alpha =0.05

1 run......

training 1 epoch...

69.552011013031 seconds used......

1 training finished! average train loss: 0.5719388969178892

total clf loss: 138.84337720274925 total adv loss: 100.19350834935904

evaluating...

Confusion matrix, without normalization

[[ 16 63 17]

[ 4 291 60]

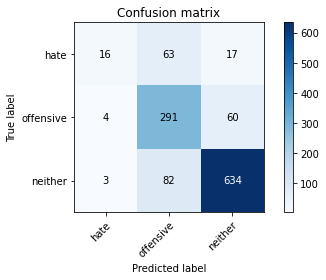
[ 3 82 634]]

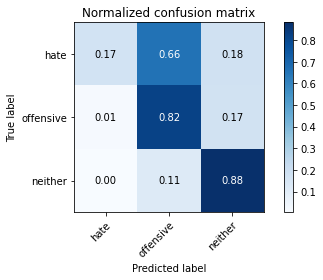
Normalized confusion matrix

[[0.16666667 0.65625 0.17708333]

[0.01126761 0.81971831 0.16901408]

[0.00417246 0.11404729 0.88178025]]





precision recall f1-score support

hate 0.70 0.17 0.27 96

offensive 0.67 0.82 0.74 355

neither 0.89 0.88 0.89 719

accuracy 0.80 1170

macro avg 0.75 0.62 0.63 1170

weighted avg 0.81 0.80 0.79 1170

average eval\_loss: 0.4446942976723283, accuracy: 0.8042735042735043

training 2 epoch...

69.49307084083557 seconds used......

2 training finished! average train loss: 0.3027657947988592

total clf loss: 73.39230266958475 total adv loss: 50.90213775075972

evaluating...

Confusion matrix, without normalization

[[ 61 19 16]

[ 38 255 62]

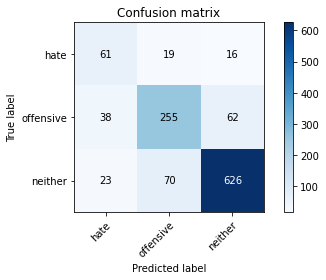
[ 23 70 626]]

Normalized confusion matrix

[[0.63541667 0.19791667 0.16666667]

[0.10704225 0.71830986 0.17464789]

[0.03198887 0.09735744 0.87065369]]





precision recall f1-score support

hate 0.50 0.64 0.56 96

offensive 0.74 0.72 0.73 355

neither 0.89 0.87 0.88 719

accuracy 0.81 1170

macro avg 0.71 0.74 0.72 1170

weighted avg 0.81 0.81 0.81 1170

average eval\_loss: 0.49478128110453234, accuracy: 0.8051282051282052

training 3 epoch...

69.50350832939148 seconds used......

3 training finished! average train loss: 0.1282540150285245

total clf loss: 31.042107515968382 total adv loss: 20.613359054666944

evaluating...

Confusion matrix, without normalization

[[ 52 29 15]

[ 20 260 75]

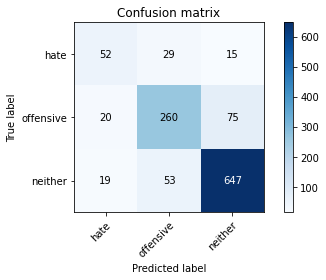
[ 19 53 647]]

Normalized confusion matrix

[[0.54166667 0.30208333 0.15625 ]

[0.05633803 0.73239437 0.21126761]

[0.02642559 0.07371349 0.89986092]]





precision recall f1-score support

hate 0.57 0.54 0.56 96

offensive 0.76 0.73 0.75 355

neither 0.88 0.90 0.89 719

accuracy 0.82 1170

macro avg 0.74 0.72 0.73 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.5996698984536895, accuracy: 0.8196581196581196

training 4 epoch...

69.6311182975769 seconds used......

4 training finished! average train loss: 0.0683792880955192

total clf loss: 16.542306101182476 total adv loss: 10.83105386243551

evaluating...

Confusion matrix, without normalization

[[ 47 35 14]

[ 17 281 57]

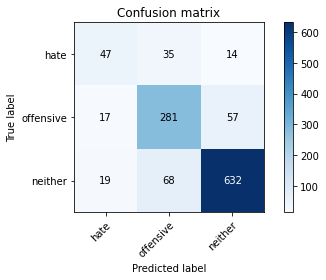
[ 19 68 632]]

Normalized confusion matrix

[[0.48958333 0.36458333 0.14583333]

[0.04788732 0.7915493 0.16056338]

[0.02642559 0.0945758 0.87899861]]





precision recall f1-score support

hate 0.57 0.49 0.53 96

offensive 0.73 0.79 0.76 355

neither 0.90 0.88 0.89 719

accuracy 0.82 1170

macro avg 0.73 0.72 0.72 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.6367450505968625, accuracy: 0.8205128205128205

training 5 epoch...

69.60935521125793 seconds used......

5 training finished! average train loss: 0.037438959376060635

total clf loss: 9.144024677574635 total adv loss: 7.666164068476064

evaluating...

Confusion matrix, without normalization

[[ 44 35 17]

[ 16 280 59]

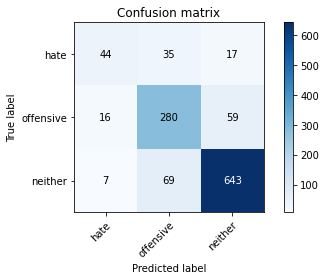
[ 7 69 643]]

Normalized confusion matrix

[[0.45833333 0.36458333 0.17708333]

[0.04507042 0.78873239 0.16619718]

[0.00973574 0.09596662 0.89429764]]





precision recall f1-score support

hate 0.66 0.46 0.54 96

offensive 0.73 0.79 0.76 355

neither 0.89 0.89 0.89 719

accuracy 0.83 1170

macro avg 0.76 0.71 0.73 1170

weighted avg 0.82 0.83 0.82 1170

average eval\_loss: 0.7364361675237148, accuracy: 0.8264957264957264

training 6 epoch...

69.53117823600769 seconds used......

6 training finished! average train loss: 0.02619151040413178

total clf loss: 6.442998212733073 total adv loss: 6.283695590565912

evaluating...

Confusion matrix, without normalization

[[ 67 11 18]

[ 55 221 79]

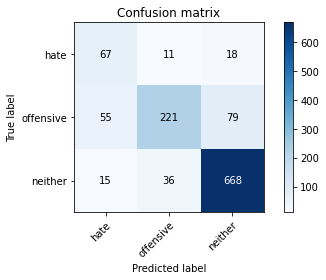
[ 15 36 668]]

Normalized confusion matrix

[[0.69791667 0.11458333 0.1875 ]

[0.15492958 0.62253521 0.22253521]

[0.02086231 0.05006954 0.92906815]]





precision recall f1-score support

hate 0.49 0.70 0.58 96

offensive 0.82 0.62 0.71 355

neither 0.87 0.93 0.90 719

accuracy 0.82 1170

macro avg 0.73 0.75 0.73 1170

weighted avg 0.83 0.82 0.82 1170

average eval\_loss: 0.8165815729032255, accuracy: 0.8170940170940171

training 7 epoch...

69.58477926254272 seconds used......

7 training finished! average train loss: 0.025565876535803646

total clf loss: 6.172426556513528 total adv loss: 3.8002294679499755

evaluating...

Confusion matrix, without normalization

[[ 28 47 21]

[ 4 269 82]

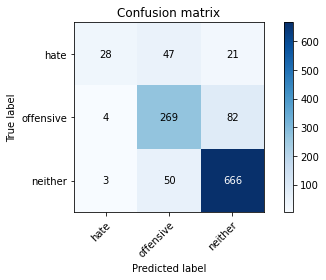
[ 3 50 666]]

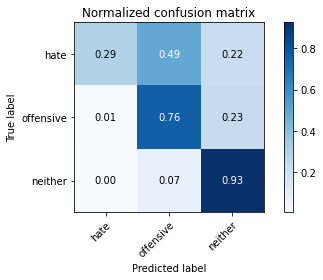
Normalized confusion matrix

[[0.29166667 0.48958333 0.21875 ]

[0.01126761 0.75774648 0.23098592]

[0.00417246 0.06954103 0.92628651]]





precision recall f1-score support

hate 0.80 0.29 0.43 96

offensive 0.73 0.76 0.75 355

neither 0.87 0.93 0.90 719

accuracy 0.82 1170

macro avg 0.80 0.66 0.69 1170

weighted avg 0.82 0.82 0.81 1170

average eval\_loss: 0.9489218945737329, accuracy: 0.823076923076923

training 8 epoch...

69.53666067123413 seconds used......

8 training finished! average train loss: 0.034486421288148314

total clf loss: 8.29962750300183 total adv loss: 4.596098337315198

evaluating...

Confusion matrix, without normalization

[[ 60 19 17]

[ 32 242 81]

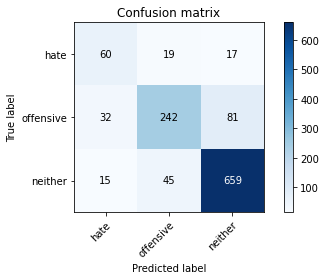
[ 15 45 659]]

Normalized confusion matrix

[[0.625 0.19791667 0.17708333]

[0.09014085 0.68169014 0.22816901]

[0.02086231 0.06258693 0.91655076]]





precision recall f1-score support

hate 0.56 0.62 0.59 96

offensive 0.79 0.68 0.73 355

neither 0.87 0.92 0.89 719

accuracy 0.82 1170

macro avg 0.74 0.74 0.74 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.8186660284764465, accuracy: 0.8213675213675213

training 9 epoch...

69.64234232902527 seconds used......

9 training finished! average train loss: 0.01667152396964186

total clf loss: 4.091282634821255 total adv loss: 3.802920234964404

evaluating...

Confusion matrix, without normalization

[[ 51 30 15]

[ 21 275 59]

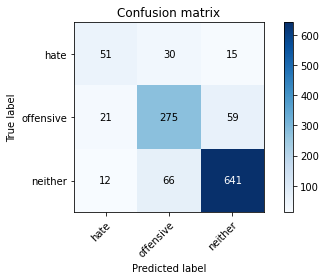
[ 12 66 641]]

Normalized confusion matrix

[[0.53125 0.3125 0.15625 ]

[0.05915493 0.77464789 0.16619718]

[0.01668985 0.09179416 0.89151599]]





precision recall f1-score support

hate 0.61 0.53 0.57 96

offensive 0.74 0.77 0.76 355

neither 0.90 0.89 0.89 719

accuracy 0.83 1170

macro avg 0.75 0.73 0.74 1170

weighted avg 0.83 0.83 0.83 1170

average eval\_loss: 0.8634809823746209, accuracy: 0.8264957264957264

training 10 epoch...

69.38720679283142 seconds used......

10 training finished! average train loss: 0.020949388145577088

total clf loss: 5.125597318605287 total adv loss: 4.468808952429754

evaluating...

Confusion matrix, without normalization

[[ 62 16 18]

[ 33 255 67]

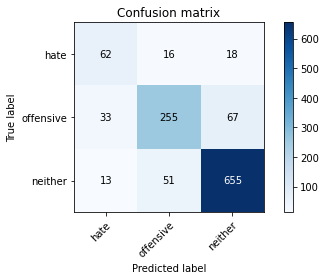
[ 13 51 655]]

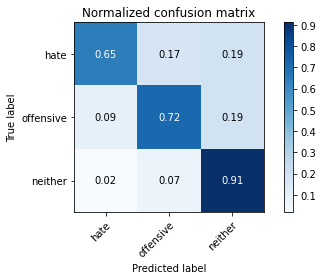
Normalized confusion matrix

[[0.64583333 0.16666667 0.1875 ]

[0.09295775 0.71830986 0.18873239]

[0.01808067 0.07093185 0.91098748]]





precision recall f1-score support

hate 0.57 0.65 0.61 96

offensive 0.79 0.72 0.75 355

neither 0.89 0.91 0.90 719

accuracy 0.83 1170

macro avg 0.75 0.76 0.75 1170

weighted avg 0.83 0.83 0.83 1170

average eval\_loss: 0.7672245595209554, accuracy: 0.8307692307692308

2 run......

training 1 epoch...

69.27910089492798 seconds used......

1 training finished! average train loss: 0.5822967774847634

total clf loss: 141.49632713198662 total adv loss: 104.77762109041214

evaluating...

Confusion matrix, without normalization

[[ 11 67 18]

[ 1 286 68]

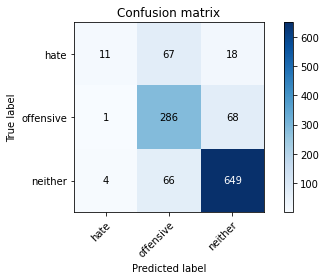
[ 4 66 649]]

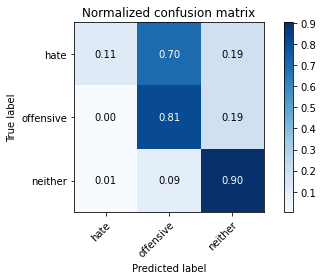
Normalized confusion matrix

[[0.11458333 0.69791667 0.1875 ]

[0.0028169 0.8056338 0.1915493 ]

[0.00556328 0.09179416 0.90264256]]





precision recall f1-score support

hate 0.69 0.11 0.20 96

offensive 0.68 0.81 0.74 355

neither 0.88 0.90 0.89 719

accuracy 0.81 1170

macro avg 0.75 0.61 0.61 1170

weighted avg 0.81 0.81 0.79 1170

average eval\_loss: 0.4750349498906378, accuracy: 0.8085470085470086

training 2 epoch...

69.31092286109924 seconds used......

2 training finished! average train loss: 0.3041772559189644

total clf loss: 73.71338918805122 total adv loss: 50.71822631265968

evaluating...

Confusion matrix, without normalization

[[ 55 28 13]

[ 19 278 58]

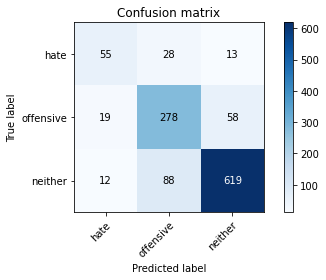
[ 12 88 619]]

Normalized confusion matrix

[[0.57291667 0.29166667 0.13541667]

[0.05352113 0.78309859 0.16338028]

[0.01668985 0.12239221 0.86091794]]





precision recall f1-score support

hate 0.64 0.57 0.60 96

offensive 0.71 0.78 0.74 355

neither 0.90 0.86 0.88 719

accuracy 0.81 1170

macro avg 0.75 0.74 0.74 1170

weighted avg 0.82 0.81 0.81 1170

average eval\_loss: 0.4621093203708277, accuracy: 0.8136752136752137

training 3 epoch...

69.44119596481323 seconds used......

3 training finished! average train loss: 0.1284271133187999

total clf loss: 31.048198928125203 total adv loss: 19.925089426222257

evaluating...

Confusion matrix, without normalization

[[ 38 36 22]

[ 13 243 99]

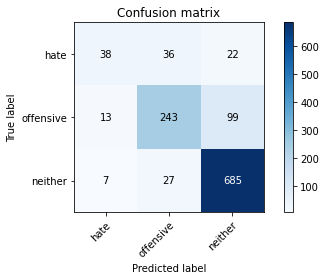
[ 7 27 685]]

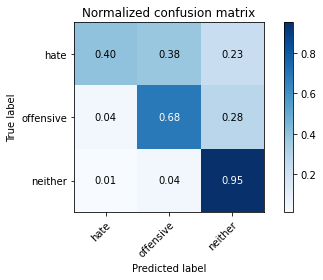
Normalized confusion matrix

[[0.39583333 0.375 0.22916667]

[0.03661972 0.68450704 0.27887324]

[0.00973574 0.03755216 0.9527121 ]]





precision recall f1-score support

hate 0.66 0.40 0.49 96

offensive 0.79 0.68 0.74 355

neither 0.85 0.95 0.90 719

accuracy 0.83 1170

macro avg 0.77 0.68 0.71 1170

weighted avg 0.82 0.83 0.82 1170

average eval\_loss: 0.5939800034955902, accuracy: 0.8256410256410256

training 4 epoch...

69.34638595581055 seconds used......

4 training finished! average train loss: 0.053600105046071746

total clf loss: 12.936191066051833 total adv loss: 7.8753296261202195

evaluating...

Confusion matrix, without normalization

[[ 55 32 9]

[ 31 280 44]

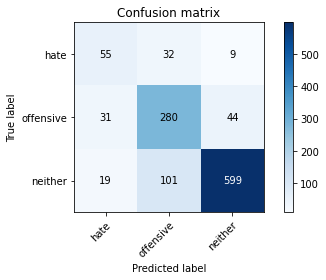
[ 19 101 599]]

Normalized confusion matrix

[[0.57291667 0.33333333 0.09375 ]

[0.08732394 0.78873239 0.12394366]

[0.02642559 0.14047288 0.83310153]]





precision recall f1-score support

hate 0.52 0.57 0.55 96

offensive 0.68 0.79 0.73 355

neither 0.92 0.83 0.87 719

accuracy 0.80 1170

macro avg 0.71 0.73 0.72 1170

weighted avg 0.81 0.80 0.80 1170

average eval\_loss: 0.8084007473820347, accuracy: 0.7982905982905983

training 5 epoch...

69.39657425880432 seconds used......

5 training finished! average train loss: 0.0394934239746549

total clf loss: 9.64807782479329 total adv loss: 8.132332365210459

evaluating...

Confusion matrix, without normalization

[[ 53 26 17]

[ 25 263 67]

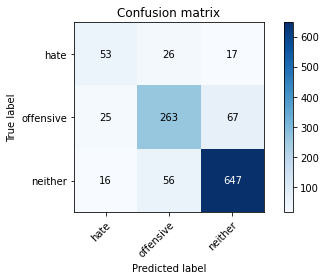
[ 16 56 647]]

Normalized confusion matrix

[[0.55208333 0.27083333 0.17708333]

[0.07042254 0.74084507 0.18873239]

[0.02225313 0.07788595 0.89986092]]





precision recall f1-score support

hate 0.56 0.55 0.56 96

offensive 0.76 0.74 0.75 355

neither 0.89 0.90 0.89 719

accuracy 0.82 1170

macro avg 0.74 0.73 0.73 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.7549779097849535, accuracy: 0.823076923076923

training 6 epoch...

69.48443460464478 seconds used......

6 training finished! average train loss: 0.034286407752083935

total clf loss: 8.271359908569138 total adv loss: 4.966810186564544

evaluating...

Confusion matrix, without normalization

[[ 54 26 16]

[ 26 268 61]

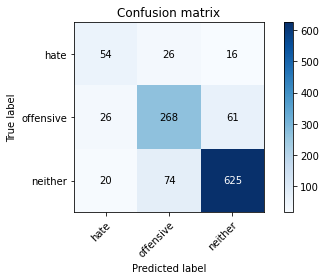
[ 20 74 625]]

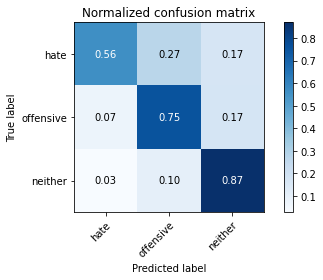
Normalized confusion matrix

[[0.5625 0.27083333 0.16666667]

[0.07323944 0.75492958 0.17183099]

[0.02781641 0.10292072 0.86926287]]





precision recall f1-score support

hate 0.54 0.56 0.55 96

offensive 0.73 0.75 0.74 355

neither 0.89 0.87 0.88 719

accuracy 0.81 1170

macro avg 0.72 0.73 0.72 1170

weighted avg 0.81 0.81 0.81 1170

average eval\_loss: 0.83491876873314, accuracy: 0.8094017094017094

training 7 epoch...

69.49952936172485 seconds used......

7 training finished! average train loss: 0.026472747702878885

total clf loss: 6.440902510861633 total adv loss: 4.92559150096713

evaluating...

Confusion matrix, without normalization

[[ 50 32 14]

[ 13 294 48]

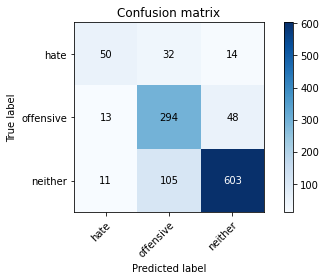
[ 11 105 603]]

Normalized confusion matrix

[[0.52083333 0.33333333 0.14583333]

[0.03661972 0.82816901 0.13521127]

[0.01529903 0.14603616 0.83866481]]





precision recall f1-score support

hate 0.68 0.52 0.59 96

offensive 0.68 0.83 0.75 355

neither 0.91 0.84 0.87 719

accuracy 0.81 1170

macro avg 0.75 0.73 0.74 1170

weighted avg 0.82 0.81 0.81 1170

average eval\_loss: 0.9107847223531896, accuracy: 0.8094017094017094

training 8 epoch...

69.58138966560364 seconds used......

8 training finished! average train loss: 0.02001954840854259

total clf loss: 4.8478574818436755 total adv loss: 3.265663868518459

evaluating...

Confusion matrix, without normalization

[[ 41 38 17]

[ 8 283 64]

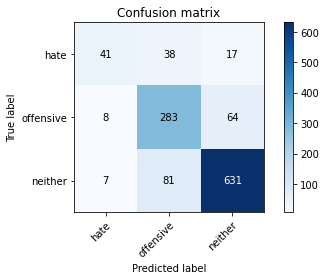
[ 7 81 631]]

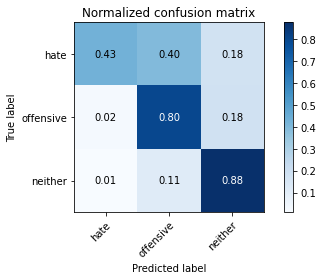
Normalized confusion matrix

[[0.42708333 0.39583333 0.17708333]

[0.02253521 0.7971831 0.18028169]

[0.00973574 0.11265647 0.87760779]]





precision recall f1-score support

hate 0.73 0.43 0.54 96

offensive 0.70 0.80 0.75 355

neither 0.89 0.88 0.88 719

accuracy 0.82 1170

macro avg 0.77 0.70 0.72 1170

weighted avg 0.82 0.82 0.81 1170

average eval\_loss: 0.965181680309381, accuracy: 0.8162393162393162

training 9 epoch...

69.7139003276825 seconds used......

9 training finished! average train loss: 0.013844433823290966

total clf loss: 3.3665607828006614 total adv loss: 2.539265213199542

evaluating...

Confusion matrix, without normalization

[[ 48 32 16]

[ 28 261 66]

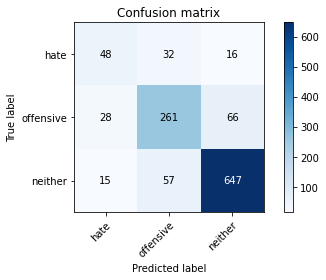
[ 15 57 647]]

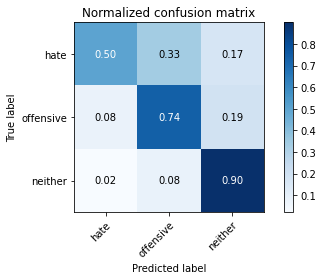
Normalized confusion matrix

[[0.5 0.33333333 0.16666667]

[0.07887324 0.73521127 0.18591549]

[0.02086231 0.07927677 0.89986092]]





precision recall f1-score support

hate 0.53 0.50 0.51 96

offensive 0.75 0.74 0.74 355

neither 0.89 0.90 0.89 719

accuracy 0.82 1170

macro avg 0.72 0.71 0.72 1170

weighted avg 0.81 0.82 0.82 1170

average eval\_loss: 0.8643945363356133, accuracy: 0.8170940170940171

training 10 epoch...

69.62840819358826 seconds used......

10 training finished! average train loss: 0.01474560476083589

total clf loss: 3.640624924970325 total adv loss: 3.803068541419634

evaluating...

Confusion matrix, without normalization

[[ 56 29 11]

[ 30 283 42]

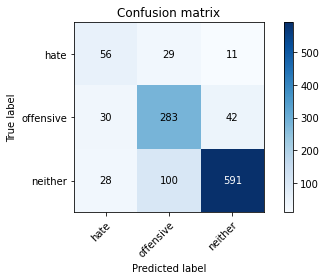
[ 28 100 591]]

Normalized confusion matrix

[[0.58333333 0.30208333 0.11458333]

[0.08450704 0.7971831 0.11830986]

[0.03894298 0.13908206 0.82197497]]





precision recall f1-score support

hate 0.49 0.58 0.53 96

offensive 0.69 0.80 0.74 355

neither 0.92 0.82 0.87 719

accuracy 0.79 1170

macro avg 0.70 0.73 0.71 1170

weighted avg 0.81 0.79 0.80 1170

average eval\_loss: 1.026931500867388, accuracy: 0.7948717948717948

3 run......

training 1 epoch...

69.50370764732361 seconds used......

1 training finished! average train loss: 0.6182239695619314

total clf loss: 150.10058717429638 total adv loss: 108.72356007993221

evaluating...

Confusion matrix, without normalization

[[ 13 54 29]

[ 2 246 107]

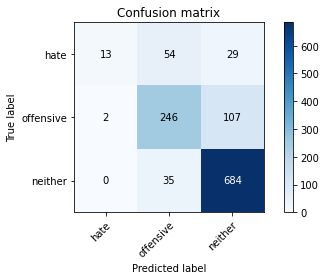
[ 0 35 684]]

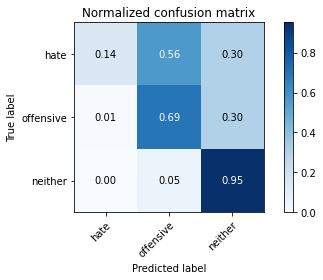
Normalized confusion matrix

[[0.13541667 0.5625 0.30208333]

[0.0056338 0.69295775 0.30140845]

[0. 0.04867872 0.95132128]]





precision recall f1-score support

hate 0.87 0.14 0.23 96

offensive 0.73 0.69 0.71 355

neither 0.83 0.95 0.89 719

accuracy 0.81 1170

macro avg 0.81 0.59 0.61 1170

weighted avg 0.81 0.81 0.78 1170

average eval\_loss: 0.46326032558740193, accuracy: 0.805982905982906

training 2 epoch...

69.56705474853516 seconds used......

2 training finished! average train loss: 0.32413207655215365

total clf loss: 78.54945583641529 total adv loss: 54.051000060513616

evaluating...

Confusion matrix, without normalization

[[ 29 57 10]

[ 3 318 34]

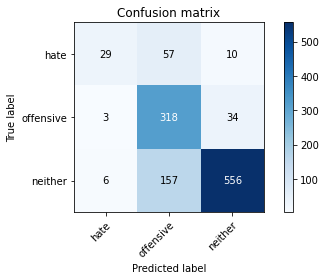
[ 6 157 556]]

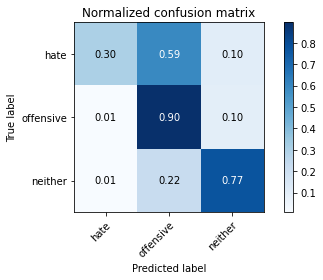
Normalized confusion matrix

[[0.30208333 0.59375 0.10416667]

[0.0084507 0.89577465 0.09577465]

[0.00834492 0.21835883 0.77329624]]





precision recall f1-score support

hate 0.76 0.30 0.43 96

offensive 0.60 0.90 0.72 355

neither 0.93 0.77 0.84 719

accuracy 0.77 1170

macro avg 0.76 0.66 0.66 1170

weighted avg 0.81 0.77 0.77 1170

average eval\_loss: 0.5340472187026072, accuracy: 0.7717948717948718

training 3 epoch...

69.52487754821777 seconds used......

3 training finished! average train loss: 0.15077808403417978

total clf loss: 36.44705206807703 total adv loss: 23.29960901965387

evaluating...

Confusion matrix, without normalization

[[ 61 24 11]

[ 36 259 60]

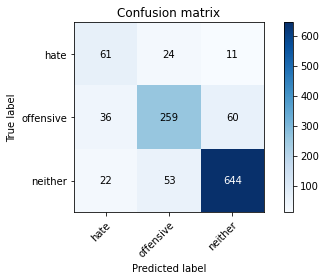
[ 22 53 644]]

Normalized confusion matrix

[[0.63541667 0.25 0.11458333]

[0.10140845 0.72957746 0.16901408]

[0.03059805 0.07371349 0.89568846]]





precision recall f1-score support

hate 0.51 0.64 0.57 96

offensive 0.77 0.73 0.75 355

neither 0.90 0.90 0.90 719

accuracy 0.82 1170

macro avg 0.73 0.75 0.74 1170

weighted avg 0.83 0.82 0.83 1170

average eval\_loss: 0.5941028897571614, accuracy: 0.8239316239316239

training 4 epoch...

69.48050355911255 seconds used......

4 training finished! average train loss: 0.05897883335416977

total clf loss: 14.265094051952474 total adv loss: 9.28094191828859

evaluating...

Confusion matrix, without normalization

[[ 51 29 16]

[ 17 271 67]

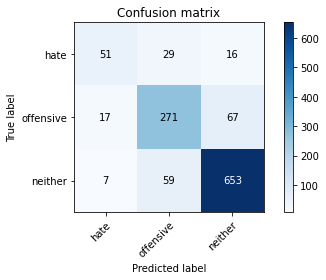
[ 7 59 653]]

Normalized confusion matrix

[[0.53125 0.30208333 0.16666667]

[0.04788732 0.76338028 0.18873239]

[0.00973574 0.08205841 0.90820584]]





precision recall f1-score support

hate 0.68 0.53 0.60 96

offensive 0.75 0.76 0.76 355

neither 0.89 0.91 0.90 719

accuracy 0.83 1170

macro avg 0.77 0.73 0.75 1170

weighted avg 0.83 0.83 0.83 1170

average eval\_loss: 0.6620889148848542, accuracy: 0.8333333333333334

training 5 epoch...

69.65709233283997 seconds used......

5 training finished! average train loss: 0.04223392395806125

total clf loss: 10.258796235837508 total adv loss: 7.521161415075767

evaluating...

Confusion matrix, without normalization

[[ 36 29 31]

[ 11 234 110]

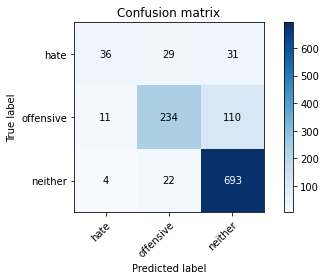
[ 4 22 693]]

Normalized confusion matrix

[[0.375 0.30208333 0.32291667]

[0.03098592 0.65915493 0.30985915]

[0.00556328 0.03059805 0.96383866]]





precision recall f1-score support

hate 0.71 0.38 0.49 96

offensive 0.82 0.66 0.73 355

neither 0.83 0.96 0.89 719

accuracy 0.82 1170

macro avg 0.79 0.67 0.70 1170

weighted avg 0.82 0.82 0.81 1170

average eval\_loss: 0.8567163866751737, accuracy: 0.823076923076923

training 6 epoch...

69.54965281486511 seconds used......

6 training finished! average train loss: 0.034907732494215235

total clf loss: 8.517930948291905 total adv loss: 6.9904314595405594

evaluating...

Confusion matrix, without normalization

[[ 56 30 10]

[ 26 286 43]

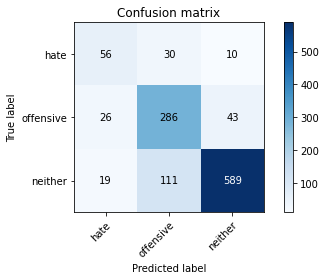
[ 19 111 589]]

Normalized confusion matrix

[[0.58333333 0.3125 0.10416667]

[0.07323944 0.8056338 0.12112676]

[0.02642559 0.15438108 0.81919332]]





precision recall f1-score support

hate 0.55 0.58 0.57 96

offensive 0.67 0.81 0.73 355

neither 0.92 0.82 0.87 719

accuracy 0.80 1170

macro avg 0.71 0.74 0.72 1170

weighted avg 0.81 0.80 0.80 1170

average eval\_loss: 0.8258468232164948, accuracy: 0.7957264957264957

training 7 epoch...

69.64589309692383 seconds used......

7 training finished! average train loss: 0.029066200695959773

total clf loss: 7.095607142604422 total adv loss: 5.882324000092922

evaluating...

Confusion matrix, without normalization

[[ 58 22 16]

[ 20 252 83]

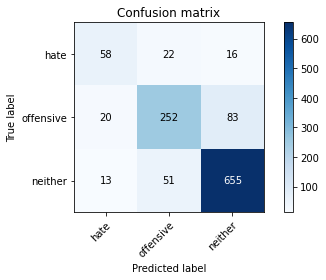
[ 13 51 655]]

Normalized confusion matrix

[[0.60416667 0.22916667 0.16666667]

[0.05633803 0.70985915 0.23380282]

[0.01808067 0.07093185 0.91098748]]





precision recall f1-score support

hate 0.64 0.60 0.62 96

offensive 0.78 0.71 0.74 355

neither 0.87 0.91 0.89 719

accuracy 0.82 1170

macro avg 0.76 0.74 0.75 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.7447664068727675, accuracy: 0.8247863247863247

training 8 epoch...

69.58306741714478 seconds used......

8 training finished! average train loss: 0.026981805635227635

total clf loss: 6.616690924522118 total adv loss: 6.058967103081159

evaluating...

Confusion matrix, without normalization

[[ 67 20 9]

[ 40 263 52]

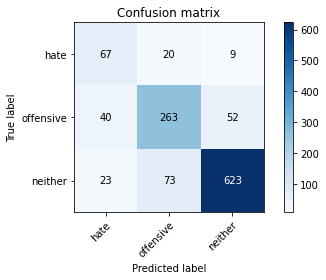
[ 23 73 623]]

Normalized confusion matrix

[[0.69791667 0.20833333 0.09375 ]

[0.11267606 0.74084507 0.14647887]

[0.03198887 0.1015299 0.86648122]]





precision recall f1-score support

hate 0.52 0.70 0.59 96

offensive 0.74 0.74 0.74 355

neither 0.91 0.87 0.89 719

accuracy 0.81 1170

macro avg 0.72 0.77 0.74 1170

weighted avg 0.83 0.81 0.82 1170

average eval\_loss: 0.6826080058590841, accuracy: 0.8145299145299145

training 9 epoch...

69.53523468971252 seconds used......

9 training finished! average train loss: 0.02812182388487809

total clf loss: 6.852826802933123 total adv loss: 5.44640061039172

evaluating...

Confusion matrix, without normalization

[[ 52 29 15]

[ 22 286 47]

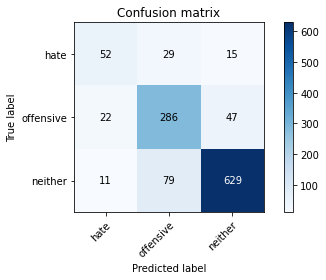
[ 11 79 629]]

Normalized confusion matrix

[[0.54166667 0.30208333 0.15625 ]

[0.06197183 0.8056338 0.13239437]

[0.01529903 0.10987483 0.87482615]]





precision recall f1-score support

hate 0.61 0.54 0.57 96

offensive 0.73 0.81 0.76 355

neither 0.91 0.87 0.89 719

accuracy 0.83 1170

macro avg 0.75 0.74 0.74 1170

weighted avg 0.83 0.83 0.83 1170

average eval\_loss: 0.6942630275073698, accuracy: 0.8264957264957264

training 10 epoch...

69.56602764129639 seconds used......

10 training finished! average train loss: 0.01682909626267954

total clf loss: 4.068939212389523 total adv loss: 2.6186136835567595

evaluating...

Confusion matrix, without normalization

[[ 35 42 19]

[ 7 268 80]

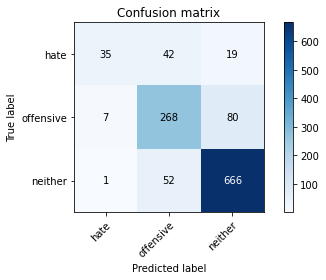
[ 1 52 666]]

Normalized confusion matrix

[[0.36458333 0.4375 0.19791667]

[0.01971831 0.75492958 0.22535211]

[0.00139082 0.07232267 0.92628651]]





precision recall f1-score support

hate 0.81 0.36 0.50 96

offensive 0.74 0.75 0.75 355

neither 0.87 0.93 0.90 719

accuracy 0.83 1170

macro avg 0.81 0.68 0.72 1170

weighted avg 0.83 0.83 0.82 1170

average eval\_loss: 0.8365717293887983, accuracy: 0.8282051282051283

4 run......

training 1 epoch...

69.69380164146423 seconds used......

1 training finished! average train loss: 0.589562410982246

total clf loss: 143.18343745172024 total adv loss: 104.51666784286499

evaluating...

Confusion matrix, without normalization

[[ 46 19 31]

[ 28 187 140]

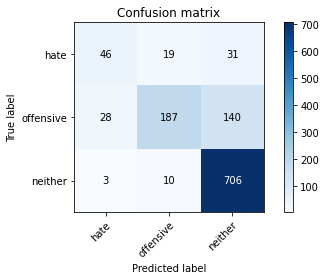
[ 3 10 706]]

Normalized confusion matrix

[[0.47916667 0.19791667 0.32291667]

[0.07887324 0.52676056 0.3943662 ]

[0.00417246 0.01390821 0.98191933]]





precision recall f1-score support

hate 0.60 0.48 0.53 96

offensive 0.87 0.53 0.65 355

neither 0.81 0.98 0.88 719

accuracy 0.80 1170

macro avg 0.76 0.66 0.69 1170

weighted avg 0.81 0.80 0.79 1170

average eval\_loss: 0.5049867819426423, accuracy: 0.8025641025641026

training 2 epoch...

69.66934418678284 seconds used......

2 training finished! average train loss: 0.3451157856541566

total clf loss: 83.74051504954696 total adv loss: 59.66842366848141

evaluating...

Confusion matrix, without normalization

[[ 39 46 11]

[ 7 303 45]

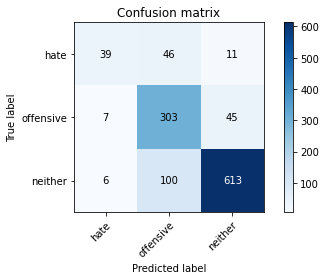
[ 6 100 613]]

Normalized confusion matrix

[[0.40625 0.47916667 0.11458333]

[0.01971831 0.85352113 0.12676056]

[0.00834492 0.13908206 0.85257302]]





precision recall f1-score support

hate 0.75 0.41 0.53 96

offensive 0.67 0.85 0.75 355

neither 0.92 0.85 0.88 719

accuracy 0.82 1170

macro avg 0.78 0.70 0.72 1170

weighted avg 0.83 0.82 0.81 1170

average eval\_loss: 0.4634056779540191, accuracy: 0.8162393162393162

training 3 epoch...

69.65421891212463 seconds used......

3 training finished! average train loss: 0.18317729613592482

total clf loss: 44.259647119790316 total adv loss: 27.923197479452938

evaluating...

Confusion matrix, without normalization

[[ 43 31 22]

[ 16 260 79]

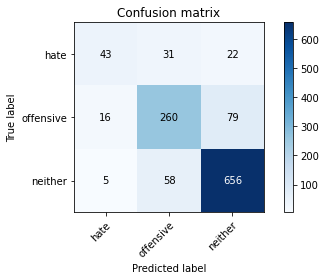
[ 5 58 656]]

Normalized confusion matrix

[[0.44791667 0.32291667 0.22916667]

[0.04507042 0.73239437 0.22253521]

[0.0069541 0.08066759 0.9123783 ]]





precision recall f1-score support

hate 0.67 0.45 0.54 96

offensive 0.74 0.73 0.74 355

neither 0.87 0.91 0.89 719

accuracy 0.82 1170

macro avg 0.76 0.70 0.72 1170

weighted avg 0.81 0.82 0.81 1170

average eval\_loss: 0.5018627870259649, accuracy: 0.8196581196581196

training 4 epoch...

69.59373354911804 seconds used......

4 training finished! average train loss: 0.08940752376961467

total clf loss: 21.693164257798344 total adv loss: 15.436071478528902

evaluating...

Confusion matrix, without normalization

[[ 37 43 16]

[ 9 286 60]

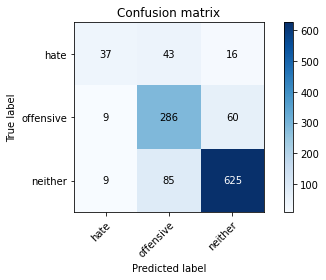
[ 9 85 625]]

Normalized confusion matrix

[[0.38541667 0.44791667 0.16666667]

[0.02535211 0.8056338 0.16901408]

[0.01251739 0.11821975 0.86926287]]





precision recall f1-score support

hate 0.67 0.39 0.49 96

offensive 0.69 0.81 0.74 355

neither 0.89 0.87 0.88 719

accuracy 0.81 1170

macro avg 0.75 0.69 0.70 1170

weighted avg 0.81 0.81 0.81 1170

average eval\_loss: 0.6188943857737517, accuracy: 0.8102564102564103

training 5 epoch...

69.65118026733398 seconds used......

5 training finished! average train loss: 0.05064919825845286

total clf loss: 12.331248883157969 total adv loss: 9.58672927270527

evaluating...

Confusion matrix, without normalization

[[ 38 31 27]

[ 12 261 82]

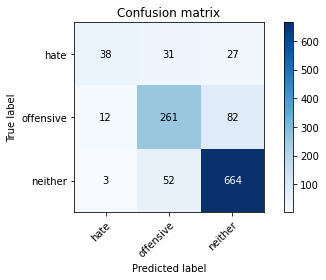
[ 3 52 664]]

Normalized confusion matrix

[[0.39583333 0.32291667 0.28125 ]

[0.03380282 0.73521127 0.23098592]

[0.00417246 0.07232267 0.92350487]]





precision recall f1-score support

hate 0.72 0.40 0.51 96

offensive 0.76 0.74 0.75 355

neither 0.86 0.92 0.89 719

accuracy 0.82 1170

macro avg 0.78 0.68 0.72 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.7498299516461221, accuracy: 0.823076923076923

training 6 epoch...

69.77603936195374 seconds used......

6 training finished! average train loss: 0.03523704818881248

total clf loss: 8.587051093403716 total adv loss: 6.831636300259561

evaluating...

Confusion matrix, without normalization

[[ 43 33 20]

[ 13 272 70]

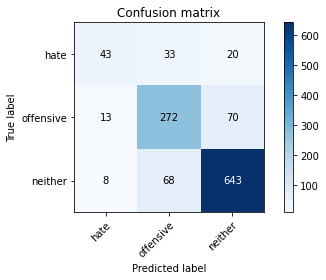
[ 8 68 643]]

Normalized confusion matrix

[[0.44791667 0.34375 0.20833333]

[0.03661972 0.76619718 0.1971831 ]

[0.01112656 0.0945758 0.89429764]]





precision recall f1-score support

hate 0.67 0.45 0.54 96

offensive 0.73 0.77 0.75 355

neither 0.88 0.89 0.89 719

accuracy 0.82 1170

macro avg 0.76 0.70 0.72 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.8460438477814639, accuracy: 0.8188034188034188

training 7 epoch...

69.63683938980103 seconds used......

7 training finished! average train loss: 0.04296445419950188

total clf loss: 10.442098582381732 total adv loss: 7.768325610923057

evaluating...

Confusion matrix, without normalization

[[ 33 39 24]

[ 5 273 77]

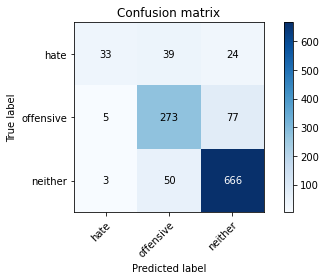
[ 3 50 666]]

Normalized confusion matrix

[[0.34375 0.40625 0.25 ]

[0.01408451 0.76901408 0.21690141]

[0.00417246 0.06954103 0.92628651]]





precision recall f1-score support

hate 0.80 0.34 0.48 96

offensive 0.75 0.77 0.76 355

neither 0.87 0.93 0.90 719

accuracy 0.83 1170

macro avg 0.81 0.68 0.71 1170

weighted avg 0.83 0.83 0.82 1170

average eval\_loss: 0.7817658180217022, accuracy: 0.8307692307692308

training 8 epoch...

69.6518964767456 seconds used......

8 training finished! average train loss: 0.025083662777503308

total clf loss: 6.122543017670978 total adv loss: 5.0593185257093864

evaluating...

Confusion matrix, without normalization

[[ 40 35 21]

[ 15 266 74]

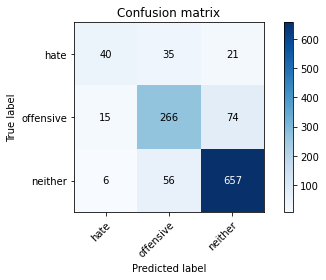
[ 6 56 657]]

Normalized confusion matrix

[[0.41666667 0.36458333 0.21875 ]

[0.04225352 0.74929577 0.2084507 ]

[0.00834492 0.07788595 0.91376912]]





precision recall f1-score support

hate 0.66 0.42 0.51 96

offensive 0.75 0.75 0.75 355

neither 0.87 0.91 0.89 719

accuracy 0.82 1170

macro avg 0.76 0.69 0.72 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.8781884210628174, accuracy: 0.823076923076923

training 9 epoch...

69.79190731048584 seconds used......

9 training finished! average train loss: 0.020084239648389887

total clf loss: 4.872493370537995 total adv loss: 3.4556253160408232

evaluating...

Confusion matrix, without normalization

[[ 26 48 22]

[ 5 287 63]

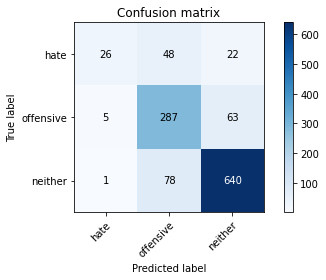
[ 1 78 640]]

Normalized confusion matrix

[[0.27083333 0.5 0.22916667]

[0.01408451 0.8084507 0.17746479]

[0.00139082 0.10848401 0.89012517]]





precision recall f1-score support

hate 0.81 0.27 0.41 96

offensive 0.69 0.81 0.75 355

neither 0.88 0.89 0.89 719

accuracy 0.81 1170

macro avg 0.80 0.66 0.68 1170

weighted avg 0.82 0.81 0.80 1170

average eval\_loss: 0.9339817558115316, accuracy: 0.8145299145299145

training 10 epoch...

69.66400218009949 seconds used......

10 training finished! average train loss: 0.02383531880509392

total clf loss: 5.816419193426555 total adv loss: 4.779091834912833

evaluating...

Confusion matrix, without normalization

[[ 59 18 19]

[ 47 242 66]

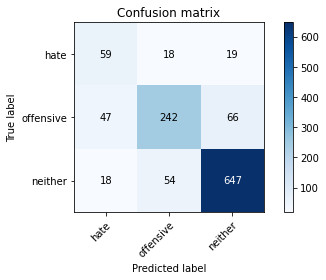
[ 18 54 647]]

Normalized confusion matrix

[[0.61458333 0.1875 0.19791667]

[0.13239437 0.68169014 0.18591549]

[0.02503477 0.07510431 0.89986092]]





precision recall f1-score support

hate 0.48 0.61 0.54 96

offensive 0.77 0.68 0.72 355

neither 0.88 0.90 0.89 719

accuracy 0.81 1170

macro avg 0.71 0.73 0.72 1170

weighted avg 0.82 0.81 0.81 1170

average eval\_loss: 0.9723987375604092, accuracy: 0.8102564102564103

5 run......

training 1 epoch...

69.51811075210571 seconds used......

1 training finished! average train loss: 0.6051977359586291

total clf loss: 146.97895480692387 total adv loss: 107.25369618833065

evaluating...

Confusion matrix, without normalization

[[ 28 49 19]

[ 3 278 74]

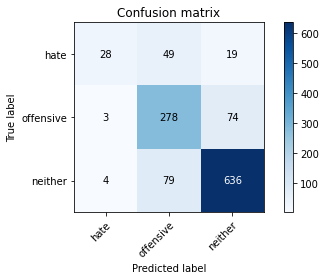
[ 4 79 636]]

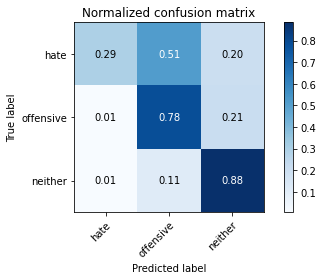
Normalized confusion matrix

[[0.29166667 0.51041667 0.19791667]

[0.0084507 0.78309859 0.2084507 ]

[0.00556328 0.10987483 0.88456189]]





precision recall f1-score support

hate 0.80 0.29 0.43 96

offensive 0.68 0.78 0.73 355

neither 0.87 0.88 0.88 719

accuracy 0.81 1170

macro avg 0.79 0.65 0.68 1170

weighted avg 0.81 0.81 0.80 1170

average eval\_loss: 0.4768974599191698, accuracy: 0.8051282051282052

training 2 epoch...

69.57428884506226 seconds used......

2 training finished! average train loss: 0.36474652846272176

total clf loss: 88.51444888859987 total adv loss: 63.27522253617644

evaluating...

Confusion matrix, without normalization

[[ 62 19 15]

[ 33 246 76]

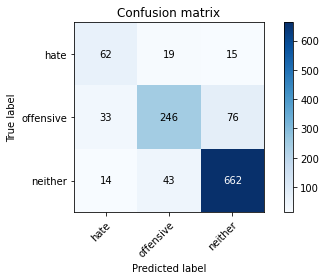
[ 14 43 662]]

Normalized confusion matrix

[[0.64583333 0.19791667 0.15625 ]

[0.09295775 0.69295775 0.21408451]

[0.01947149 0.05980529 0.92072323]]





precision recall f1-score support

hate 0.57 0.65 0.60 96

offensive 0.80 0.69 0.74 355

neither 0.88 0.92 0.90 719

accuracy 0.83 1170

macro avg 0.75 0.75 0.75 1170

weighted avg 0.83 0.83 0.83 1170

average eval\_loss: 0.43004880915001287, accuracy: 0.8290598290598291

training 3 epoch...

69.64427280426025 seconds used......

3 training finished! average train loss: 0.17289539963468656

total clf loss: 41.93206695187837 total adv loss: 29.490868457593024

evaluating...

Confusion matrix, without normalization

[[ 58 24 14]

[ 18 270 67]

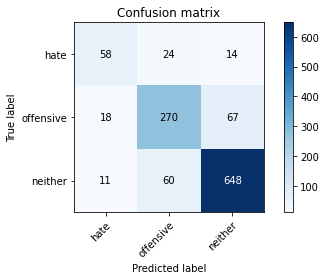
[ 11 60 648]]

Normalized confusion matrix

[[0.60416667 0.25 0.14583333]

[0.05070423 0.76056338 0.18873239]

[0.01529903 0.08344924 0.90125174]]





precision recall f1-score support

hate 0.67 0.60 0.63 96

offensive 0.76 0.76 0.76 355

neither 0.89 0.90 0.90 719

accuracy 0.83 1170

macro avg 0.77 0.76 0.76 1170

weighted avg 0.83 0.83 0.83 1170

average eval\_loss: 0.49337433647932644, accuracy: 0.8341880341880342

training 4 epoch...

69.63672018051147 seconds used......

4 training finished! average train loss: 0.0955622341368411

total clf loss: 23.117181729059666 total adv loss: 15.11238022299949

evaluating...

Confusion matrix, without normalization

[[ 48 24 24]

[ 16 260 79]

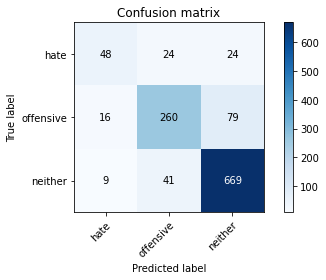
[ 9 41 669]]

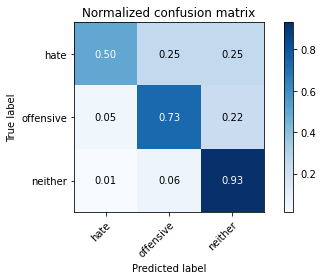
Normalized confusion matrix

[[0.5 0.25 0.25 ]

[0.04507042 0.73239437 0.22253521]

[0.01251739 0.05702364 0.93045897]]





precision recall f1-score support

hate 0.66 0.50 0.57 96

offensive 0.80 0.73 0.76 355

neither 0.87 0.93 0.90 719

accuracy 0.84 1170

macro avg 0.77 0.72 0.74 1170

weighted avg 0.83 0.84 0.83 1170

average eval\_loss: 0.707722237784518, accuracy: 0.8350427350427351

training 5 epoch...

69.59898447990417 seconds used......

5 training finished! average train loss: 0.04610383540638683

total clf loss: 11.268397033796646 total adv loss: 9.60199174634181

evaluating...

Confusion matrix, without normalization

[[ 48 33 15]

[ 15 279 61]

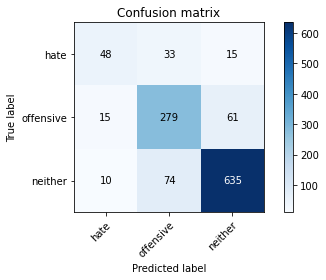
[ 10 74 635]]

Normalized confusion matrix

[[0.5 0.34375 0.15625 ]

[0.04225352 0.78591549 0.17183099]

[0.01390821 0.10292072 0.88317107]]





precision recall f1-score support

hate 0.66 0.50 0.57 96

offensive 0.72 0.79 0.75 355

neither 0.89 0.88 0.89 719

accuracy 0.82 1170

macro avg 0.76 0.72 0.74 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.764222967959309, accuracy: 0.8222222222222222

training 6 epoch...

69.6528549194336 seconds used......

6 training finished! average train loss: 0.0263378333621738

total clf loss: 6.425555746245664 total adv loss: 5.250054767224356

evaluating...

Confusion matrix, without normalization

[[ 35 35 26]

[ 5 278 72]

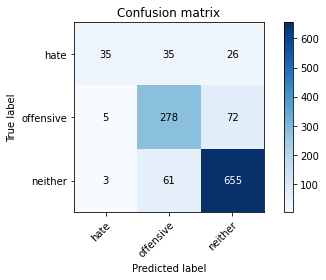
[ 3 61 655]]

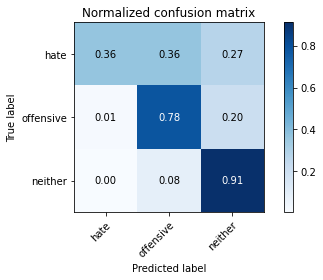
Normalized confusion matrix

[[0.36458333 0.36458333 0.27083333]

[0.01408451 0.78309859 0.2028169 ]

[0.00417246 0.08484006 0.91098748]]





precision recall f1-score support

hate 0.81 0.36 0.50 96

offensive 0.74 0.78 0.76 355

neither 0.87 0.91 0.89 719

accuracy 0.83 1170

macro avg 0.81 0.69 0.72 1170

weighted avg 0.83 0.83 0.82 1170

average eval\_loss: 0.8561692608054727, accuracy: 0.8273504273504273

training 7 epoch...

69.66049647331238 seconds used......

7 training finished! average train loss: 0.03681430137247588

total clf loss: 8.953348587499931 total adv loss: 6.776040748783998

evaluating...

Confusion matrix, without normalization

[[ 56 26 14]

[ 30 267 58]

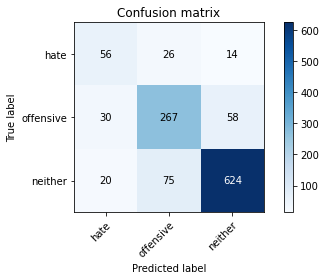
[ 20 75 624]]

Normalized confusion matrix

[[0.58333333 0.27083333 0.14583333]

[0.08450704 0.75211268 0.16338028]

[0.02781641 0.10431154 0.86787204]]





precision recall f1-score support

hate 0.53 0.58 0.55 96

offensive 0.73 0.75 0.74 355

neither 0.90 0.87 0.88 719

accuracy 0.81 1170

macro avg 0.72 0.73 0.73 1170

weighted avg 0.81 0.81 0.81 1170

average eval\_loss: 0.7775762681375747, accuracy: 0.8094017094017094

training 8 epoch...

69.62484431266785 seconds used......

8 training finished! average train loss: 0.023658570232273787

total clf loss: 5.7243576314649545 total adv loss: 3.765043737425003

evaluating...

Confusion matrix, without normalization

[[ 38 40 18]

[ 9 280 66]

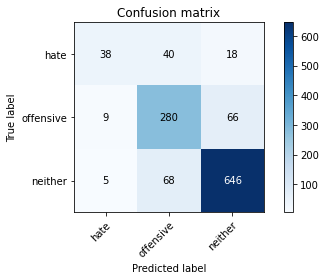
[ 5 68 646]]

Normalized confusion matrix

[[0.39583333 0.41666667 0.1875 ]

[0.02535211 0.78873239 0.18591549]

[0.0069541 0.0945758 0.8984701 ]]





precision recall f1-score support

hate 0.73 0.40 0.51 96

offensive 0.72 0.79 0.75 355

neither 0.88 0.90 0.89 719

accuracy 0.82 1170

macro avg 0.78 0.69 0.72 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.9003030006147051, accuracy: 0.8239316239316239

training 9 epoch...

69.60434556007385 seconds used......

9 training finished! average train loss: 0.027623064719109684

total clf loss: 6.691122298740083 total adv loss: 4.54650299957575

evaluating...

Confusion matrix, without normalization

[[ 52 25 19]

[ 19 262 74]

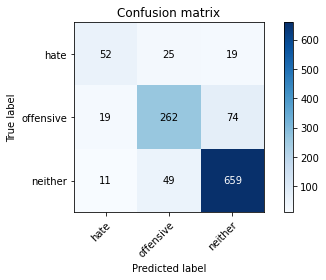
[ 11 49 659]]

Normalized confusion matrix

[[0.54166667 0.26041667 0.19791667]

[0.05352113 0.73802817 0.2084507 ]

[0.01529903 0.06815021 0.91655076]]





precision recall f1-score support

hate 0.63 0.54 0.58 96

offensive 0.78 0.74 0.76 355

neither 0.88 0.92 0.90 719

accuracy 0.83 1170

macro avg 0.76 0.73 0.75 1170

weighted avg 0.83 0.83 0.83 1170

average eval\_loss: 0.9047097259692422, accuracy: 0.8316239316239317

training 10 epoch...

69.47082018852234 seconds used......

10 training finished! average train loss: 0.023192077351641506

total clf loss: 5.701999441880616 total adv loss: 5.501067031014827

evaluating...

Confusion matrix, without normalization

[[ 67 18 11]

[ 36 253 66]

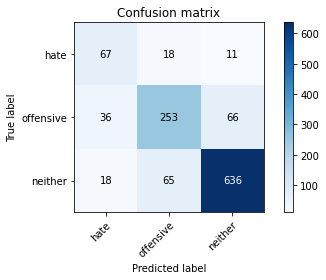
[ 18 65 636]]

Normalized confusion matrix

[[0.69791667 0.1875 0.11458333]

[0.10140845 0.71267606 0.18591549]

[0.02503477 0.09040334 0.88456189]]





precision recall f1-score support

hate 0.55 0.70 0.62 96

offensive 0.75 0.71 0.73 355

neither 0.89 0.88 0.89 719

accuracy 0.82 1170

macro avg 0.73 0.77 0.75 1170

weighted avg 0.82 0.82 0.82 1170

average eval\_loss: 0.7744145888314282, accuracy: 0.8170940170940171