.52 0.70 0.59 164

neither 0.94 0.46 0.62 865

accuracy 0.52 1153

macro avg 0.54 0.62 0.49 1153

weighted avg 0.79 0.52 0.58 1153

average eval\_loss: 1.0766947947699448, accuracy: 0.5186470078057241

training 5 epoch...

114.57575607299805 seconds used......

5 training finished! average train loss: 0.5540625834394068

total clf loss: 63.51088182628155 total adv loss: 14.785816302290186

evaluating...

Confusion matrix, without normalization

[[ 94 22 8]

[ 40 117 7]

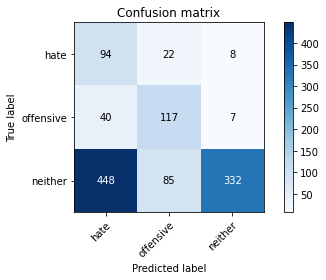
[448 85 332]]

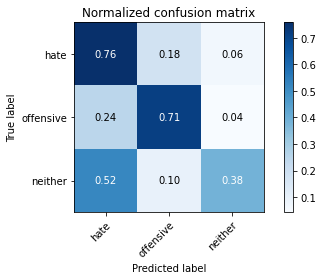
Normalized confusion matrix

[[0.75806452 0.17741935 0.06451613]

[0.24390244 0.71341463 0.04268293]

[0.51791908 0.0982659 0.38381503]]





precision recall f1-score support

hate 0.16 0.76 0.27 124

offensive 0.52 0.71 0.60 164

neither 0.96 0.38 0.55 865

accuracy 0.47 1153

macro avg 0.55 0.62 0.47 1153

weighted avg 0.81 0.47 0.53 1153

average eval\_loss: 1.4978711132345528, accuracy: 0.47094535993061576

training 6 epoch...

114.53816604614258 seconds used......

6 training finished! average train loss: 0.3670449746950042

total clf loss: 43.8540650382638 total adv loss: 8.904745412917691

evaluating...

Confusion matrix, without normalization

[[ 97 13 14]

[ 50 92 22]

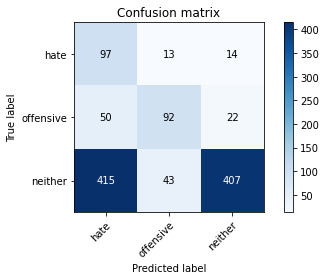
[415 43 407]]

Normalized confusion matrix

[[0.78225806 0.10483871 0.11290323]

[0.30487805 0.56097561 0.13414634]

[0.47976879 0.04971098 0.47052023]]





precision recall f1-score support

hate 0.17 0.78 0.28 124

offensive 0.62 0.56 0.59 164

neither 0.92 0.47 0.62 865

accuracy 0.52 1153

macro avg 0.57 0.60 0.50 1153

weighted avg 0.80 0.52 0.58 1153

average eval\_loss: 1.5371038327957023, accuracy: 0.5169124024284475

training 7 epoch...

114.61533188819885 seconds used......

7 training finished! average train loss: 0.22613243816331738

total clf loss: 27.20023643039167 total adv loss: 5.395006676655612

evaluating...

Confusion matrix, without normalization

[[ 91 20 13]

[ 41 113 10]

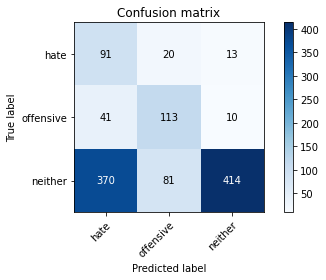
[370 81 414]]

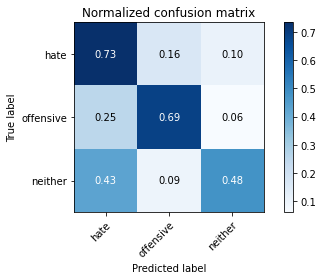
Normalized confusion matrix

[[0.73387097 0.16129032 0.10483871]

[0.25 0.68902439 0.06097561]

[0.42774566 0.09364162 0.47861272]]





precision recall f1-score support

hate 0.18 0.73 0.29 124

offensive 0.53 0.69 0.60 164

neither 0.95 0.48 0.64 865

accuracy 0.54 1153

macro avg 0.55 0.63 0.51 1153

weighted avg 0.81 0.54 0.59 1153

average eval\_loss: 1.7879368858090763, accuracy: 0.5359930615784909

training 8 epoch...

114.5340530872345 seconds used......

8 training finished! average train loss: 0.18435911443971453

total clf loss: 21.21163182798773 total adv loss: 4.880349740255042

evaluating...

Confusion matrix, without normalization

[[ 80 30 14]

[ 36 115 13]

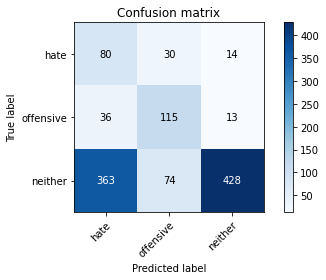
[363 74 428]]

Normalized confusion matrix

[[0.64516129 0.24193548 0.11290323]

[0.2195122 0.70121951 0.07926829]

[0.41965318 0.08554913 0.49479769]]





precision recall f1-score support

hate 0.17 0.65 0.27 124

offensive 0.53 0.70 0.60 164

neither 0.94 0.49 0.65 865

accuracy 0.54 1153

macro avg 0.54 0.61 0.50 1153

weighted avg 0.80 0.54 0.60 1153

average eval\_loss: 1.8467692773917626, accuracy: 0.5403295750216826

training 9 epoch...

114.49794006347656 seconds used......

9 training finished! average train loss: 0.14380594957730777

total clf loss: 14.64457167033106 total adv loss: 4.757413931933115

evaluating...

Confusion matrix, without normalization

[[ 81 31 12]

[ 30 126 8]

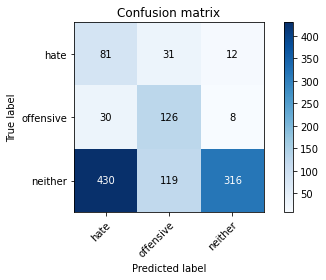
[430 119 316]]

Normalized confusion matrix

[[0.65322581 0.25 0.09677419]

[0.18292683 0.76829268 0.04878049]

[0.49710983 0.13757225 0.36531792]]





precision recall f1-score support

hate 0.15 0.65 0.24 124

offensive 0.46 0.77 0.57 164

neither 0.94 0.37 0.53 865

accuracy 0.45 1153

macro avg 0.52 0.60 0.45 1153

weighted avg 0.79 0.45 0.50 1153

average eval\_loss: 2.6156308465990525, accuracy: 0.4535993061578491

training 10 epoch...

114.54886102676392 seconds used......

10 training finished! average train loss: 0.10614105961230096

total clf loss: 10.00250406586565 total adv loss: 3.9145969197998056

evaluating...

Confusion matrix, without normalization

[[ 86 22 16]

[ 31 120 13]

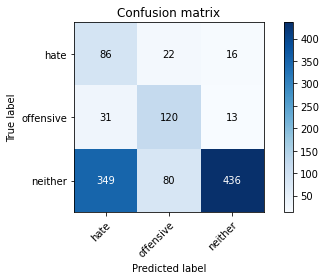
[349 80 436]]

Normalized confusion matrix

[[0.69354839 0.17741935 0.12903226]

[0.18902439 0.73170732 0.07926829]

[0.40346821 0.09248555 0.50404624]]





precision recall f1-score support

hate 0.18 0.69 0.29 124

offensive 0.54 0.73 0.62 164

neither 0.94 0.50 0.66 865

accuracy 0.56 1153

macro avg 0.55 0.64 0.52 1153

weighted avg 0.80 0.56 0.61 1153

average eval\_loss: 2.111476649498117, accuracy: 0.5568083261058109

4 run......

training 1 epoch...

114.58350920677185 seconds used......

1 training finished! average train loss: 2.026328279858544

total clf loss: 164.90441220998764 total adv loss: 87.75936861336231

evaluating...

Confusion matrix, without normalization

[[ 93 26 5]

[ 47 112 5]

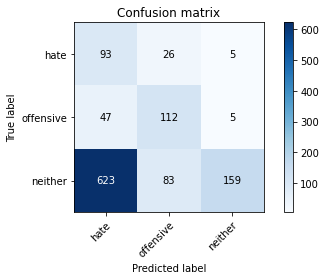
[623 83 159]]

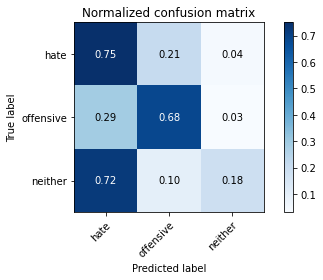
Normalized confusion matrix

[[0.75 0.20967742 0.04032258]

[0.28658537 0.68292683 0.0304878 ]

[0.72023121 0.09595376 0.18381503]]





precision recall f1-score support

hate 0.12 0.75 0.21 124

offensive 0.51 0.68 0.58 164

neither 0.94 0.18 0.31 865

accuracy 0.32 1153

macro avg 0.52 0.54 0.37 1153

weighted avg 0.79 0.32 0.34 1153

average eval\_loss: 0.9894700769720406, accuracy: 0.31569817866435385

training 2 epoch...

114.5764217376709 seconds used......

2 training finished! average train loss: 1.686427999820028

total clf loss: 142.4285079240799 total adv loss: 70.44569757580757

evaluating...

Confusion matrix, without normalization

[[ 90 33 1]

[ 37 126 1]

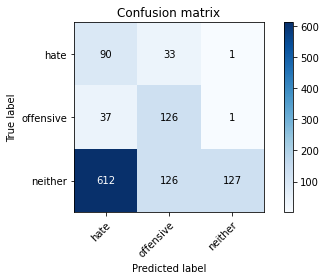
[612 126 127]]

Normalized confusion matrix

[[0.72580645 0.26612903 0.00806452]

[0.22560976 0.76829268 0.00609756]

[0.70751445 0.14566474 0.14682081]]





precision recall f1-score support

hate 0.12 0.73 0.21 124

offensive 0.44 0.77 0.56 164

neither 0.98 0.15 0.26 865

accuracy 0.30 1153

macro avg 0.52 0.55 0.34 1153

weighted avg 0.81 0.30 0.29 1153

average eval\_loss: 1.0928224366286705, accuracy: 0.2974848222029488

training 3 epoch...

114.53277373313904 seconds used......

3 training finished! average train loss: 1.3415589364511626

total clf loss: 120.46263536810875 total adv loss: 52.459632992744446

evaluating...

Confusion matrix, without normalization

[[ 78 34 12]

[ 34 119 11]

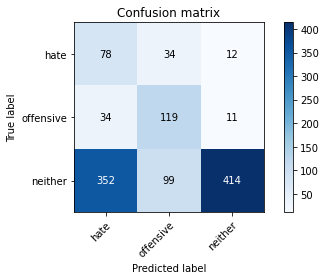
[352 99 414]]

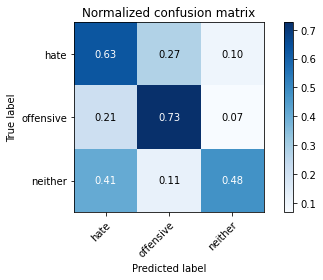
Normalized confusion matrix

[[0.62903226 0.27419355 0.09677419]

[0.20731707 0.72560976 0.06707317]

[0.40693642 0.11445087 0.47861272]]





precision recall f1-score support

hate 0.17 0.63 0.27 124

offensive 0.47 0.73 0.57 164

neither 0.95 0.48 0.64 865

accuracy 0.53 1153

macro avg 0.53 0.61 0.49 1153

weighted avg 0.80 0.53 0.59 1153

average eval\_loss: 0.9354192357638786, accuracy: 0.5299219427580225

training 4 epoch...

114.5626437664032 seconds used......

4 training finished! average train loss: 0.8759844301357156

total clf loss: 90.65427856147289 total adv loss: 28.255552885122597

evaluating...

Confusion matrix, without normalization

[[ 82 34 8]

[ 36 114 14]

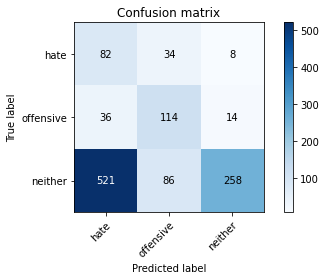
[521 86 258]]

Normalized confusion matrix

[[0.66129032 0.27419355 0.06451613]

[0.2195122 0.69512195 0.08536585]

[0.60231214 0.09942197 0.2982659 ]]





precision recall f1-score support

hate 0.13 0.66 0.21 124

offensive 0.49 0.70 0.57 164

neither 0.92 0.30 0.45 865

accuracy 0.39 1153

macro avg 0.51 0.55 0.41 1153

weighted avg 0.77 0.39 0.44 1153

average eval\_loss: 1.2721142635263245, accuracy: 0.39375542064180397

training 5 epoch...

114.50736355781555 seconds used......

5 training finished! average train loss: 0.5953521438475166

total clf loss: 68.09930609166622 total adv loss: 15.959927000571042

evaluating...

Confusion matrix, without normalization

[[ 86 29 9]

[ 31 123 10]

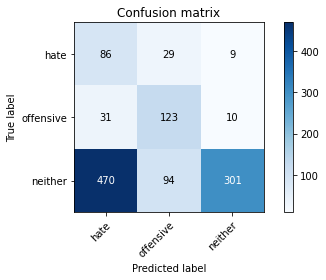
[470 94 301]]

Normalized confusion matrix

[[0.69354839 0.23387097 0.07258065]

[0.18902439 0.75 0.06097561]

[0.5433526 0.10867052 0.34797688]]





precision recall f1-score support

hate 0.15 0.69 0.24 124

offensive 0.50 0.75 0.60 164

neither 0.94 0.35 0.51 865

accuracy 0.44 1153

macro avg 0.53 0.60 0.45 1153

weighted avg 0.79 0.44 0.49 1153

average eval\_loss: 1.4797877700164401, accuracy: 0.44232437120555074

training 6 epoch...

114.52873945236206 seconds used......

6 training finished! average train loss: 0.421637333308657

total clf loss: 49.63259346783161 total adv loss: 10.601239201554563

evaluating...

Confusion matrix, without normalization

[[ 89 28 7]

[ 40 117 7]

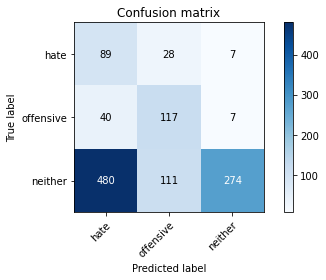
[480 111 274]]

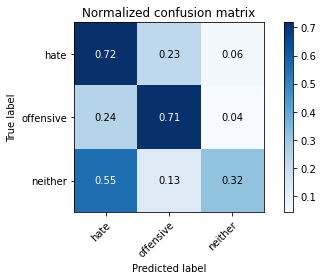
Normalized confusion matrix

[[0.71774194 0.22580645 0.05645161]

[0.24390244 0.71341463 0.04268293]

[0.55491329 0.1283237 0.31676301]]





precision recall f1-score support

hate 0.15 0.72 0.24 124

offensive 0.46 0.71 0.56 164

neither 0.95 0.32 0.48 865

accuracy 0.42 1153

macro avg 0.52 0.58 0.43 1153

weighted avg 0.79 0.42 0.46 1153

average eval\_loss: 2.018302154952082, accuracy: 0.4163052905464007

training 7 epoch...

114.60871577262878 seconds used......

7 training finished! average train loss: 0.24551821242840516

total clf loss: 31.804641380906105 total adv loss: 4.721209212708345

evaluating...

Confusion matrix, without normalization

[[ 93 18 13]

[ 37 111 16]

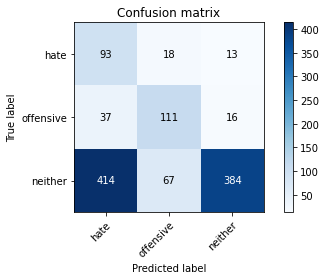
[414 67 384]]

Normalized confusion matrix

[[0.75 0.14516129 0.10483871]

[0.22560976 0.67682927 0.09756098]

[0.47861272 0.07745665 0.44393064]]





precision recall f1-score support

hate 0.17 0.75 0.28 124

offensive 0.57 0.68 0.62 164

neither 0.93 0.44 0.60 865

accuracy 0.51 1153

macro avg 0.56 0.62 0.50 1153

weighted avg 0.80 0.51 0.57 1153

average eval\_loss: 1.9197935071484795, accuracy: 0.5099739809193409

training 8 epoch...

114.78949022293091 seconds used......

8 training finished! average train loss: 0.16952299808950297

total clf loss: 21.368650311604142 total adv loss: 3.5556067079942295

evaluating...

Confusion matrix, without normalization

[[ 96 16 12]

[ 43 103 18]

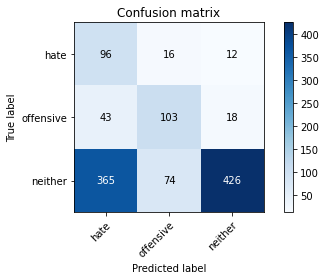
[365 74 426]]

Normalized confusion matrix

[[0.77419355 0.12903226 0.09677419]

[0.26219512 0.62804878 0.1097561 ]

[0.42196532 0.08554913 0.49248555]]





precision recall f1-score support

hate 0.19 0.77 0.31 124

offensive 0.53 0.63 0.58 164

neither 0.93 0.49 0.64 865

accuracy 0.54 1153

macro avg 0.55 0.63 0.51 1153

weighted avg 0.80 0.54 0.60 1153

average eval\_loss: 1.9495381692360187, accuracy: 0.5420641803989592

training 9 epoch...

114.7078857421875 seconds used......

9 training finished! average train loss: 0.12584930553012305

total clf loss: 14.421558442525566 total adv loss: 3.360562489390759

evaluating...

Confusion matrix, without normalization

[[ 94 21 9]

[ 38 115 11]

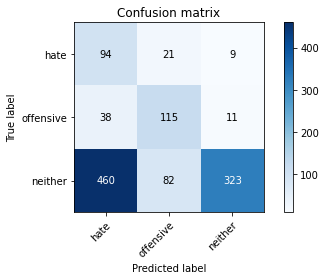
[460 82 323]]

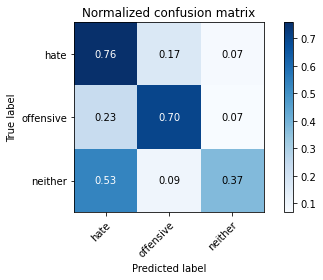
Normalized confusion matrix

[[0.75806452 0.16935484 0.07258065]

[0.23170732 0.70121951 0.06707317]

[0.53179191 0.09479769 0.3734104 ]]





precision recall f1-score support

hate 0.16 0.76 0.26 124

offensive 0.53 0.70 0.60 164

neither 0.94 0.37 0.53 865

accuracy 0.46 1153

macro avg 0.54 0.61 0.47 1153

weighted avg 0.80 0.46 0.52 1153

average eval\_loss: 2.6325297848931672, accuracy: 0.4614050303555941

training 10 epoch...

114.49547696113586 seconds used......

10 training finished! average train loss: 0.11811063719298025

total clf loss: 12.281156838638708 total adv loss: 3.7807151573761075

evaluating...

Confusion matrix, without normalization

[[ 93 20 11]

[ 49 100 15]

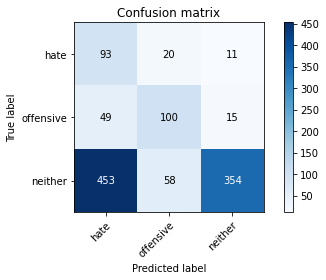
[453 58 354]]

Normalized confusion matrix

[[0.75 0.16129032 0.08870968]

[0.29878049 0.6097561 0.09146341]

[0.52369942 0.06705202 0.40924855]]





precision recall f1-score support

hate 0.16 0.75 0.26 124

offensive 0.56 0.61 0.58 164

neither 0.93 0.41 0.57 865

accuracy 0.47 1153

macro avg 0.55 0.59 0.47 1153

weighted avg 0.80 0.47 0.54 1153

average eval\_loss: 2.4475145422179123, accuracy: 0.4744145706851691

5 run......

training 1 epoch...

114.68809342384338 seconds used......

1 training finished! average train loss: 2.0410567131780444

total clf loss: 165.39491027593613 total adv loss: 88.75130864977837

evaluating...

Confusion matrix, without normalization

[[100 18 6]

[ 52 111 1]

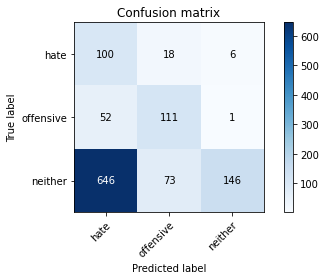
[646 73 146]]

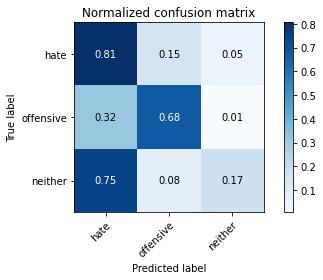
Normalized confusion matrix

[[0.80645161 0.14516129 0.0483871 ]

[0.31707317 0.67682927 0.00609756]

[0.74682081 0.08439306 0.16878613]]





precision recall f1-score support

hate 0.13 0.81 0.22 124

offensive 0.55 0.68 0.61 164

neither 0.95 0.17 0.29 865

accuracy 0.31 1153

macro avg 0.54 0.55 0.37 1153

weighted avg 0.81 0.31 0.32 1153

average eval\_loss: 0.9558369733136276, accuracy: 0.3096270598438855

training 2 epoch...

114.50062489509583 seconds used......

2 training finished! average train loss: 1.6989705520016807

total clf loss: 142.42284733057022 total adv loss: 71.50210273265839

evaluating...

Confusion matrix, without normalization

[[ 89 20 15]

[ 43 114 7]

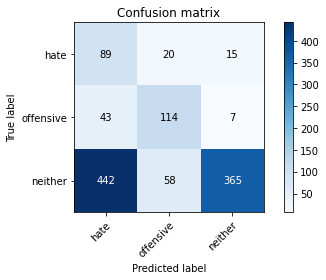
[442 58 365]]

Normalized confusion matrix

[[0.71774194 0.16129032 0.12096774]

[0.26219512 0.69512195 0.04268293]

[0.51098266 0.06705202 0.42196532]]





precision recall f1-score support

hate 0.16 0.72 0.26 124

offensive 0.59 0.70 0.64 164

neither 0.94 0.42 0.58 865

accuracy 0.49 1153

macro avg 0.56 0.61 0.49 1153

weighted avg 0.81 0.49 0.56 1153

average eval\_loss: 0.8842591345310211, accuracy: 0.49262792714657416

training 3 epoch...

114.66035842895508 seconds used......

3 training finished! average train loss: 1.3697040400334768

total clf loss: 122.36485573649406 total adv loss: 53.87271188944578

evaluating...

Confusion matrix, without normalization

[[ 79 31 14]

[ 39 118 7]

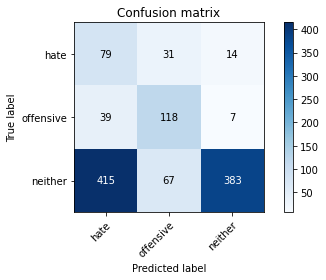
[415 67 383]]

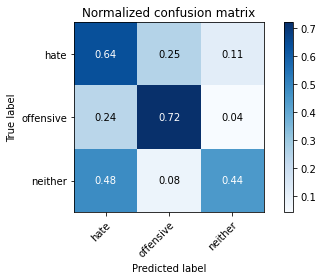
Normalized confusion matrix

[[0.63709677 0.25 0.11290323]

[0.23780488 0.7195122 0.04268293]

[0.47976879 0.07745665 0.44277457]]





precision recall f1-score support

hate 0.15 0.64 0.24 124

offensive 0.55 0.72 0.62 164

neither 0.95 0.44 0.60 865

accuracy 0.50 1153

macro avg 0.55 0.60 0.49 1153

weighted avg 0.80 0.50 0.57 1153

average eval\_loss: 0.9061302172726599, accuracy: 0.5030355594102341

training 4 epoch...

114.64798855781555 seconds used......

4 training finished! average train loss: 0.9412947960552716

total clf loss: 95.61694601178169 total adv loss: 31.26028998894617

evaluating...

Confusion matrix, without normalization

[[105 8 11]

[ 67 82 15]

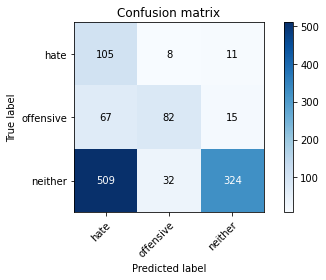
[509 32 324]]

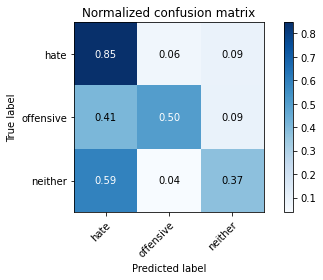
Normalized confusion matrix

[[0.84677419 0.06451613 0.08870968]

[0.40853659 0.5 0.09146341]

[0.58843931 0.03699422 0.37456647]]





precision recall f1-score support

hate 0.15 0.85 0.26 124

offensive 0.67 0.50 0.57 164

neither 0.93 0.37 0.53 865

accuracy 0.44 1153

macro avg 0.58 0.57 0.46 1153

weighted avg 0.81 0.44 0.51 1153

average eval\_loss: 1.1560435171785026, accuracy: 0.44319167389418906

training 5 epoch...

114.68313789367676 seconds used......

5 training finished! average train loss: 0.6085706430354288

total clf loss: 71.26427172124386 total adv loss: 15.487798060523346

evaluating...

Confusion matrix, without normalization

[[ 73 29 22]

[ 37 109 18]

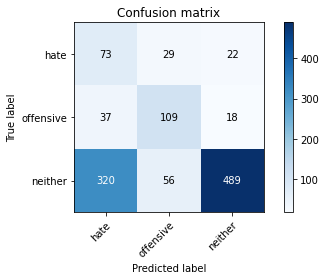
[320 56 489]]

Normalized confusion matrix

[[0.58870968 0.23387097 0.17741935]

[0.22560976 0.66463415 0.1097561 ]

[0.3699422 0.06473988 0.56531792]]





precision recall f1-score support

hate 0.17 0.59 0.26 124

offensive 0.56 0.66 0.61 164

neither 0.92 0.57 0.70 865

accuracy 0.58 1153

macro avg 0.55 0.61 0.52 1153

weighted avg 0.79 0.58 0.64 1153

average eval\_loss: 1.0754659535556004, accuracy: 0.5819601040763226

training 6 epoch...

114.60984754562378 seconds used......

6 training finished! average train loss: 0.40081636749562766

total clf loss: 50.75151140987873 total adv loss: 8.292819142196095

evaluating...

Confusion matrix, without normalization

[[ 80 21 23]

[ 45 99 20]

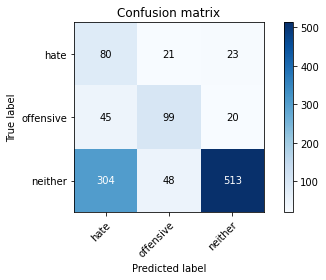
[304 48 513]]

Normalized confusion matrix

[[0.64516129 0.16935484 0.18548387]

[0.27439024 0.60365854 0.12195122]

[0.35144509 0.05549133 0.59306358]]





precision recall f1-score support

hate 0.19 0.65 0.29 124

offensive 0.59 0.60 0.60 164

neither 0.92 0.59 0.72 865

accuracy 0.60 1153

macro avg 0.57 0.61 0.54 1153

weighted avg 0.80 0.60 0.66 1153

average eval\_loss: 1.2490207943423042, accuracy: 0.6001734605377277

training 7 epoch...

114.62884950637817 seconds used......

7 training finished! average train loss: 0.30077141758409287

total clf loss: 34.940955236554146 total adv loss: 7.794321412045974

evaluating...

Confusion matrix, without normalization

[[ 74 40 10]

[ 32 125 7]

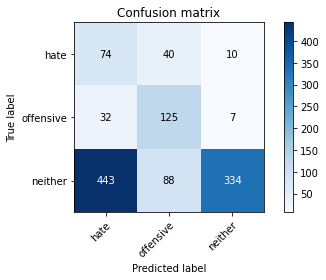
[443 88 334]]

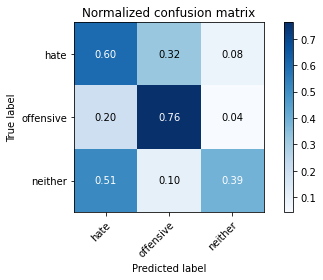
Normalized confusion matrix

[[0.59677419 0.32258065 0.08064516]

[0.19512195 0.76219512 0.04268293]

[0.51213873 0.1017341 0.38612717]]





precision recall f1-score support

hate 0.13 0.60 0.22 124

offensive 0.49 0.76 0.60 164

neither 0.95 0.39 0.55 865

accuracy 0.46 1153

macro avg 0.53 0.58 0.46 1153

weighted avg 0.80 0.46 0.52 1153

average eval\_loss: 1.982189499098679, accuracy: 0.46227233304423243

training 8 epoch...

114.58327555656433 seconds used......

8 training finished! average train loss: 0.19086608319498954

total clf loss: 21.94289688579738 total adv loss: 5.061302539106691

evaluating...

Confusion matrix, without normalization

[[104 11 9]

[ 62 91 11]

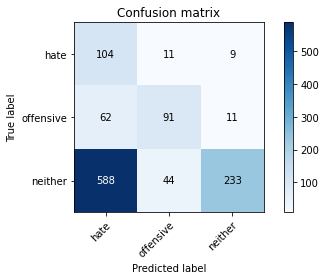
[588 44 233]]

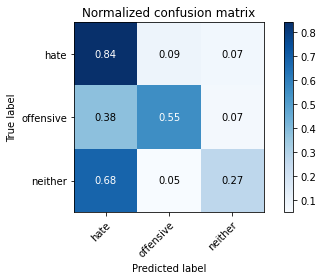
Normalized confusion matrix

[[0.83870968 0.08870968 0.07258065]

[0.37804878 0.55487805 0.06707317]

[0.67976879 0.05086705 0.26936416]]





precision recall f1-score support

hate 0.14 0.84 0.24 124

offensive 0.62 0.55 0.59 164

neither 0.92 0.27 0.42 865

accuracy 0.37 1153

macro avg 0.56 0.55 0.41 1153

weighted avg 0.79 0.37 0.42 1153

average eval\_loss: 2.6831710934638977, accuracy: 0.3712055507372073

training 9 epoch...

114.7206380367279 seconds used......

9 training finished! average train loss: 0.13004096099896179

total clf loss: 15.615692295134068 total adv loss: 3.115594556441465

evaluating...

Confusion matrix, without normalization

[[ 80 31 13]

[ 39 114 11]

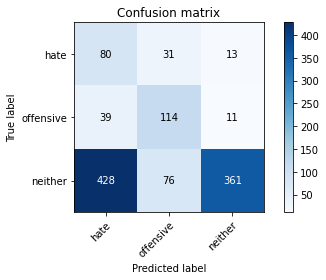
[428 76 361]]

Normalized confusion matrix

[[0.64516129 0.25 0.10483871]

[0.23780488 0.69512195 0.06707317]

[0.49479769 0.08786127 0.41734104]]





precision recall f1-score support

hate 0.15 0.65 0.24 124

offensive 0.52 0.70 0.59 164

neither 0.94 0.42 0.58 865

accuracy 0.48 1153

macro avg 0.53 0.59 0.47 1153

weighted avg 0.79 0.48 0.54 1153

average eval\_loss: 2.4098214137143104, accuracy: 0.4813529921942758

training 10 epoch...

114.93237709999084 seconds used......

10 training finished! average train loss: 0.09273843707550052

total clf loss: 10.2876127040945 total adv loss: 2.6462223193893806

evaluating...

Confusion matrix, without normalization

[[ 75 23 26]

[ 38 106 20]

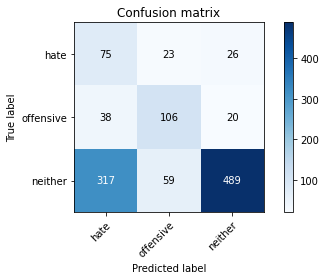
[317 59 489]]

Normalized confusion matrix

[[0.60483871 0.18548387 0.20967742]

[0.23170732 0.64634146 0.12195122]

[0.36647399 0.06820809 0.56531792]]





precision recall f1-score support

hate 0.17 0.60 0.27 124

offensive 0.56 0.65 0.60 164

neither 0.91 0.57 0.70 865

accuracy 0.58 1153

macro avg 0.55 0.61 0.52 1153

weighted avg 0.78 0.58 0.64 1153

average eval\_loss: 1.9167790104602944, accuracy: 0.5810928013876843