alpha=0.05

1 run......

Downloading: 100%

411/411 [00:00<00:00, 14.7kB/s]

Downloading: 100%

263M/263M [00:06<00:00, 41.3MB/s]

training 1 epoch...

115.54195046424866 seconds used......

1 training finished! average train loss: 0.9684507126609484

total clf loss: 167.85964113473892 total adv loss: 103.19843098521233

evaluating...

Confusion matrix, without normalization

[[ 62 50 12]

[ 33 118 13]

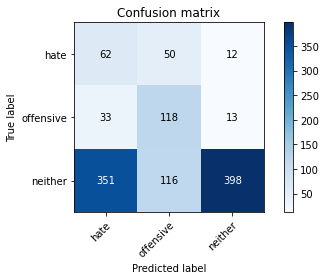
[351 116 398]]

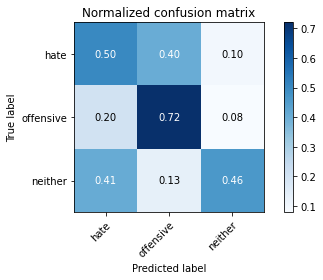
Normalized confusion matrix

[[0.5 0.40322581 0.09677419]

[0.20121951 0.7195122 0.07926829]

[0.40578035 0.13410405 0.46011561]]





precision recall f1-score support

hate 0.14 0.50 0.22 124

offensive 0.42 0.72 0.53 164

neither 0.94 0.46 0.62 865

accuracy 0.50 1153

macro avg 0.50 0.56 0.45 1153

weighted avg 0.78 0.50 0.56 1153

average eval\_loss: 0.9455026244295055, accuracy: 0.5013009540329575

training 2 epoch...

114.78833842277527 seconds used......

2 training finished! average train loss: 0.8266617311608224

total clf loss: 143.5467083454132 total adv loss: 93.3507604598999

evaluating...

Confusion matrix, without normalization

[[ 70 31 23]

[ 33 117 14]

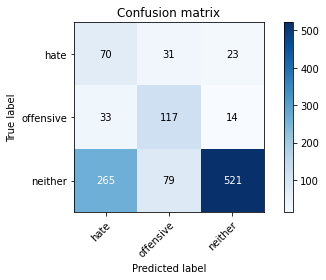
[265 79 521]]

Normalized confusion matrix

[[0.56451613 0.25 0.18548387]

[0.20121951 0.71341463 0.08536585]

[0.30635838 0.09132948 0.60231214]]





precision recall f1-score support

hate 0.19 0.56 0.28 124

offensive 0.52 0.71 0.60 164

neither 0.93 0.60 0.73 865

accuracy 0.61 1153

macro avg 0.55 0.63 0.54 1153

weighted avg 0.79 0.61 0.67 1153

average eval\_loss: 0.831448015467874, accuracy: 0.614050303555941

training 3 epoch...

114.99815964698792 seconds used......

3 training finished! average train loss: 0.6708195716852233

total clf loss: 116.68914258480072 total adv loss: 79.82909382879734

evaluating...

Confusion matrix, without normalization

[[ 64 40 20]

[ 25 129 10]

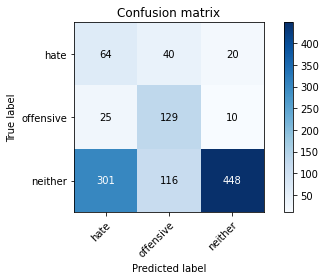
[301 116 448]]

Normalized confusion matrix

[[0.51612903 0.32258065 0.16129032]

[0.15243902 0.78658537 0.06097561]

[0.34797688 0.13410405 0.51791908]]





precision recall f1-score support

hate 0.16 0.52 0.25 124

offensive 0.45 0.79 0.57 164

neither 0.94 0.52 0.67 865

accuracy 0.56 1153

macro avg 0.52 0.61 0.50 1153

weighted avg 0.79 0.56 0.61 1153

average eval\_loss: 0.897761142459409, accuracy: 0.5559410234171726

training 4 epoch...

115.07660961151123 seconds used......

4 training finished! average train loss: 0.4327858179985058

total clf loss: 75.50366067886353 total adv loss: 55.9128705188632

evaluating...

Confusion matrix, without normalization

[[ 63 49 12]

[ 27 128 9]

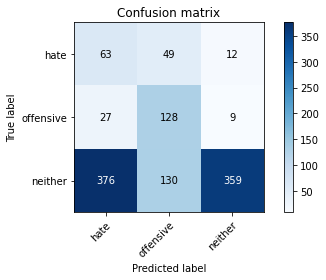
[376 130 359]]

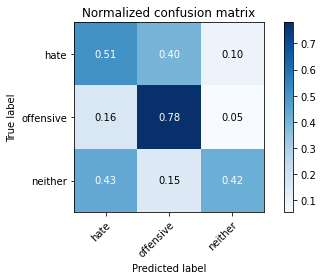
Normalized confusion matrix

[[0.50806452 0.39516129 0.09677419]

[0.16463415 0.7804878 0.05487805]

[0.43468208 0.15028902 0.4150289 ]]





precision recall f1-score support

hate 0.14 0.51 0.21 124

offensive 0.42 0.78 0.54 164

neither 0.94 0.42 0.58 865

accuracy 0.48 1153

macro avg 0.50 0.57 0.44 1153

weighted avg 0.78 0.48 0.53 1153

average eval\_loss: 1.2895860558953778, accuracy: 0.4770164787510841

training 5 epoch...

115.02832269668579 seconds used......

5 training finished! average train loss: 0.23031455171959742

total clf loss: 40.154483024030924 total adv loss: 29.23276569508016

evaluating...

Confusion matrix, without normalization

[[ 93 19 12]

[ 49 106 9]

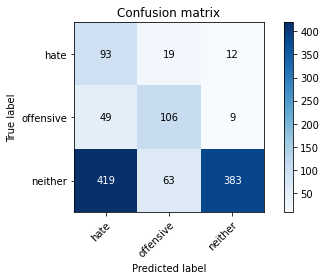
[419 63 383]]

Normalized confusion matrix

[[0.75 0.15322581 0.09677419]

[0.29878049 0.64634146 0.05487805]

[0.48439306 0.07283237 0.44277457]]





precision recall f1-score support

hate 0.17 0.75 0.27 124

offensive 0.56 0.65 0.60 164

neither 0.95 0.44 0.60 865

accuracy 0.50 1153

macro avg 0.56 0.61 0.49 1153

weighted avg 0.81 0.50 0.57 1153

average eval\_loss: 1.6021652499149586, accuracy: 0.5047701647875108

training 6 epoch...

114.96877527236938 seconds used......

6 training finished! average train loss: 0.11060818962730645

total clf loss: 19.2911744043231 total adv loss: 14.17997151601594

evaluating...

Confusion matrix, without normalization

[[ 92 17 15]

[ 47 103 14]

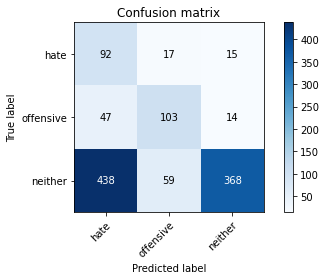
[438 59 368]]

Normalized confusion matrix

[[0.74193548 0.13709677 0.12096774]

[0.28658537 0.62804878 0.08536585]

[0.50635838 0.06820809 0.42543353]]





precision recall f1-score support

hate 0.16 0.74 0.26 124

offensive 0.58 0.63 0.60 164

neither 0.93 0.43 0.58 865

accuracy 0.49 1153

macro avg 0.55 0.60 0.48 1153

weighted avg 0.79 0.49 0.55 1153

average eval\_loss: 1.951048717416566, accuracy: 0.48829141370338247

training 7 epoch...

115.08650135993958 seconds used......

7 training finished! average train loss: 0.07695338791624333

total clf loss: 13.431994144339114 total adv loss: 10.076499323127791

evaluating...

Confusion matrix, without normalization

[[ 94 16 14]

[ 50 98 16]

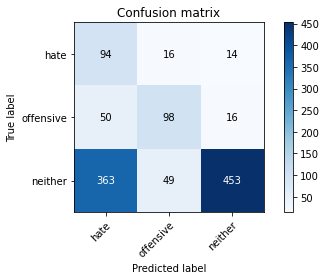
[363 49 453]]

Normalized confusion matrix

[[0.75806452 0.12903226 0.11290323]

[0.30487805 0.59756098 0.09756098]

[0.41965318 0.0566474 0.52369942]]





precision recall f1-score support

hate 0.19 0.76 0.30 124

offensive 0.60 0.60 0.60 164

neither 0.94 0.52 0.67 865

accuracy 0.56 1153

macro avg 0.57 0.63 0.52 1153

weighted avg 0.81 0.56 0.62 1153

average eval\_loss: 1.8352769387179408, accuracy: 0.559410234171726

training 8 epoch...

115.06198763847351 seconds used......

8 training finished! average train loss: 0.06009545722571645

total clf loss: 10.529715293319896 total adv loss: 8.673569480888546

evaluating...

Confusion matrix, without normalization

[[ 99 12 13]

[ 53 100 11]

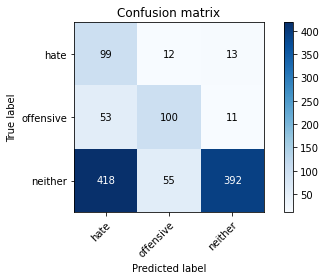
[418 55 392]]

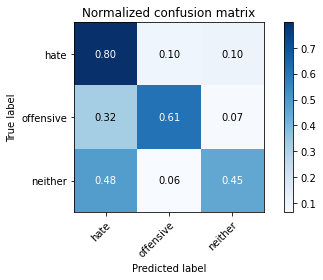
Normalized confusion matrix

[[0.7983871 0.09677419 0.10483871]

[0.32317073 0.6097561 0.06707317]

[0.48323699 0.06358382 0.45317919]]





precision recall f1-score support

hate 0.17 0.80 0.29 124

offensive 0.60 0.61 0.60 164

neither 0.94 0.45 0.61 865

accuracy 0.51 1153

macro avg 0.57 0.62 0.50 1153

weighted avg 0.81 0.51 0.58 1153

average eval\_loss: 2.280013308442872, accuracy: 0.5125758889852559

training 9 epoch...

115.10627293586731 seconds used......

9 training finished! average train loss: 0.048461790720466524

total clf loss: 8.470037260325626 total adv loss: 6.5691278919548495

evaluating...

Confusion matrix, without normalization

[[ 82 30 12]

[ 34 119 11]

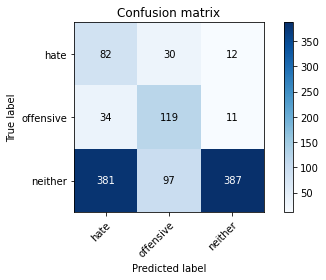
[381 97 387]]

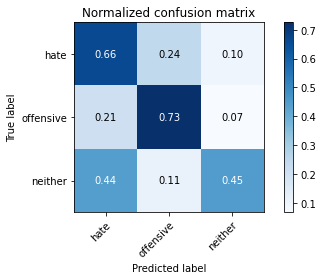
Normalized confusion matrix

[[0.66129032 0.24193548 0.09677419]

[0.20731707 0.72560976 0.06707317]

[0.44046243 0.11213873 0.44739884]]





precision recall f1-score support

hate 0.16 0.66 0.26 124

offensive 0.48 0.73 0.58 164

neither 0.94 0.45 0.61 865

accuracy 0.51 1153

macro avg 0.53 0.61 0.48 1153

weighted avg 0.79 0.51 0.57 1153

average eval\_loss: 2.526641916612099, accuracy: 0.5099739809193409

training 10 epoch...

115.05857753753662 seconds used......

10 training finished! average train loss: 0.06333584673328525

total clf loss: 11.14450944319833 total adv loss: 10.08174444094766

evaluating...

Confusion matrix, without normalization

[[ 76 24 24]

[ 42 103 19]

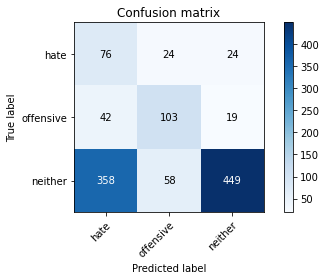
[358 58 449]]

Normalized confusion matrix

[[0.61290323 0.19354839 0.19354839]

[0.25609756 0.62804878 0.11585366]

[0.41387283 0.06705202 0.51907514]]





precision recall f1-score support

hate 0.16 0.61 0.25 124

offensive 0.56 0.63 0.59 164

neither 0.91 0.52 0.66 865

accuracy 0.54 1153

macro avg 0.54 0.59 0.50 1153

weighted avg 0.78 0.54 0.61 1153

average eval\_loss: 1.8834614887319763, accuracy: 0.5446660884648742

2 run......

training 1 epoch...

115.13926792144775 seconds used......

1 training finished! average train loss: 0.9839748598280407

total clf loss: 170.49233442544937 total adv loss: 103.69116815924644

evaluating...

Confusion matrix, without normalization

[[ 83 33 8]

[ 46 113 5]

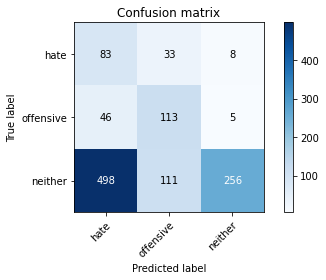
[498 111 256]]

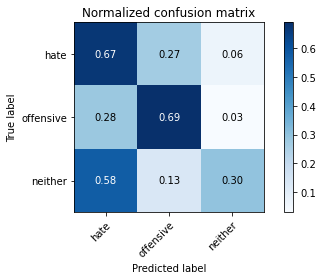
Normalized confusion matrix

[[0.66935484 0.26612903 0.06451613]

[0.2804878 0.68902439 0.0304878 ]

[0.57572254 0.1283237 0.29595376]]





precision recall f1-score support

hate 0.13 0.67 0.22 124

offensive 0.44 0.69 0.54 164

neither 0.95 0.30 0.45 865

accuracy 0.39 1153

macro avg 0.51 0.55 0.40 1153

weighted avg 0.79 0.39 0.44 1153

average eval\_loss: 0.9940552701210154, accuracy: 0.3920208152645273

training 2 epoch...

115.19846796989441 seconds used......

2 training finished! average train loss: 0.8369698648651441

total clf loss: 145.38150811195374 total adv loss: 95.41141450405121

evaluating...

Confusion matrix, without normalization

[[ 98 17 9]

[ 54 104 6]

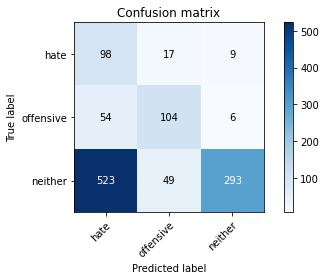
[523 49 293]]

Normalized confusion matrix

[[0.79032258 0.13709677 0.07258065]

[0.32926829 0.63414634 0.03658537]

[0.60462428 0.0566474 0.33872832]]





precision recall f1-score support

hate 0.15 0.79 0.25 124

offensive 0.61 0.63 0.62 164

neither 0.95 0.34 0.50 865

accuracy 0.43 1153

macro avg 0.57 0.59 0.46 1153

weighted avg 0.82 0.43 0.49 1153

average eval\_loss: 0.9288980138712916, accuracy: 0.4293148308759757

training 3 epoch...

115.13748240470886 seconds used......

3 training finished! average train loss: 0.7096443392691159

total clf loss: 123.47619894146919 total adv loss: 85.11899399757385

evaluating...

Confusion matrix, without normalization

[[ 96 17 11]

[ 50 110 4]

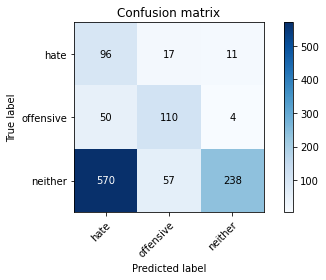
[570 57 238]]

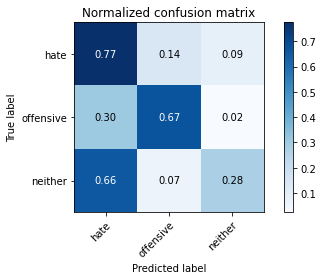
Normalized confusion matrix

[[0.77419355 0.13709677 0.08870968]

[0.30487805 0.67073171 0.02439024]

[0.65895954 0.06589595 0.27514451]]





precision recall f1-score support

hate 0.13 0.77 0.23 124

offensive 0.60 0.67 0.63 164

neither 0.94 0.28 0.43 865

accuracy 0.39 1153

macro avg 0.56 0.57 0.43 1153

weighted avg 0.81 0.39 0.43 1153

average eval\_loss: 1.059898653934742, accuracy: 0.38508239375542064

training 4 epoch...

115.10250163078308 seconds used......

4 training finished! average train loss: 0.4913114674744152

total clf loss: 85.78033992648125 total adv loss: 64.80027041584253

evaluating...

Confusion matrix, without normalization

[[ 84 39 1]

[ 43 115 6]

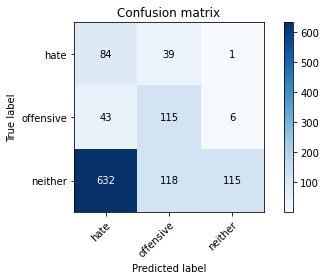
[632 118 115]]

Normalized confusion matrix

[[0.67741935 0.31451613 0.00806452]

[0.26219512 0.70121951 0.03658537]

[0.73063584 0.13641618 0.13294798]]





precision recall f1-score support

hate 0.11 0.68 0.19 124

offensive 0.42 0.70 0.53 164

neither 0.94 0.13 0.23 865

accuracy 0.27 1153

macro avg 0.49 0.50 0.32 1153

weighted avg 0.78 0.27 0.27 1153

average eval\_loss: 1.7657903679485978, accuracy: 0.2723330442324371

training 5 epoch...

115.07380270957947 seconds used......

5 training finished! average train loss: 0.25350961095786523

total clf loss: 44.35451990738511 total adv loss: 35.29810153134167

evaluating...

Confusion matrix, without normalization

[[ 91 19 14]

[ 46 105 13]

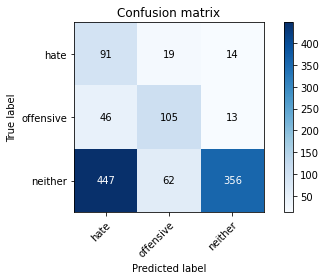
[447 62 356]]

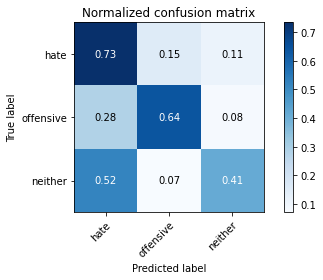
Normalized confusion matrix

[[0.73387097 0.15322581 0.11290323]

[0.2804878 0.6402439 0.07926829]

[0.51676301 0.0716763 0.41156069]]





precision recall f1-score support

hate 0.16 0.73 0.26 124

offensive 0.56 0.64 0.60 164

neither 0.93 0.41 0.57 865

accuracy 0.48 1153

macro avg 0.55 0.60 0.48 1153

weighted avg 0.79 0.48 0.54 1153

average eval\_loss: 1.6878395039459755, accuracy: 0.4787510841283608

training 6 epoch...

115.04976010322571 seconds used......

6 training finished! average train loss: 0.11754875291427154

total clf loss: 20.59087879769504 total adv loss: 16.853765568695962

evaluating...

Confusion matrix, without normalization

[[ 81 27 16]

[ 37 118 9]

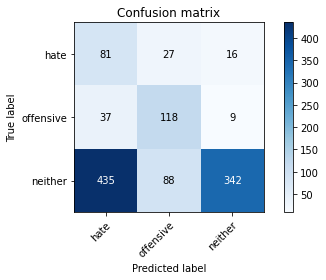
[435 88 342]]

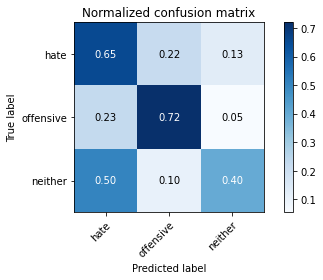
Normalized confusion matrix

[[0.65322581 0.21774194 0.12903226]

[0.22560976 0.7195122 0.05487805]

[0.50289017 0.1017341 0.39537572]]





precision recall f1-score support

hate 0.15 0.65 0.24 124

offensive 0.51 0.72 0.59 164

neither 0.93 0.40 0.56 865

accuracy 0.47 1153

macro avg 0.53 0.59 0.46 1153

weighted avg 0.79 0.47 0.53 1153

average eval\_loss: 1.9286449202175797, accuracy: 0.4692107545533391

training 7 epoch...

115.00120997428894 seconds used......

7 training finished! average train loss: 0.0731229269544461

total clf loss: 12.764817877206951 total adv loss: 9.603323542745784

evaluating...

Confusion matrix, without normalization

[[ 82 27 15]

[ 45 104 15]

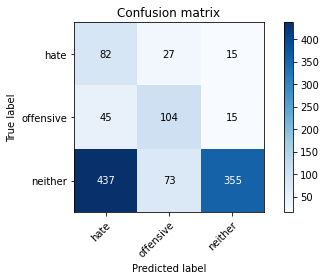
[437 73 355]]

Normalized confusion matrix

[[0.66129032 0.21774194 0.12096774]

[0.27439024 0.63414634 0.09146341]

[0.50520231 0.08439306 0.41040462]]





precision recall f1-score support

hate 0.15 0.66 0.24 124

offensive 0.51 0.63 0.57 164

neither 0.92 0.41 0.57 865

accuracy 0.47 1153

macro avg 0.53 0.57 0.46 1153

weighted avg 0.78 0.47 0.53 1153

average eval\_loss: 2.2820473744951446, accuracy: 0.4692107545533391

training 8 epoch...

114.93622756004333 seconds used......

8 training finished! average train loss: 0.04588977624024708

total clf loss: 8.034055622993037 total adv loss: 6.4914638670452405

evaluating...

Confusion matrix, without normalization

[[ 93 16 15]

[ 49 101 14]

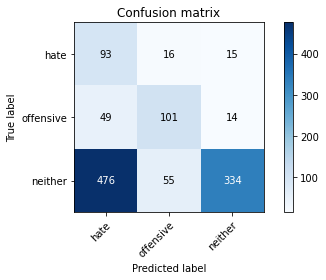
[476 55 334]]

Normalized confusion matrix

[[0.75 0.12903226 0.12096774]

[0.29878049 0.61585366 0.08536585]

[0.55028902 0.06358382 0.38612717]]





precision recall f1-score support

hate 0.15 0.75 0.25 124

offensive 0.59 0.62 0.60 164

neither 0.92 0.39 0.54 865

accuracy 0.46 1153

macro avg 0.55 0.58 0.47 1153

weighted avg 0.79 0.46 0.52 1153

average eval\_loss: 2.4142217410021813, accuracy: 0.45793581960104074

training 9 epoch...

114.97091317176819 seconds used......

9 training finished! average train loss: 0.034139603922431844

total clf loss: 6.008663484361023 total adv loss: 5.464200560352765

evaluating...

Confusion matrix, without normalization

[[ 89 17 18]

[ 46 100 18]

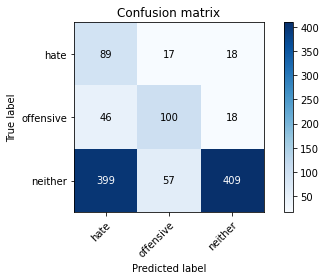
[399 57 409]]

Normalized confusion matrix

[[0.71774194 0.13709677 0.14516129]

[0.2804878 0.6097561 0.1097561 ]

[0.46127168 0.06589595 0.47283237]]





precision recall f1-score support

hate 0.17 0.72 0.27 124

offensive 0.57 0.61 0.59 164

neither 0.92 0.47 0.62 865

accuracy 0.52 1153

macro avg 0.55 0.60 0.50 1153

weighted avg 0.79 0.52 0.58 1153

average eval\_loss: 2.4218823601459634, accuracy: 0.5186470078057241

training 10 epoch...

114.9393219947815 seconds used......

10 training finished! average train loss: 0.033302880447687734

total clf loss: 5.829179689404555 total adv loss: 4.685915825073607

evaluating...

Confusion matrix, without normalization

[[ 82 25 17]

[ 42 107 15]

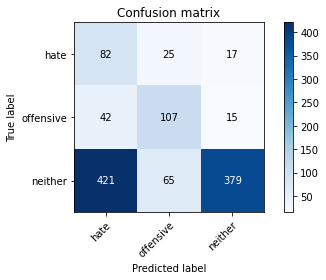
[421 65 379]]

Normalized confusion matrix

[[0.66129032 0.2016129 0.13709677]

[0.25609756 0.65243902 0.09146341]

[0.4867052 0.07514451 0.43815029]]





precision recall f1-score support

hate 0.15 0.66 0.25 124

offensive 0.54 0.65 0.59 164

neither 0.92 0.44 0.59 865

accuracy 0.49 1153

macro avg 0.54 0.58 0.48 1153

weighted avg 0.79 0.49 0.56 1153

average eval\_loss: 2.661443291039302, accuracy: 0.49262792714657416

3 run......

training 1 epoch...

114.986248254776 seconds used......

1 training finished! average train loss: 0.9627276453233901

total clf loss: 166.8495767712593 total adv loss: 102.22666290402412

evaluating...

Confusion matrix, without normalization

[[ 92 29 3]

[ 40 119 5]

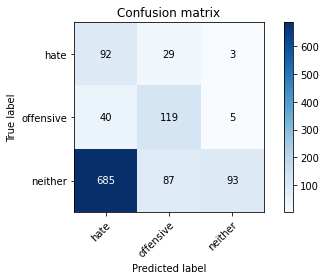
[685 87 93]]

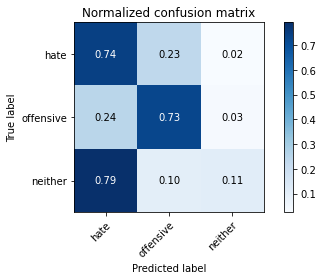
Normalized confusion matrix

[[0.74193548 0.23387097 0.02419355]

[0.24390244 0.72560976 0.0304878 ]

[0.79190751 0.10057803 0.10751445]]





precision recall f1-score support

hate 0.11 0.74 0.20 124

offensive 0.51 0.73 0.60 164

neither 0.92 0.11 0.19 865

accuracy 0.26 1153

macro avg 0.51 0.53 0.33 1153

weighted avg 0.77 0.26 0.25 1153

average eval\_loss: 1.0592372756579826, accuracy: 0.26366001734605377

training 2 epoch...

114.92720746994019 seconds used......

2 training finished! average train loss: 0.8200166633441335

total clf loss: 142.52671766281128 total adv loss: 95.27835914492607

evaluating...

Confusion matrix, without normalization

[[100 14 10]

[ 52 104 8]

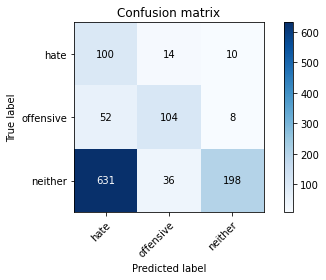
[631 36 198]]

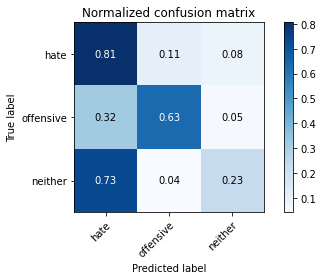
Normalized confusion matrix

[[0.80645161 0.11290323 0.08064516]

[0.31707317 0.63414634 0.04878049]

[0.72947977 0.0416185 0.22890173]]





precision recall f1-score support

hate 0.13 0.81 0.22 124

offensive 0.68 0.63 0.65 164

neither 0.92 0.23 0.37 865

accuracy 0.35 1153

macro avg 0.57 0.56 0.41 1153

weighted avg 0.80 0.35 0.39 1153

average eval\_loss: 0.9674996928922062, accuracy: 0.34865568083261056

training 3 epoch...

115.10630965232849 seconds used......

3 training finished! average train loss: 0.6762336616714796

total clf loss: 117.77866581082344 total adv loss: 83.42822290956974

evaluating...

Confusion matrix, without normalization

[[ 94 17 13]

[ 39 116 9]

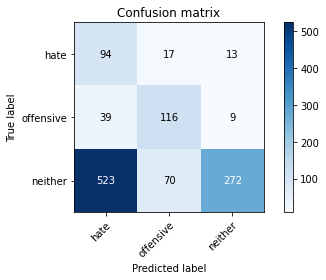
[523 70 272]]

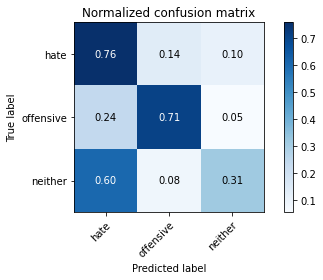
Normalized confusion matrix

[[0.75806452 0.13709677 0.10483871]

[0.23780488 0.70731707 0.05487805]

[0.60462428 0.08092486 0.31445087]]





precision recall f1-score support

hate 0.14 0.76 0.24 124

offensive 0.57 0.71 0.63 164

neither 0.93 0.31 0.47 865

accuracy 0.42 1153

macro avg 0.55 0.59 0.45 1153

weighted avg 0.79 0.42 0.47 1153

average eval\_loss: 1.0652772362889915, accuracy: 0.41803989592367735

training 4 epoch...

114.96850943565369 seconds used......

4 training finished! average train loss: 0.42902005765409695

total clf loss: 74.97882288694382 total adv loss: 58.06905998289585

evaluating...

Confusion matrix, without normalization

[[ 89 21 14]

[ 35 116 13]

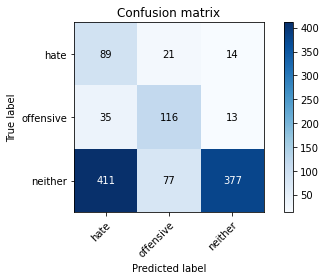
[411 77 377]]

Normalized confusion matrix

[[0.71774194 0.16935484 0.11290323]

[0.21341463 0.70731707 0.07926829]

[0.47514451 0.08901734 0.43583815]]





precision recall f1-score support

hate 0.17 0.72 0.27 124

offensive 0.54 0.71 0.61 164

neither 0.93 0.44 0.59 865

accuracy 0.50 1153

macro avg 0.55 0.62 0.49 1153

weighted avg 0.80 0.50 0.56 1153

average eval\_loss: 1.2400651820774735, accuracy: 0.5047701647875108

training 5 epoch...

114.78778553009033 seconds used......

5 training finished! average train loss: 0.1989989249878341

total clf loss: 34.78381256759167 total adv loss: 27.03986321669072

evaluating...

Confusion matrix, without normalization

[[ 96 15 13]

[ 54 97 13]

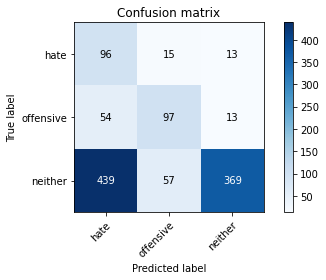
[439 57 369]]

Normalized confusion matrix

[[0.77419355 0.12096774 0.10483871]

[0.32926829 0.59146341 0.07926829]

[0.50751445 0.06589595 0.4265896 ]]





precision recall f1-score support

hate 0.16 0.77 0.27 124

offensive 0.57 0.59 0.58 164

neither 0.93 0.43 0.59 865

accuracy 0.49 1153

macro avg 0.56 0.60 0.48 1153

weighted avg 0.80 0.49 0.55 1153

average eval\_loss: 1.631223756691505, accuracy: 0.48742411101474414

training 6 epoch...

114.87399387359619 seconds used......

6 training finished! average train loss: 0.1063391964361515

total clf loss: 18.614280438050628 total adv loss: 14.98590946663171

evaluating...

Confusion matrix, without normalization

[[ 90 22 12]

[ 47 110 7]

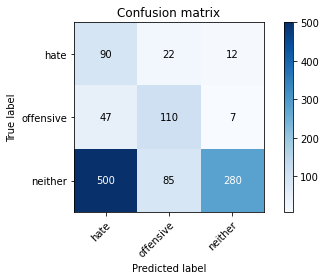
[500 85 280]]

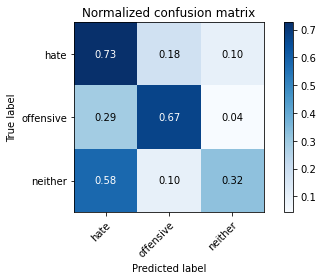
Normalized confusion matrix

[[0.72580645 0.17741935 0.09677419]

[0.28658537 0.67073171 0.04268293]

[0.57803468 0.0982659 0.32369942]]





precision recall f1-score support

hate 0.14 0.73 0.24 124

offensive 0.51 0.67 0.58 164

neither 0.94 0.32 0.48 865

accuracy 0.42 1153

macro avg 0.53 0.57 0.43 1153

weighted avg 0.79 0.42 0.47 1153

average eval\_loss: 2.3953320959518694, accuracy: 0.4163052905464007

training 7 epoch...

114.89125323295593 seconds used......

7 training finished! average train loss: 0.07235879788086527

total clf loss: 12.65514906309545 total adv loss: 9.977419900824316

evaluating...

Confusion matrix, without normalization

[[ 86 21 17]

[ 41 108 15]

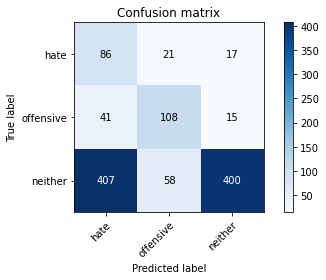
[407 58 400]]

Normalized confusion matrix

[[0.69354839 0.16935484 0.13709677]

[0.25 0.65853659 0.09146341]

[0.47052023 0.06705202 0.46242775]]





precision recall f1-score support

hate 0.16 0.69 0.26 124

offensive 0.58 0.66 0.62 164

neither 0.93 0.46 0.62 865

accuracy 0.52 1153

macro avg 0.55 0.60 0.50 1153

weighted avg 0.79 0.52 0.58 1153

average eval\_loss: 1.8814305128722355, accuracy: 0.5151777970511708

training 8 epoch...

114.94829297065735 seconds used......

8 training finished! average train loss: 0.04297730372248528

total clf loss: 7.505190182244405 total adv loss: 5.700063944503199

evaluating...

Confusion matrix, without normalization

[[ 92 23 9]

[ 47 107 10]

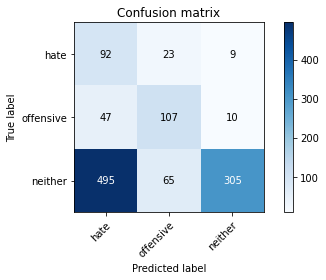
[495 65 305]]

Normalized confusion matrix

[[0.74193548 0.18548387 0.07258065]

[0.28658537 0.65243902 0.06097561]

[0.57225434 0.07514451 0.35260116]]





precision recall f1-score support

hate 0.15 0.74 0.24 124

offensive 0.55 0.65 0.60 164

neither 0.94 0.35 0.51 865

accuracy 0.44 1153

macro avg 0.55 0.58 0.45 1153

weighted avg 0.80 0.44 0.50 1153

average eval\_loss: 2.757419559462317, accuracy: 0.4371205550737207

training 9 epoch...

114.95241451263428 seconds used......

9 training finished! average train loss: 0.039809216274685845

total clf loss: 7.010317545733415 total adv loss: 6.447384498940664

evaluating...

Confusion matrix, without normalization

[[ 83 23 18]

[ 40 109 15]

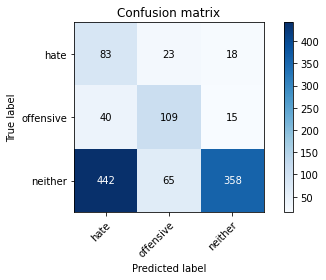
[442 65 358]]

Normalized confusion matrix

[[0.66935484 0.18548387 0.14516129]

[0.24390244 0.66463415 0.09146341]

[0.51098266 0.07514451 0.41387283]]





precision recall f1-score support

hate 0.15 0.67 0.24 124

offensive 0.55 0.66 0.60 164

neither 0.92 0.41 0.57 865

accuracy 0.48 1153

macro avg 0.54 0.58 0.47 1153

weighted avg 0.78 0.48 0.54 1153

average eval\_loss: 2.3165568462733566, accuracy: 0.4770164787510841

training 10 epoch...

114.99628114700317 seconds used......

10 training finished! average train loss: 0.031637832955138515

total clf loss: 5.518107191892341 total adv loss: 4.059024578586104

evaluating...

Confusion matrix, without normalization

[[ 93 15 16]

[ 45 107 12]

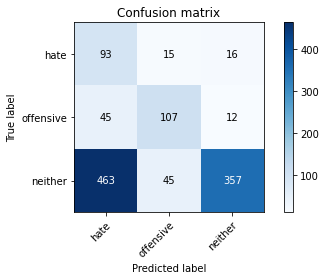
[463 45 357]]

Normalized confusion matrix

[[0.75 0.12096774 0.12903226]

[0.27439024 0.65243902 0.07317073]

[0.53526012 0.05202312 0.41271676]]





precision recall f1-score support

hate 0.15 0.75 0.26 124

offensive 0.64 0.65 0.65 164

neither 0.93 0.41 0.57 865

accuracy 0.48 1153

macro avg 0.57 0.61 0.49 1153

weighted avg 0.80 0.48 0.55 1153

average eval\_loss: 2.5829316603726356, accuracy: 0.48308759757155245

4 run......

training 1 epoch...

114.9965832233429 seconds used......

1 training finished! average train loss: 0.968968180673463

total clf loss: 167.96245050430298 total adv loss: 103.51591739058495

evaluating...

Confusion matrix, without normalization

[[ 97 20 7]

[ 43 109 12]

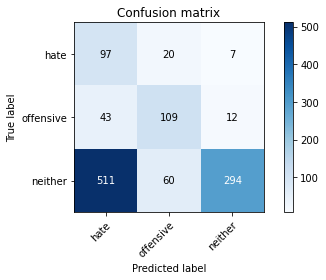
[511 60 294]]

Normalized confusion matrix

[[0.78225806 0.16129032 0.05645161]

[0.26219512 0.66463415 0.07317073]

[0.59075145 0.06936416 0.33988439]]





precision recall f1-score support

hate 0.15 0.78 0.25 124

offensive 0.58 0.66 0.62 164

neither 0.94 0.34 0.50 865

accuracy 0.43 1153

macro avg 0.56 0.60 0.46 1153

weighted avg 0.80 0.43 0.49 1153

average eval\_loss: 0.921824706011805, accuracy: 0.4336513443191674

training 2 epoch...

114.98986482620239 seconds used......

2 training finished! average train loss: 0.8323759383388928

total clf loss: 144.61921948194504 total adv loss: 95.60122632980347

evaluating...

Confusion matrix, without normalization

[[ 82 24 18]

[ 34 118 12]

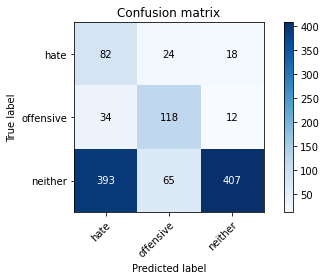
[393 65 407]]

Normalized confusion matrix

[[0.66129032 0.19354839 0.14516129]

[0.20731707 0.7195122 0.07317073]

[0.45433526 0.07514451 0.47052023]]





precision recall f1-score support

hate 0.16 0.66 0.26 124

offensive 0.57 0.72 0.64 164

neither 0.93 0.47 0.63 865

accuracy 0.53 1153

macro avg 0.55 0.62 0.51 1153

weighted avg 0.80 0.53 0.59 1153

average eval\_loss: 0.8729222303834455, accuracy: 0.5264527320034692

training 3 epoch...

114.94324493408203 seconds used......

3 training finished! average train loss: 0.676994030319509

total clf loss: 117.90395948290825 total adv loss: 83.37924663722515

evaluating...

Confusion matrix, without normalization

[[ 67 47 10]

[ 25 129 10]

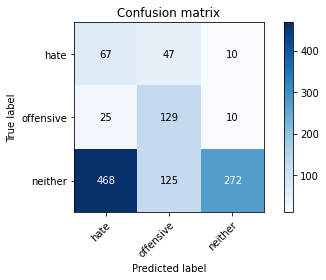
[468 125 272]]

Normalized confusion matrix

[[0.54032258 0.37903226 0.08064516]

[0.15243902 0.78658537 0.06097561]

[0.54104046 0.14450867 0.31445087]]





precision recall f1-score support

hate 0.12 0.54 0.20 124

offensive 0.43 0.79 0.55 164

neither 0.93 0.31 0.47 865

accuracy 0.41 1153

macro avg 0.49 0.55 0.41 1153

weighted avg 0.77 0.41 0.45 1153

average eval\_loss: 1.1292586131342526, accuracy: 0.40589765828274066

training 4 epoch...

115.00396585464478 seconds used......

4 training finished! average train loss: 0.41970874626366866

total clf loss: 73.34215313196182 total adv loss: 56.621676839888096

evaluating...

Confusion matrix, without normalization

[[ 87 14 23]

[ 46 99 19]

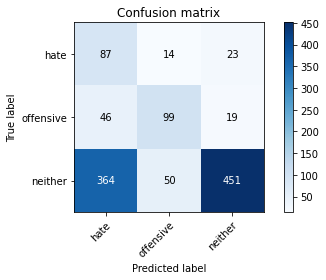
[364 50 451]]

Normalized confusion matrix

[[0.7016129 0.11290323 0.18548387]

[0.2804878 0.60365854 0.11585366]

[0.42080925 0.05780347 0.52138728]]





precision recall f1-score support

hate 0.18 0.70 0.28 124

offensive 0.61 0.60 0.61 164

neither 0.91 0.52 0.66 865

accuracy 0.55 1153

macro avg 0.57 0.61 0.52 1153

weighted avg 0.79 0.55 0.61 1153

average eval\_loss: 1.1355488012576926, accuracy: 0.5524718126626192

training 5 epoch...

114.94903111457825 seconds used......

5 training finished! average train loss: 0.20313391113831175

total clf loss: 35.552057802677155 total adv loss: 28.51121521834284

evaluating...

Confusion matrix, without normalization

[[ 69 35 20]

[ 28 123 13]

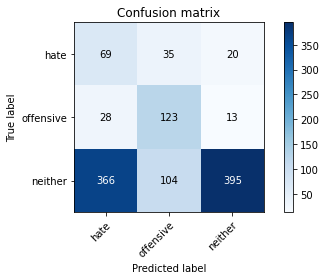
[366 104 395]]

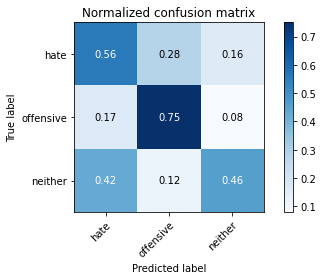
Normalized confusion matrix

[[0.55645161 0.28225806 0.16129032]

[0.17073171 0.75 0.07926829]

[0.42312139 0.12023121 0.4566474 ]]





precision recall f1-score support

hate 0.15 0.56 0.24 124

offensive 0.47 0.75 0.58 164

neither 0.92 0.46 0.61 865

accuracy 0.51 1153

macro avg 0.51 0.59 0.47 1153

weighted avg 0.78 0.51 0.57 1153

average eval\_loss: 1.5390047533758755, accuracy: 0.5091066782307025

training 6 epoch...

114.87534523010254 seconds used......

6 training finished! average train loss: 0.11855897343429249

total clf loss: 20.777509736828506 total adv loss: 17.192045137984678

evaluating...

Confusion matrix, without normalization

[[ 79 24 21]

[ 38 109 17]

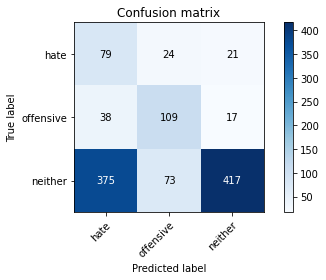
[375 73 417]]

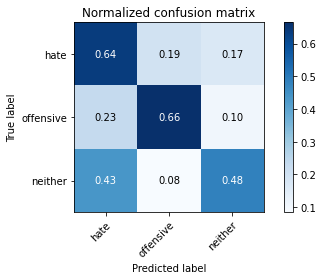
Normalized confusion matrix

[[0.63709677 0.19354839 0.16935484]

[0.23170732 0.66463415 0.10365854]

[0.43352601 0.08439306 0.48208092]]





precision recall f1-score support

hate 0.16 0.64 0.26 124

offensive 0.53 0.66 0.59 164

neither 0.92 0.48 0.63 865

accuracy 0.52 1153

macro avg 0.54 0.59 0.49 1153

weighted avg 0.78 0.52 0.59 1153

average eval\_loss: 1.8460986891697193, accuracy: 0.5247181266261925

training 7 epoch...

114.91853880882263 seconds used......

7 training finished! average train loss: 0.058902437964986475

total clf loss: 10.314213850535452 total adv loss: 8.372084604634438

evaluating...

Confusion matrix, without normalization

[[ 80 26 18]

[ 33 112 19]

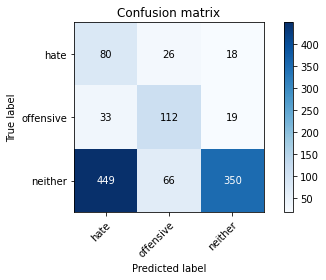
[449 66 350]]

Normalized confusion matrix

[[0.64516129 0.20967742 0.14516129]

[0.20121951 0.68292683 0.11585366]

[0.51907514 0.07630058 0.40462428]]





precision recall f1-score support

hate 0.14 0.65 0.23 124

offensive 0.55 0.68 0.61 164

neither 0.90 0.40 0.56 865

accuracy 0.47 1153

macro avg 0.53 0.58 0.47 1153

weighted avg 0.77 0.47 0.53 1153

average eval\_loss: 2.2350300529907488, accuracy: 0.47007805724197743

training 8 epoch...

114.85334277153015 seconds used......

8 training finished! average train loss: 0.04857519064708391

total clf loss: 8.50561073049903 total adv loss: 6.899574347859016

evaluating...

Confusion matrix, without normalization

[[ 76 17 31]

[ 34 108 22]

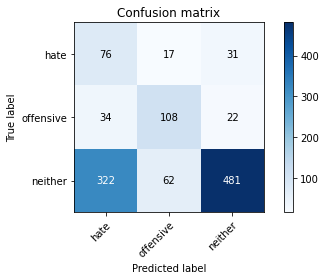
[322 62 481]]

Normalized confusion matrix

[[0.61290323 0.13709677 0.25 ]

[0.20731707 0.65853659 0.13414634]

[0.37225434 0.0716763 0.55606936]]





precision recall f1-score support

hate 0.18 0.61 0.27 124

offensive 0.58 0.66 0.62 164

neither 0.90 0.56 0.69 865

accuracy 0.58 1153

macro avg 0.55 0.61 0.53 1153

weighted avg 0.78 0.58 0.63 1153

average eval\_loss: 1.7823218086670185, accuracy: 0.5767562879444926

training 9 epoch...

114.91850304603577 seconds used......

9 training finished! average train loss: 0.034106857514762806

total clf loss: 5.961234241491184 total adv loss: 4.6256436568219215

evaluating...

Confusion matrix, without normalization

[[ 78 21 25]

[ 38 105 21]

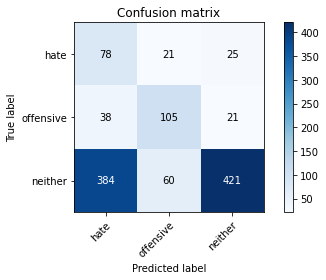
[384 60 421]]

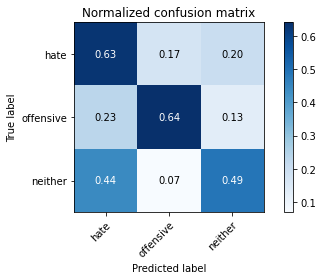
Normalized confusion matrix

[[0.62903226 0.16935484 0.2016129 ]

[0.23170732 0.6402439 0.12804878]

[0.44393064 0.06936416 0.4867052 ]]





precision recall f1-score support

hate 0.16 0.63 0.25 124

offensive 0.56 0.64 0.60 164

neither 0.90 0.49 0.63 865

accuracy 0.52 1153

macro avg 0.54 0.59 0.49 1153

weighted avg 0.77 0.52 0.59 1153

average eval\_loss: 2.2126803613942245, accuracy: 0.5238508239375542

training 10 epoch...

114.95726871490479 seconds used......

10 training finished! average train loss: 0.028960885078482705

total clf loss: 5.088130006042775 total adv loss: 4.4540260296926135

evaluating...

Confusion matrix, without normalization

[[ 69 39 16]

[ 32 121 11]

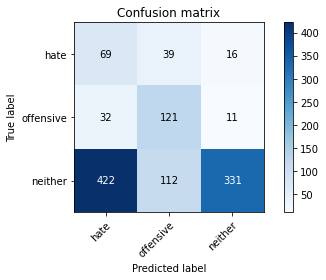
[422 112 331]]

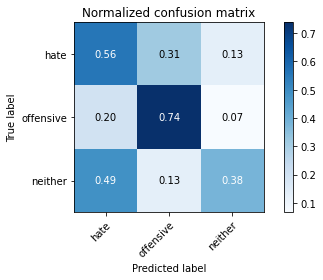
Normalized confusion matrix

[[0.55645161 0.31451613 0.12903226]

[0.19512195 0.73780488 0.06707317]

[0.48786127 0.12947977 0.38265896]]





precision recall f1-score support

hate 0.13 0.56 0.21 124

offensive 0.44 0.74 0.56 164

neither 0.92 0.38 0.54 865

accuracy 0.45 1153

macro avg 0.50 0.56 0.44 1153

weighted avg 0.77 0.45 0.51 1153

average eval\_loss: 2.864778146661561, accuracy: 0.4518647007805724

5 run......

training 1 epoch...

115.02659153938293 seconds used......

1 training finished! average train loss: 0.9800354171366918

total clf loss: 169.8233104944229 total adv loss: 103.54721775650978

evaluating...

Confusion matrix, without normalization

[[107 16 1]

[ 65 98 1]

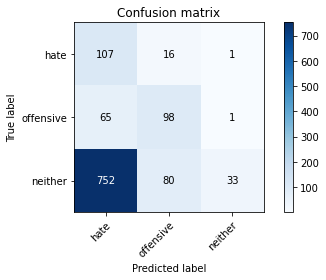
[752 80 33]]

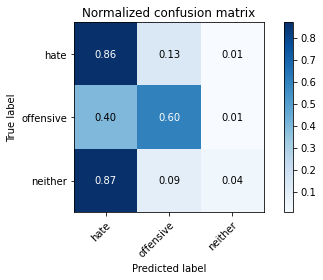
Normalized confusion matrix

[[0.86290323 0.12903226 0.00806452]

[0.39634146 0.59756098 0.00609756]

[0.86936416 0.09248555 0.03815029]]





precision recall f1-score support

hate 0.12 0.86 0.20 124

offensive 0.51 0.60 0.55 164

neither 0.94 0.04 0.07 865

accuracy 0.21 1153

macro avg 0.52 0.50 0.28 1153

weighted avg 0.79 0.21 0.15 1153

average eval\_loss: 1.110636191121463, accuracy: 0.20641803989592367

training 2 epoch...

114.94830012321472 seconds used......

2 training finished! average train loss: 0.8499583399721554

total clf loss: 147.61757135391235 total adv loss: 96.4913980960846

evaluating...

Confusion matrix, without normalization

[[ 93 21 10]

[ 30 124 10]

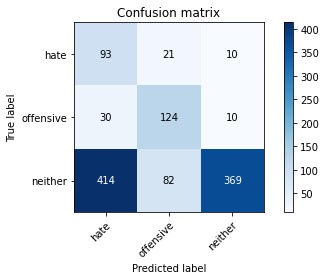
[414 82 369]]

Normalized confusion matrix

[[0.75 0.16935484 0.08064516]

[0.18292683 0.75609756 0.06097561]

[0.47861272 0.09479769 0.4265896 ]]





precision recall f1-score support

hate 0.17 0.75 0.28 124

offensive 0.55 0.76 0.63 164

neither 0.95 0.43 0.59 865

accuracy 0.51 1153

macro avg 0.56 0.64 0.50 1153

weighted avg 0.81 0.51 0.56 1153

average eval\_loss: 0.8889471559688963, accuracy: 0.5082393755420642

training 3 epoch...

114.95724487304688 seconds used......

3 training finished! average train loss: 0.7246699033393746

total clf loss: 126.08511817455292 total adv loss: 86.81148555874825

evaluating...

Confusion matrix, without normalization

[[ 94 16 14]

[ 41 108 15]

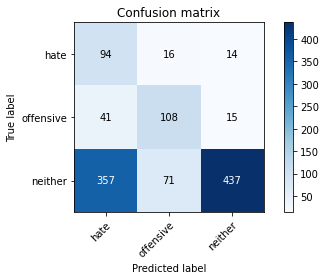
[357 71 437]]

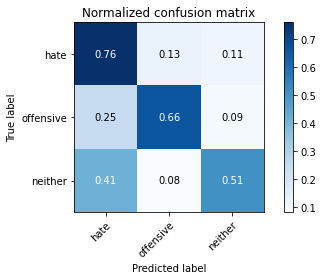
Normalized confusion matrix

[[0.75806452 0.12903226 0.11290323]

[0.25 0.65853659 0.09146341]

[0.41271676 0.08208092 0.50520231]]





precision recall f1-score support

hate 0.19 0.76 0.31 124

offensive 0.55 0.66 0.60 164

neither 0.94 0.51 0.66 865

accuracy 0.55 1153

macro avg 0.56 0.64 0.52 1153

weighted avg 0.80 0.55 0.61 1153

average eval\_loss: 0.8854703276321806, accuracy: 0.5542064180398959

training 4 epoch...

114.97390866279602 seconds used......

4 training finished! average train loss: 0.49827328971808865

total clf loss: 86.9215023368597 total adv loss: 64.2317945510149

evaluating...

Confusion matrix, without normalization

[[ 62 30 32]

[ 31 111 22]

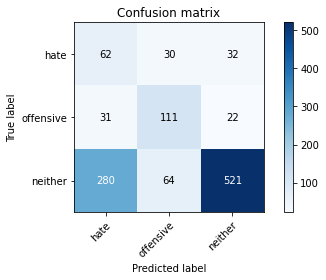
[280 64 521]]

Normalized confusion matrix

[[0.5 0.24193548 0.25806452]

[0.18902439 0.67682927 0.13414634]

[0.32369942 0.07398844 0.60231214]]





precision recall f1-score support

hate 0.17 0.50 0.25 124

offensive 0.54 0.68 0.60 164

neither 0.91 0.60 0.72 865

accuracy 0.60 1153

macro avg 0.54 0.59 0.52 1153

weighted avg 0.77 0.60 0.66 1153

average eval\_loss: 0.9119806793229334, accuracy: 0.6019080659150043

training 5 epoch...

114.96411323547363 seconds used......

5 training finished! average train loss: 0.26248946510964916

total clf loss: 45.795103304088116 total adv loss: 33.93746338225901

evaluating...

Confusion matrix, without normalization

[[ 93 17 14]

[ 45 105 14]

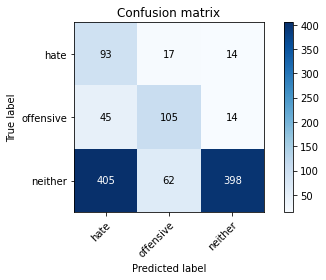
[405 62 398]]

Normalized confusion matrix

[[0.75 0.13709677 0.11290323]

[0.27439024 0.6402439 0.08536585]

[0.46820809 0.0716763 0.46011561]]





precision recall f1-score support

hate 0.17 0.75 0.28 124

offensive 0.57 0.64 0.60 164

neither 0.93 0.46 0.62 865

accuracy 0.52 1153

macro avg 0.56 0.62 0.50 1153

weighted avg 0.80 0.52 0.58 1153

average eval\_loss: 1.514472494865286, accuracy: 0.5169124024284475

training 6 epoch...

115.02051305770874 seconds used......

6 training finished! average train loss: 0.13069979955131808

total clf loss: 22.829557033255696 total adv loss: 17.439814590848982

evaluating...

Confusion matrix, without normalization

[[ 46 27 51]

[ 21 108 35]

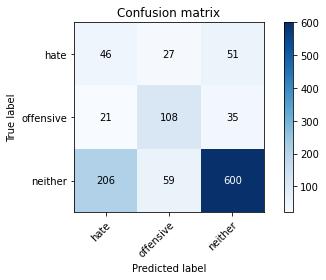
[206 59 600]]

Normalized confusion matrix

[[0.37096774 0.21774194 0.41129032]

[0.12804878 0.65853659 0.21341463]

[0.23815029 0.06820809 0.69364162]]





precision recall f1-score support

hate 0.17 0.37 0.23 124

offensive 0.56 0.66 0.60 164

neither 0.87 0.69 0.77 865

accuracy 0.65 1153

macro avg 0.53 0.57 0.54 1153

weighted avg 0.75 0.65 0.69 1153

average eval\_loss: 1.1097293732495144, accuracy: 0.6539462272333044

training 7 epoch...

115.09779477119446 seconds used......

7 training finished! average train loss: 0.0841950534820734

total clf loss: 14.706847682595253 total adv loss: 11.241574265761301

evaluating...

Confusion matrix, without normalization

[[ 94 18 12]

[ 50 98 16]

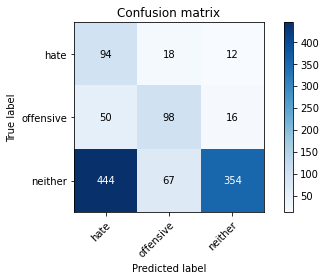
[444 67 354]]

Normalized confusion matrix

[[0.75806452 0.14516129 0.09677419]

[0.30487805 0.59756098 0.09756098]

[0.5132948 0.07745665 0.40924855]]





precision recall f1-score support

hate 0.16 0.76 0.26 124

offensive 0.54 0.60 0.56 164

neither 0.93 0.41 0.57 865

accuracy 0.47 1153

macro avg 0.54 0.59 0.47 1153

weighted avg 0.79 0.47 0.53 1153

average eval\_loss: 2.155996824132985, accuracy: 0.4735472679965308

training 8 epoch...

114.97894763946533 seconds used......

8 training finished! average train loss: 0.05526285067233922

total clf loss: 9.638751897029579 total adv loss: 7.091859771491727

evaluating...

Confusion matrix, without normalization

[[ 90 18 16]

[ 50 93 21]

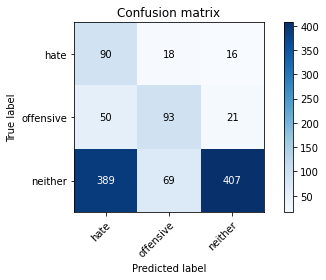
[389 69 407]]

Normalized confusion matrix

[[0.72580645 0.14516129 0.12903226]

[0.30487805 0.56707317 0.12804878]

[0.44971098 0.07976879 0.47052023]]





precision recall f1-score support

hate 0.17 0.73 0.28 124

offensive 0.52 0.57 0.54 164

neither 0.92 0.47 0.62 865

accuracy 0.51 1153

macro avg 0.53 0.59 0.48 1153

weighted avg 0.78 0.51 0.57 1153

average eval\_loss: 2.156833021805204, accuracy: 0.5117085862966175

training 9 epoch...

115.00708270072937 seconds used......

9 training finished! average train loss: 0.04517357824425701

total clf loss: 7.928498484892771 total adv loss: 6.786746718047652

evaluating...

Confusion matrix, without normalization

[[ 88 17 19]

[ 47 97 20]

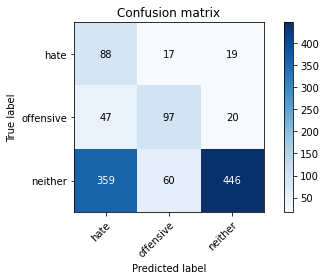
[359 60 446]]

Normalized confusion matrix

[[0.70967742 0.13709677 0.15322581]

[0.28658537 0.59146341 0.12195122]

[0.4150289 0.06936416 0.51560694]]





precision recall f1-score support

hate 0.18 0.71 0.28 124

offensive 0.56 0.59 0.57 164

neither 0.92 0.52 0.66 865

accuracy 0.55 1153

macro avg 0.55 0.61 0.51 1153

weighted avg 0.79 0.55 0.61 1153

average eval\_loss: 2.0001393679914803, accuracy: 0.5472679965307893

training 10 epoch...

114.97009754180908 seconds used......

10 training finished! average train loss: 0.03472020300839739

total clf loss: 6.084659698186442 total adv loss: 5.033311847422738

evaluating...

Confusion matrix, without normalization

[[106 9 9]

[ 71 81 12]

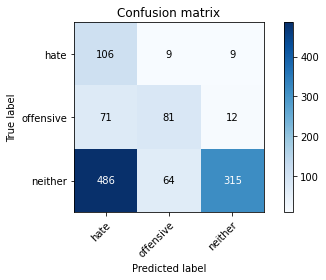
[486 64 315]]

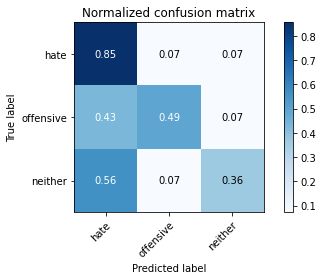
Normalized confusion matrix

[[0.85483871 0.07258065 0.07258065]

[0.43292683 0.49390244 0.07317073]

[0.56184971 0.07398844 0.36416185]]





precision recall f1-score support

hate 0.16 0.85 0.27 124

offensive 0.53 0.49 0.51 164

neither 0.94 0.36 0.52 865

accuracy 0.44 1153

macro avg 0.54 0.57 0.43 1153

weighted avg 0.80 0.44 0.49 1153

average eval\_loss: 3.0260316133499146, accuracy: 0.43538594969644406