alpha = 2

1 run...

training 1 epoch...

388.20374846458435 seconds used......

1 training finished! average train loss: 1.3099574809957035

total clf loss: 274.32753452658653 total adv loss: 152.99181487038732

evaluating...

Confusion matrix, without normalization

[[ 733 133 156]

[1021 4261 125]

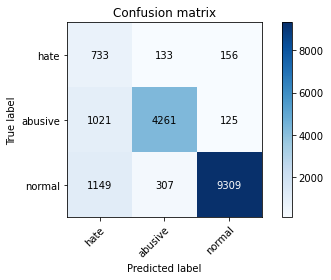
[1149 307 9309]]

Normalized confusion matrix

[[0.71722114 0.13013699 0.15264188]

[0.1888293 0.78805252 0.02311818]

[0.10673479 0.02851835 0.86474686]]





precision recall f1-score support

hate 0.25 0.72 0.37 1022

abusive 0.91 0.79 0.84 5407

normal 0.97 0.86 0.91 10765

accuracy 0.83 17194

macro avg 0.71 0.79 0.71 17194

weighted avg 0.91 0.83 0.86 17194

average eval\_loss: 0.45960437600002735, accuracy: 0.8318599511457485

training 2 epoch...

388.3958294391632 seconds used......

2 training finished! average train loss: 0.9770498096135615

total clf loss: 203.75842190533876 total adv loss: 114.53732233867049

evaluating...

Confusion matrix, without normalization

[[ 631 208 183]

[ 626 4640 141]

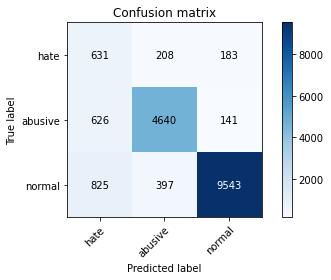
[ 825 397 9543]]

Normalized confusion matrix

[[0.61741683 0.2035225 0.17906067]

[0.11577585 0.85814685 0.02607731]

[0.07663725 0.03687877 0.88648398]]





precision recall f1-score support

hate 0.30 0.62 0.41 1022

abusive 0.88 0.86 0.87 5407

normal 0.97 0.89 0.93 10765

accuracy 0.86 17194

macro avg 0.72 0.79 0.73 17194

weighted avg 0.90 0.86 0.88 17194

average eval\_loss: 0.3918011279213567, accuracy: 0.8615796207979528

training 3 epoch...

388.43953371047974 seconds used......

3 training finished! average train loss: 0.7829454545242792

total clf loss: 163.46779028326273 total adv loss: 91.688522873912

evaluating...

Confusion matrix, without normalization

[[ 669 198 155]

[ 795 4475 137]

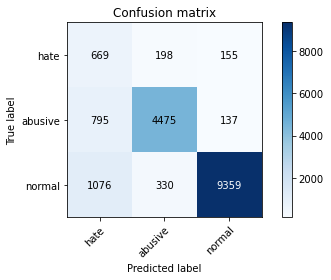
[1076 330 9359]]

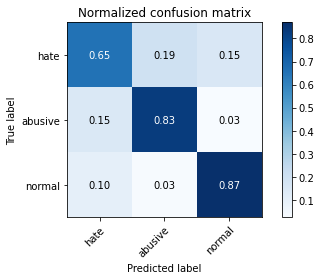
Normalized confusion matrix

[[0.65459883 0.19373777 0.15166341]

[0.14703163 0.82763085 0.02533753]

[0.09995355 0.0306549 0.86939155]]





precision recall f1-score support

hate 0.26 0.65 0.38 1022

abusive 0.89 0.83 0.86 5407

normal 0.97 0.87 0.92 10765

accuracy 0.84 17194

macro avg 0.71 0.78 0.72 17194

weighted avg 0.90 0.84 0.87 17194

average eval\_loss: 0.4392200190270709, accuracy: 0.843491915784576

training 4 epoch...

388.17922711372375 seconds used......

4 training finished! average train loss: 0.5578764057287244

total clf loss: 116.90667075663805 total adv loss: 65.11628869036213

evaluating...

Confusion matrix, without normalization

[[ 628 211 183]

[ 821 4438 148]

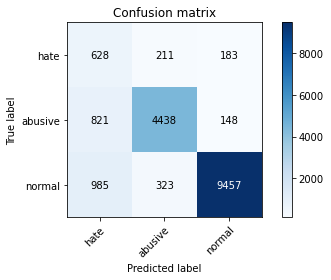
[ 985 323 9457]]

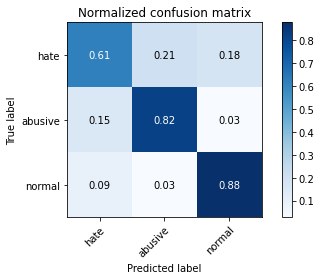
Normalized confusion matrix

[[0.61448141 0.20645793 0.17906067]

[0.15184021 0.82078787 0.02737193]

[0.09150023 0.03000464 0.87849512]]





precision recall f1-score support

hate 0.26 0.61 0.36 1022

abusive 0.89 0.82 0.86 5407

normal 0.97 0.88 0.92 10765

accuracy 0.84 17194

macro avg 0.71 0.77 0.71 17194

weighted avg 0.90 0.84 0.87 17194

average eval\_loss: 0.5026107279670446, accuracy: 0.8446551122484588

training 5 epoch...

388.34766244888306 seconds used......

5 training finished! average train loss: 0.4069455279874075

total clf loss: 81.6668545641005 total adv loss: 49.305007343529724

evaluating...

Confusion matrix, without normalization

[[ 706 200 116]

[ 897 4405 105]

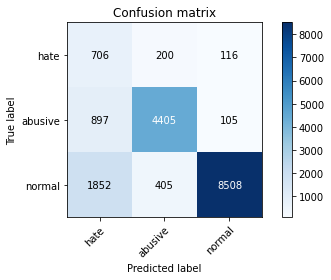
[1852 405 8508]]

Normalized confusion matrix

[[0.69080235 0.19569472 0.11350294]

[0.16589606 0.81468467 0.01941927]

[0.17203902 0.03762192 0.79033906]]





precision recall f1-score support

hate 0.20 0.69 0.32 1022

abusive 0.88 0.81 0.85 5407

normal 0.97 0.79 0.87 10765

accuracy 0.79 17194

macro avg 0.69 0.77 0.68 17194

weighted avg 0.90 0.79 0.83 17194

average eval\_loss: 0.7030194097263522, accuracy: 0.7920786320809585

2 run...

training 1 epoch...

388.24802350997925 seconds used......

1 training finished! average train loss: 1.2970297814895553

total clf loss: 271.4989049434662 total adv loss: 151.54264380037785

evaluating...

Confusion matrix, without normalization

[[ 594 240 188]

[ 516 4743 148]

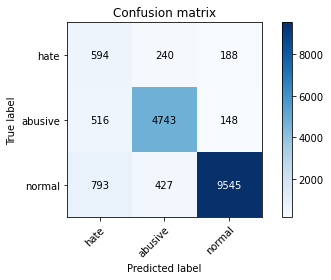
[ 793 427 9545]]

Normalized confusion matrix

[[0.58121331 0.23483366 0.18395303]

[0.09543185 0.87719623 0.02737193]

[0.07366465 0.03966558 0.88666976]]





precision recall f1-score support

hate 0.31 0.58 0.41 1022

abusive 0.88 0.88 0.88 5407

normal 0.97 0.89 0.92 10765

accuracy 0.87 17194

macro avg 0.72 0.78 0.74 17194

weighted avg 0.90 0.87 0.88 17194

average eval\_loss: 0.4059976750161759, accuracy: 0.8655344887751542

training 2 epoch...

388.42808389663696 seconds used......

2 training finished! average train loss: 0.9760745175805102

total clf loss: 203.6806470900774 total adv loss: 114.36018198356032

evaluating...

Confusion matrix, without normalization

[[ 713 172 137]

[ 916 4370 121]

[1337 294 9134]]

Normalized confusion matrix

[[0.69765166 0.16829746 0.13405088]

[0.16941002 0.80821158 0.0223784 ]

[0.12419879 0.02731073 0.84849048]]





precision recall f1-score support

hate 0.24 0.70 0.36 1022

abusive 0.90 0.81 0.85 5407

normal 0.97 0.85 0.91 10765

accuracy 0.83 17194

macro avg 0.71 0.78 0.71 17194

weighted avg 0.91 0.83 0.86 17194

average eval\_loss: 0.46725858855386115, accuracy: 0.8268582063510527

training 3 epoch...

388.484920501709 seconds used......

3 training finished! average train loss: 0.7785445390180863

total clf loss: 161.95453857257962 total adv loss: 91.47034534066916

evaluating...

Confusion matrix, without normalization

[[ 687 184 151]

[ 919 4365 123]

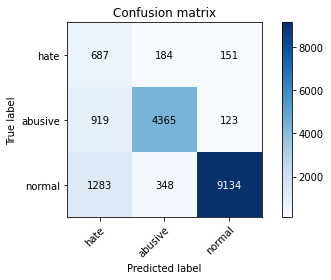
[1283 348 9134]]

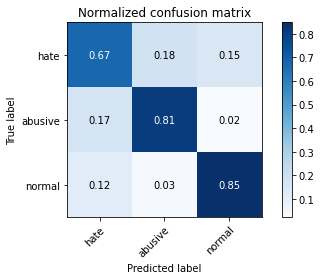
Normalized confusion matrix

[[0.67221135 0.18003914 0.14774951]

[0.16996486 0.80728685 0.02274829]

[0.11918254 0.03232699 0.84849048]]





precision recall f1-score support

hate 0.24 0.67 0.35 1022

abusive 0.89 0.81 0.85 5407

normal 0.97 0.85 0.91 10765

accuracy 0.83 17194

macro avg 0.70 0.78 0.70 17194

weighted avg 0.90 0.83 0.85 17194

average eval\_loss: 0.4822523349982708, accuracy: 0.8250552518320344

training 4 epoch...

388.48399209976196 seconds used......

4 training finished! average train loss: 0.5640368336984051

total clf loss: 117.62494169175625 total adv loss: 66.12168790539727

evaluating...

Confusion matrix, without normalization

[[ 725 174 123]

[ 996 4285 126]

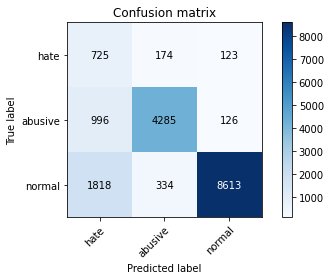
[1818 334 8613]]

Normalized confusion matrix

[[0.70939335 0.1702544 0.12035225]

[0.18420566 0.79249122 0.02330313]

[0.16888063 0.03102647 0.80009289]]





precision recall f1-score support

hate 0.20 0.71 0.32 1022

abusive 0.89 0.79 0.84 5407

normal 0.97 0.80 0.88 10765

accuracy 0.79 17194

macro avg 0.69 0.77 0.68 17194

weighted avg 0.90 0.79 0.83 17194

average eval\_loss: 0.6395691578817922, accuracy: 0.792311271373735

training 5 epoch...

388.77692198753357 seconds used......

5 training finished! average train loss: 0.396096327819317

total clf loss: 81.87036108784378 total adv loss: 46.80015591799747

evaluating...

Confusion matrix, without normalization

[[ 616 216 190]

[ 753 4466 188]

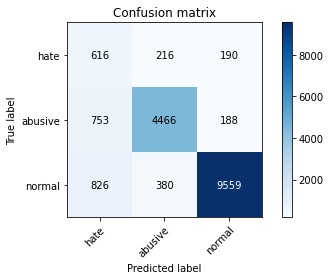
[ 826 380 9559]]

Normalized confusion matrix

[[0.60273973 0.21135029 0.18590998]

[0.13926392 0.82596634 0.03476974]

[0.07673014 0.03529958 0.88797027]]





precision recall f1-score support

hate 0.28 0.60 0.38 1022

abusive 0.88 0.83 0.85 5407

normal 0.96 0.89 0.92 10765

accuracy 0.85 17194

macro avg 0.71 0.77 0.72 17194

weighted avg 0.90 0.85 0.87 17194

average eval\_loss: 0.5320006965031457, accuracy: 0.851517971385367

3 run...

training 1 epoch...

389.0773298740387 seconds used......

1 training finished! average train loss: 1.2925557393589084

total clf loss: 270.20676662027836 total adv loss: 151.19771315343678

evaluating...

Confusion matrix, without normalization

[[ 753 137 132]

[1078 4221 108]

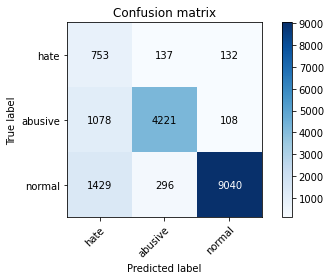
[1429 296 9040]]

Normalized confusion matrix

[[0.73679061 0.13405088 0.12915851]

[0.19937119 0.78065471 0.01997411]

[0.13274501 0.02749652 0.83975848]]





precision recall f1-score support

hate 0.23 0.74 0.35 1022

abusive 0.91 0.78 0.84 5407

normal 0.97 0.84 0.90 10765

accuracy 0.82 17194

macro avg 0.70 0.79 0.70 17194

weighted avg 0.91 0.82 0.85 17194

average eval\_loss: 0.4994280723465043, accuracy: 0.8150517622426428

training 2 epoch...

388.90748167037964 seconds used......

2 training finished! average train loss: 0.9758776652355495

total clf loss: 202.8096987605095 total adv loss: 114.75205313786864

evaluating...

Confusion matrix, without normalization

[[ 720 184 118]

[ 923 4386 98]

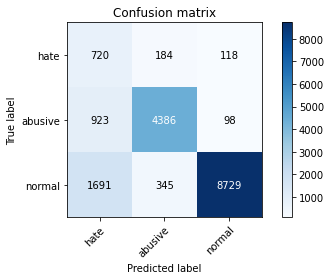
[1691 345 8729]]

Normalized confusion matrix

[[0.70450098 0.18003914 0.11545988]

[0.17070464 0.8111707 0.01812465]

[0.15708314 0.0320483 0.81086856]]





precision recall f1-score support

hate 0.22 0.70 0.33 1022

abusive 0.89 0.81 0.85 5407

normal 0.98 0.81 0.89 10765

accuracy 0.80 17194

macro avg 0.69 0.78 0.69 17194

weighted avg 0.90 0.80 0.84 17194

average eval\_loss: 0.5134441387688005, accuracy: 0.8046411538908922

training 3 epoch...

388.96313858032227 seconds used......

3 training finished! average train loss: 0.7899317829859176

total clf loss: 162.73823026940227 total adv loss: 93.60077489353716

evaluating...

Confusion matrix, without normalization

[[ 782 126 114]

[1248 4064 95]

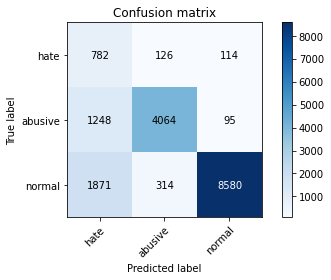
[1871 314 8580]]

Normalized confusion matrix

[[0.76516634 0.12328767 0.11154599]

[0.23081191 0.75161827 0.01756982]

[0.17380399 0.0291686 0.7970274 ]]





precision recall f1-score support

hate 0.20 0.77 0.32 1022

abusive 0.90 0.75 0.82 5407

normal 0.98 0.80 0.88 10765

accuracy 0.78 17194

macro avg 0.69 0.77 0.67 17194

weighted avg 0.91 0.78 0.83 17194

average eval\_loss: 0.6230523931252402, accuracy: 0.78085378620449

training 4 epoch...

388.51652097702026 seconds used......

4 training finished! average train loss: 0.540893701394982

total clf loss: 113.69015377014875 total adv loss: 62.962877887068316

evaluating...

Confusion matrix, without normalization

[[ 612 221 189]

[ 694 4529 184]

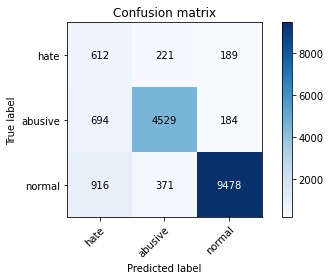
[ 916 371 9478]]

Normalized confusion matrix

[[0.59882583 0.21624266 0.18493151]

[0.12835214 0.8376179 0.03402996]

[0.08509057 0.03446354 0.88044589]]





precision recall f1-score support

hate 0.28 0.60 0.38 1022

abusive 0.88 0.84 0.86 5407

normal 0.96 0.88 0.92 10765

accuracy 0.85 17194

macro avg 0.71 0.77 0.72 17194

weighted avg 0.90 0.85 0.87 17194

average eval\_loss: 0.48660543840378523, accuracy: 0.850238455275096

training 5 epoch...

388.52517223358154 seconds used......

5 training finished! average train loss: 0.39159828573717625

total clf loss: 79.87133178673685 total adv loss: 46.80335450661369

evaluating...

Confusion matrix, without normalization

[[ 646 208 168]

[ 847 4417 143]

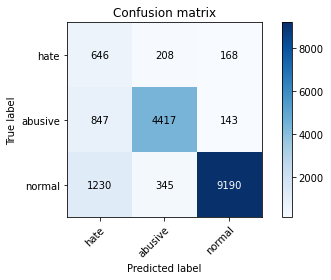
[1230 345 9190]]

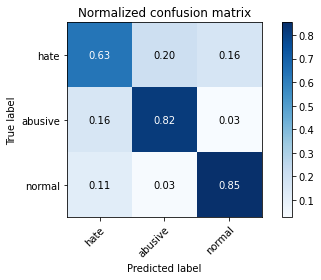
Normalized confusion matrix

[[0.63209393 0.2035225 0.16438356]

[0.15664879 0.81690401 0.0264472 ]

[0.11425917 0.0320483 0.85369252]]





precision recall f1-score support

hate 0.24 0.63 0.34 1022

abusive 0.89 0.82 0.85 5407

normal 0.97 0.85 0.91 10765

accuracy 0.83 17194

macro avg 0.70 0.77 0.70 17194

weighted avg 0.90 0.83 0.86 17194

average eval\_loss: 0.6134919847470037, accuracy: 0.8289519599860417

4 run...

training 1 epoch...

388.88596057891846 seconds used......

1 training finished! average train loss: 1.279474224718645

total clf loss: 268.2679592370987 total adv loss: 149.26956125348806

evaluating...

Confusion matrix, without normalization

[[ 631 230 161]

[ 579 4706 122]

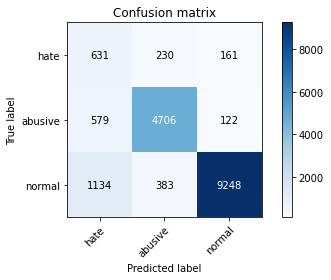
[1134 383 9248]]

Normalized confusion matrix

[[0.61741683 0.22504892 0.15753425]

[0.10708341 0.87035325 0.02256334]

[0.10534138 0.03557826 0.85908035]]





precision recall f1-score support

hate 0.27 0.62 0.37 1022

abusive 0.88 0.87 0.88 5407

normal 0.97 0.86 0.91 10765

accuracy 0.85 17194

macro avg 0.71 0.78 0.72 17194

weighted avg 0.90 0.85 0.87 17194

average eval\_loss: 0.4496383292903734, accuracy: 0.8482610212864953

training 2 epoch...

388.93143916130066 seconds used......

2 training finished! average train loss: 0.9750456777706104

total clf loss: 202.52823434770107 total adv loss: 114.70849976874888

evaluating...

Confusion matrix, without normalization

[[ 731 154 137]

[1077 4204 126]

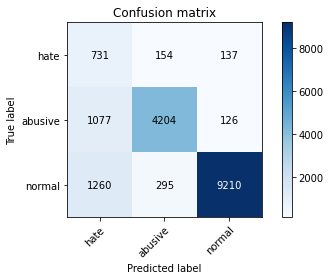
[1260 295 9210]]

Normalized confusion matrix

[[0.71526419 0.15068493 0.13405088]

[0.19918624 0.77751063 0.02330313]

[0.11704598 0.02740362 0.85555039]]





precision recall f1-score support

hate 0.24 0.72 0.36 1022

abusive 0.90 0.78 0.84 5407

normal 0.97 0.86 0.91 10765

accuracy 0.82 17194

macro avg 0.70 0.78 0.70 17194

weighted avg 0.91 0.82 0.85 17194

average eval\_loss: 0.49326849175227244, accuracy: 0.8226706990810748

training 3 epoch...

388.89395356178284 seconds used......

3 training finished! average train loss: 0.7683775745180875

total clf loss: 159.41134195029736 total adv loss: 90.48996166139841

evaluating...

Confusion matrix, without normalization

[[ 614 220 188]

[ 667 4582 158]

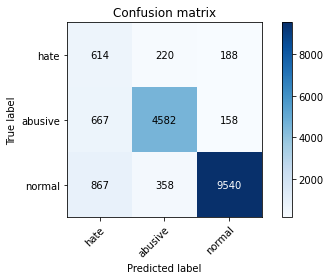
[ 867 358 9540]]

Normalized confusion matrix

[[0.60078278 0.21526419 0.18395303]

[0.12335861 0.84742001 0.02922138]

[0.08053878 0.03325592 0.88620529]]





precision recall f1-score support

hate 0.29 0.60 0.39 1022

abusive 0.89 0.85 0.87 5407

normal 0.97 0.89 0.92 10765

accuracy 0.86 17194

macro avg 0.71 0.78 0.73 17194

weighted avg 0.90 0.86 0.87 17194

average eval\_loss: 0.41980282858793816, accuracy: 0.8570431545888101

training 4 epoch...

388.8444895744324 seconds used......

4 training finished! average train loss: 0.5438843753785904

total clf loss: 111.94252711161971 total adv loss: 64.49912585271522

evaluating...

Confusion matrix, without normalization

[[ 679 196 147]

[ 936 4311 160]

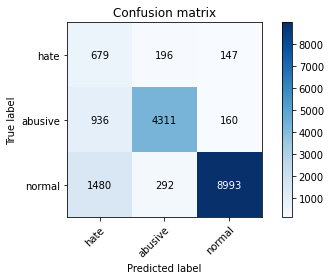
[1480 292 8993]]

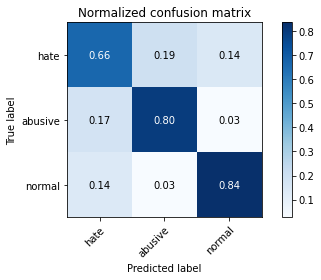
Normalized confusion matrix

[[0.66438356 0.19178082 0.14383562]

[0.17310893 0.7972998 0.02959127]

[0.13748258 0.02712494 0.83539248]]





precision recall f1-score support

hate 0.22 0.66 0.33 1022

abusive 0.90 0.80 0.84 5407

normal 0.97 0.84 0.90 10765

accuracy 0.81 17194

macro avg 0.69 0.77 0.69 17194

weighted avg 0.90 0.81 0.85 17194

average eval\_loss: 0.5699302599527115, accuracy: 0.8132488077236245

training 5 epoch...

388.8278012275696 seconds used......

5 training finished! average train loss: 0.3748973916846056

total clf loss: 75.52074952144176 total adv loss: 45.27939768007491

evaluating...

Confusion matrix, without normalization

[[ 716 184 122]

[ 920 4378 109]

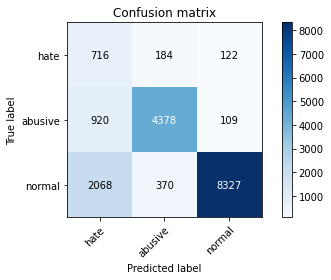
[2068 370 8327]]

Normalized confusion matrix

[[0.70058708 0.18003914 0.11937378]

[0.17014981 0.80969114 0.02015905]

[0.19210404 0.03437065 0.77352531]]





precision recall f1-score support

hate 0.19 0.70 0.30 1022

abusive 0.89 0.81 0.85 5407

normal 0.97 0.77 0.86 10765

accuracy 0.78 17194

macro avg 0.68 0.76 0.67 17194

weighted avg 0.90 0.78 0.82 17194

average eval\_loss: 0.7426745507692875, accuracy: 0.7805629870885192

5 run...

training 1 epoch...

388.76948285102844 seconds used......

1 training finished! average train loss: 1.2795705094967147

total clf loss: 268.74236853420734 total adv loss: 149.05368335172534

evaluating...

Confusion matrix, without normalization

[[ 742 154 126]

[ 918 4396 93]

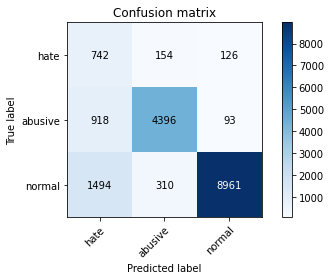
[1494 310 8961]]

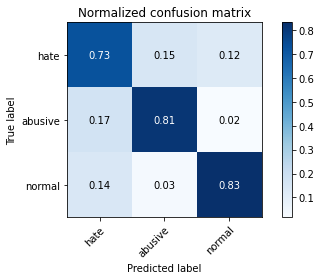
Normalized confusion matrix

[[0.7260274 0.15068493 0.12328767]

[0.16977991 0.81302016 0.01719993]

[0.13878309 0.02879703 0.83241988]]





precision recall f1-score support

hate 0.24 0.73 0.36 1022

abusive 0.90 0.81 0.86 5407

normal 0.98 0.83 0.90 10765

accuracy 0.82 17194

macro avg 0.71 0.79 0.70 17194

weighted avg 0.91 0.82 0.85 17194

average eval\_loss: 0.49814002299031546, accuracy: 0.8199953472141445

training 2 epoch...

388.45660400390625 seconds used......

2 training finished! average train loss: 0.9758382968385925

total clf loss: 203.12923155725002 total adv loss: 114.58356731384993

evaluating...

Confusion matrix, without normalization

[[ 643 218 161]

[ 563 4690 154]

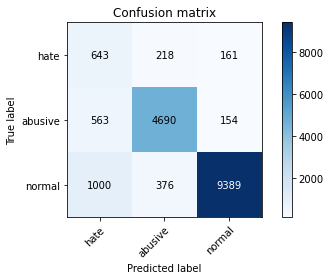
[1000 376 9389]]

Normalized confusion matrix

[[0.62915851 0.21330724 0.15753425]

[0.10412428 0.86739412 0.0284816 ]

[0.09289364 0.03492801 0.87217836]]





precision recall f1-score support

hate 0.29 0.63 0.40 1022

abusive 0.89 0.87 0.88 5407

normal 0.97 0.87 0.92 10765

accuracy 0.86 17194

macro avg 0.72 0.79 0.73 17194

weighted avg 0.90 0.86 0.87 17194

average eval\_loss: 0.41231389408714547, accuracy: 0.8562289170640921

training 3 epoch...

388.61901783943176 seconds used......

3 training finished! average train loss: 0.781544046228411

total clf loss: 163.35841096192598 total adv loss: 91.43280060775578

evaluating...

Confusion matrix, without normalization

[[ 617 205 200]

[ 737 4465 205]

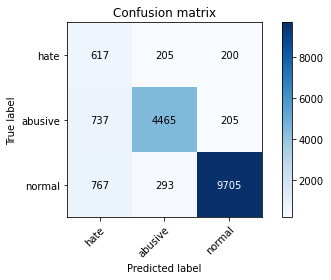
[ 767 293 9705]]

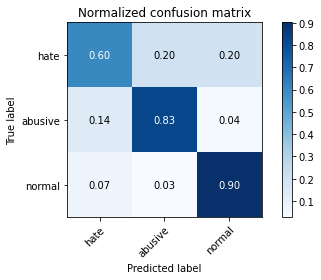
Normalized confusion matrix

[[0.6037182 0.20058708 0.19569472]

[0.13630479 0.82578139 0.03791382]

[0.07124942 0.02721784 0.90153275]]





precision recall f1-score support

hate 0.29 0.60 0.39 1022

abusive 0.90 0.83 0.86 5407

normal 0.96 0.90 0.93 10765

accuracy 0.86 17194

macro avg 0.72 0.78 0.73 17194

weighted avg 0.90 0.86 0.88 17194

average eval\_loss: 0.42408878996015287, accuracy: 0.8600093055717111

training 4 epoch...

388.48212814331055 seconds used......

4 training finished! average train loss: 0.5677535741528996

total clf loss: 117.90239072218537 total adv loss: 66.80622167885303

evaluating...

Confusion matrix, without normalization

[[ 734 155 133]

[1113 4176 118]

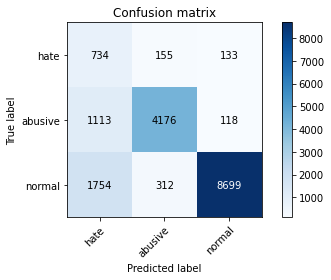
[1754 312 8699]]

Normalized confusion matrix

[[0.71819961 0.15166341 0.13013699]

[0.20584428 0.77233216 0.02182356]

[0.16293544 0.02898281 0.80808175]]





precision recall f1-score support

hate 0.20 0.72 0.32 1022

abusive 0.90 0.77 0.83 5407

normal 0.97 0.81 0.88 10765

accuracy 0.79 17194

macro avg 0.69 0.77 0.68 17194

weighted avg 0.90 0.79 0.83 17194

average eval\_loss: 0.6627791722066874, accuracy: 0.7914970338490172

training 5 epoch...

388.4187936782837 seconds used......

5 training finished! average train loss: 0.4008529708546252

total clf loss: 83.78966799378395 total adv loss: 46.89409899304155

evaluating...

Confusion matrix, without normalization

[[ 624 236 162]

[ 726 4492 189]

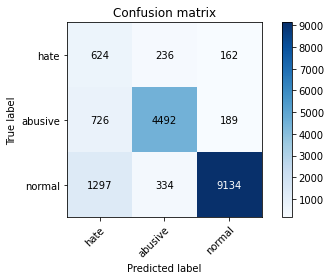
[1297 334 9134]]

Normalized confusion matrix

[[0.61056751 0.23091977 0.15851272]

[0.13427039 0.83077492 0.03495469]

[0.12048305 0.03102647 0.84849048]]





precision recall f1-score support

hate 0.24 0.61 0.34 1022

abusive 0.89 0.83 0.86 5407

normal 0.96 0.85 0.90 10765

accuracy 0.83 17194

macro avg 0.70 0.76 0.70 17194

weighted avg 0.90 0.83 0.85 17194

average eval\_loss: 0.622006008563953, accuracy: 0.8287774805164593