alpha = 2

1 run

training 1 epoch...

251.7345187664032 seconds used......

1 training finished! average train loss: 1.317704437615909

total clf loss: 273.9854424595833 total adv loss: 154.87881191819906

evaluating...

Confusion matrix, without normalization

[[ 784 159 79]

[ 950 4396 61]

[2302 389 8074]]

Normalized confusion matrix

[[0.76712329 0.1555773 0.07729941]

[0.17569817 0.81302016 0.01128167]

[0.21384115 0.03613562 0.75002322]]





precision recall f1-score support

hate 0.19 0.77 0.31 1022

abusive 0.89 0.81 0.85 5407

normal 0.98 0.75 0.85 10765

accuracy 0.77 17194

macro avg 0.69 0.78 0.67 17194

weighted avg 0.91 0.77 0.82 17194

average eval\_loss: 0.6024271737870782, accuracy: 0.7708502966150983

training 2 epoch...

251.24795866012573 seconds used......

2 training finished! average train loss: 0.9941253755283679

total clf loss: 204.65173421800137 total adv loss: 117.87290371581912

evaluating...

Confusion matrix, without normalization

[[ 733 158 131]

[ 951 4335 121]

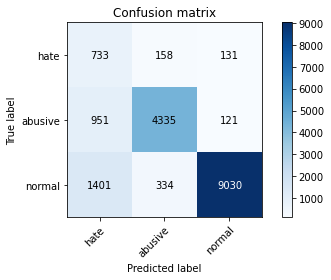
[1401 334 9030]]

Normalized confusion matrix

[[0.71722114 0.15459883 0.12818004]

[0.17588311 0.80173849 0.0223784 ]

[0.13014399 0.03102647 0.83882954]]





precision recall f1-score support

hate 0.24 0.72 0.36 1022

abusive 0.90 0.80 0.85 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: 0.4759501629357421, accuracy: 0.8199371873909503

training 3 epoch...

250.82006216049194 seconds used......

3 training finished! average train loss: 0.7733172251260577

total clf loss: 159.01547241955996 total adv loss: 91.78202951513231

evaluating...

Confusion matrix, without normalization

[[ 696 195 131]

[ 743 4534 130]

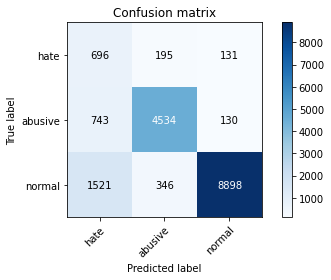
[1521 346 8898]]

Normalized confusion matrix

[[0.68101761 0.19080235 0.12818004]

[0.13741446 0.83854263 0.02404291]

[0.14129122 0.0321412 0.82656758]]





precision recall f1-score support

hate 0.24 0.68 0.35 1022

abusive 0.89 0.84 0.87 5407

normal 0.97 0.83 0.89 10765

accuracy 0.82 17194

macro avg 0.70 0.78 0.70 17194

weighted avg 0.90 0.82 0.85 17194

average eval\_loss: 0.521487490728844, accuracy: 0.8216819820867745

training 4 epoch...

250.99555087089539 seconds used......

4 training finished! average train loss: 0.5353503548971148

total clf loss: 110.13853633031249 total adv loss: 63.51083538727835

evaluating...

Confusion matrix, without normalization

[[ 713 172 137]

[ 983 4290 134]

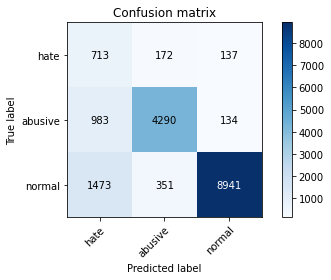
[1473 351 8941]]

Normalized confusion matrix

[[0.69765166 0.16829746 0.13405088]

[0.18180137 0.79341594 0.02478269]

[0.13683233 0.03260567 0.83056201]]





precision recall f1-score support

hate 0.22 0.70 0.34 1022

abusive 0.89 0.79 0.84 5407

normal 0.97 0.83 0.90 10765

accuracy 0.81 17194

macro avg 0.70 0.77 0.69 17194

weighted avg 0.90 0.81 0.84 17194

average eval\_loss: 0.5998458736337895, accuracy: 0.8109805746190532

training 5 epoch...

251.1270170211792 seconds used......

5 training finished! average train loss: 0.37117007254090995

total clf loss: 76.53073606826365 total adv loss: 43.94880301586818

evaluating...

Confusion matrix, without normalization

[[ 694 184 144]

[ 959 4306 142]

[1563 325 8877]]

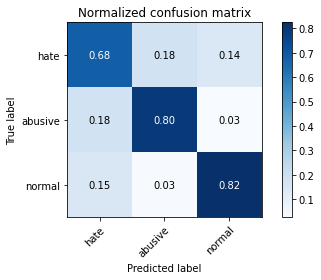
Normalized confusion matrix

[[0.67906067 0.18003914 0.1409002 ]

[0.17736268 0.79637507 0.02626225]

[0.14519275 0.03019043 0.82461681]]





precision recall f1-score support

hate 0.22 0.68 0.33 1022

abusive 0.89 0.80 0.84 5407

normal 0.97 0.82 0.89 10765

accuracy 0.81 17194

macro avg 0.69 0.77 0.69 17194

weighted avg 0.90 0.81 0.84 17194

average eval\_loss: 0.6893219620249299, accuracy: 0.8070838664650459

2 run...

training 1 epoch...

244.48889327049255 seconds used......

1 training finished! average train loss: 1.2919810208456242

total clf loss: 267.9980513751507 total adv loss: 152.17477033287287

evaluating...

Confusion matrix, without normalization

[[ 772 131 119]

[1205 4093 109]

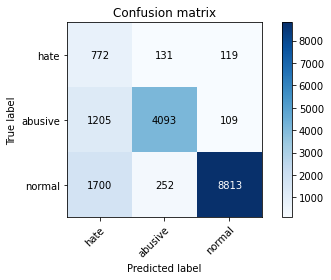
[1700 252 8813]]

Normalized confusion matrix

[[0.7553816 0.12818004 0.11643836]

[0.22285926 0.75698169 0.02015905]

[0.15791918 0.0234092 0.81867162]]





precision recall f1-score support

hate 0.21 0.76 0.33 1022

abusive 0.91 0.76 0.83 5407

normal 0.97 0.82 0.89 10765

accuracy 0.80 17194

macro avg 0.70 0.78 0.68 17194

weighted avg 0.91 0.80 0.84 17194

average eval\_loss: 0.5331492138463397, accuracy: 0.7955100616494126

training 2 epoch...

244.30796027183533 seconds used......

2 training finished! average train loss: 0.9859799715116384

total clf loss: 203.76064560562372 total adv loss: 116.51424096897244

evaluating...

Confusion matrix, without normalization

[[ 692 183 147]

[ 762 4506 139]

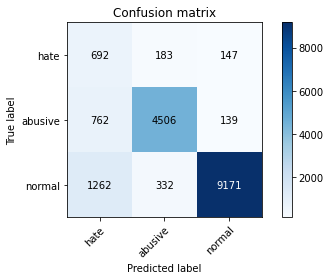
[1262 332 9171]]

Normalized confusion matrix

[[0.67710372 0.17906067 0.14383562]

[0.14092843 0.83336416 0.02570742]

[0.11723177 0.03084069 0.85192754]]





precision recall f1-score support

hate 0.25 0.68 0.37 1022

abusive 0.90 0.83 0.86 5407

normal 0.97 0.85 0.91 10765

accuracy 0.84 17194

macro avg 0.71 0.79 0.71 17194

weighted avg 0.90 0.84 0.86 17194

average eval\_loss: 0.46157180995268876, accuracy: 0.8356984994765616

training 3 epoch...

244.3346426486969 seconds used......

3 training finished! average train loss: 0.7714783096723847

total clf loss: 160.52742485702038 total adv loss: 90.61873307731003

evaluating...

Confusion matrix, without normalization

[[ 702 185 135]

[1011 4302 94]

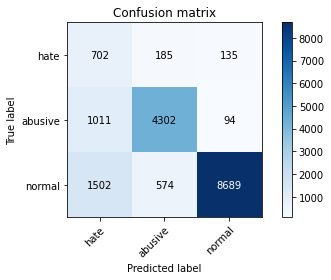
[1502 574 8689]]

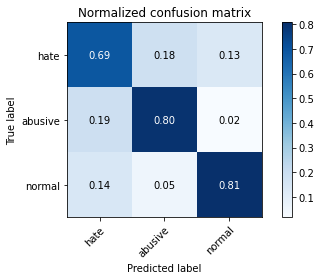
Normalized confusion matrix

[[0.68688845 0.18101761 0.13209393]

[0.18697984 0.79563529 0.01738487]

[0.13952624 0.05332095 0.80715281]]





precision recall f1-score support

hate 0.22 0.69 0.33 1022

abusive 0.85 0.80 0.82 5407

normal 0.97 0.81 0.88 10765

accuracy 0.80 17194

macro avg 0.68 0.76 0.68 17194

weighted avg 0.89 0.80 0.83 17194

average eval\_loss: 0.5547282660076784, accuracy: 0.7963824589973246

training 4 epoch...

244.50652503967285 seconds used......

4 training finished! average train loss: 0.5606817514646134

total clf loss: 117.02631874382496 total adv loss: 65.67784862802364

evaluating...

Confusion matrix, without normalization

[[ 682 166 174]

[1028 4222 157]

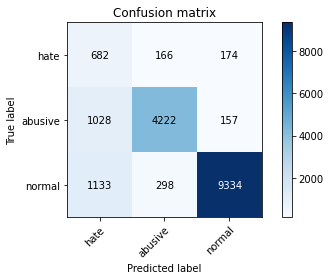
[1133 298 9334]]

Normalized confusion matrix

[[0.66731898 0.16242661 0.1702544 ]

[0.19012391 0.78083965 0.02903643]

[0.10524849 0.0276823 0.86706921]]





precision recall f1-score support

hate 0.24 0.67 0.35 1022

abusive 0.90 0.78 0.84 5407

normal 0.97 0.87 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: 0.5334538920284357, accuracy: 0.8280795626381295

training 5 epoch...

244.52348613739014 seconds used......

5 training finished! average train loss: 0.39344810657455476

total clf loss: 80.56161028984934 total adv loss: 46.867950345622376

evaluating...

Confusion matrix, without normalization

[[ 710 159 153]

[1122 4144 141]

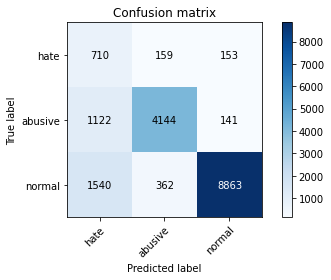
[1540 362 8863]]

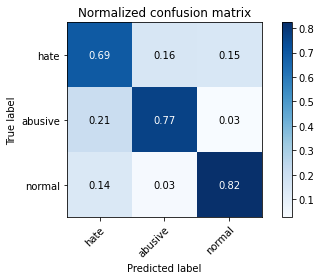
Normalized confusion matrix

[[0.69471624 0.1555773 0.14970646]

[0.20750878 0.76641391 0.02607731]

[0.1430562 0.0336275 0.8233163 ]]





precision recall f1-score support

hate 0.21 0.69 0.32 1022

abusive 0.89 0.77 0.82 5407

normal 0.97 0.82 0.89 10765

accuracy 0.80 17194

macro avg 0.69 0.76 0.68 17194

weighted avg 0.90 0.80 0.84 17194

average eval\_loss: 0.6827372788538246, accuracy: 0.797778294753984

3 run...

training 1 epoch...

244.53070497512817 seconds used......

1 training finished! average train loss: 1.3073807617746143

total clf loss: 270.7801495194435 total adv loss: 154.19476432353258

evaluating...

Confusion matrix, without normalization

[[ 660 211 151]

[ 627 4641 139]

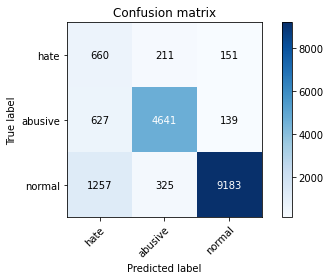
[1257 325 9183]]

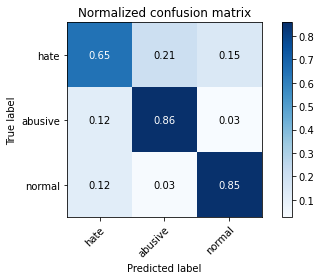
Normalized confusion matrix

[[0.64579256 0.20645793 0.14774951]

[0.11596079 0.85833179 0.02570742]

[0.1167673 0.03019043 0.85304227]]





precision recall f1-score support

hate 0.26 0.65 0.37 1022

abusive 0.90 0.86 0.88 5407

normal 0.97 0.85 0.91 10765

accuracy 0.84 17194

macro avg 0.71 0.79 0.72 17194

weighted avg 0.90 0.84 0.87 17194

average eval\_loss: 0.46023089229020964, accuracy: 0.8423868791438874

training 2 epoch...

244.54041528701782 seconds used......

2 training finished! average train loss: 0.9788175129728984

total clf loss: 202.98359286785126 total adv loss: 115.31628309562802

evaluating...

Confusion matrix, without normalization

[[ 737 165 120]

[ 988 4319 100]

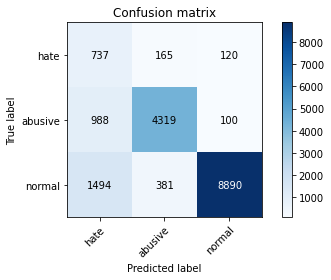
[1494 381 8890]]

Normalized confusion matrix

[[0.72113503 0.16144814 0.11741683]

[0.1827261 0.79877936 0.01849454]

[0.13878309 0.03539248 0.82582443]]





precision recall f1-score support

hate 0.23 0.72 0.35 1022

abusive 0.89 0.80 0.84 5407

normal 0.98 0.83 0.89 10765

accuracy 0.81 17194

macro avg 0.70 0.78 0.69 17194

weighted avg 0.90 0.81 0.85 17194

average eval\_loss: 0.49363109616692674, accuracy: 0.8110968942654414

training 3 epoch...

244.83341646194458 seconds used......

3 training finished! average train loss: 0.7740377126173295

total clf loss: 162.587564997375 total adv loss: 90.15557075664401

evaluating...

244.8402132987976 seconds used......

4 training finished! average train loss: 0.5537304427668149

total clf loss: 117.60736109688878 total adv loss: 63.84761265106499

evaluating...

Confusion matrix, without normalization

[[ 619 216 187]

[ 648 4600 159]

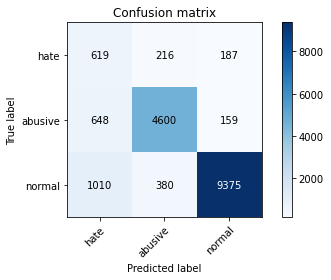
[1010 380 9375]]

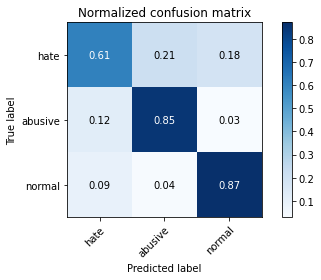
Normalized confusion matrix

[[0.60567515 0.21135029 0.18297456]

[0.11984465 0.85074903 0.02940633]

[0.09382257 0.03529958 0.87087784]]





precision recall f1-score support

hate 0.27 0.61 0.38 1022

abusive 0.89 0.85 0.87 5407

normal 0.96 0.87 0.92 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: 0.49135318499188435, accuracy: 0.8487844596952425

training 5 epoch...

244.72014164924622 seconds used......

5 training finished! average train loss: 0.4093677394478908

total clf loss: 86.16832070425153 total adv loss: 47.590793962357566

evaluating...

Confusion matrix, without normalization

[[ 701 179 142]

[1119 4102 186]

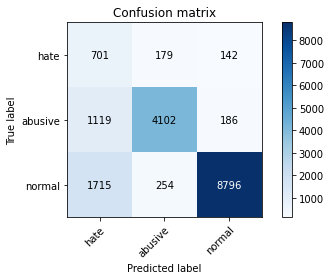
[1715 254 8796]]

Normalized confusion matrix

[[0.68590998 0.17514677 0.13894325]

[0.20695395 0.7586462 0.03439985]

[0.15931259 0.02359498 0.81709243]]





precision recall f1-score support

hate 0.20 0.69 0.31 1022

abusive 0.90 0.76 0.83 5407

normal 0.96 0.82 0.88 10765

accuracy 0.79 17194

macro avg 0.69 0.75 0.67 17194

weighted avg 0.90 0.79 0.83 17194

average eval\_loss: 0.6695523923133002, accuracy: 0.7909154356170758

4 run...

training 1 epoch...

244.65195083618164 seconds used......

1 training finished! average train loss: 1.3198100652705467

total clf loss: 273.5507931858301 total adv loss: 155.56253326684237

evaluating...

Confusion matrix, without normalization

[[ 622 232 168]

[ 571 4724 112]

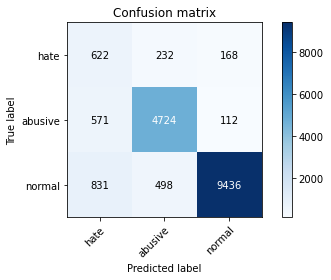
[ 831 498 9436]]

Normalized confusion matrix

[[0.60861057 0.22700587 0.16438356]

[0.10560385 0.87368226 0.02071389]

[0.07719461 0.04626103 0.87654436]]





precision recall f1-score support

hate 0.31 0.61 0.41 1022

abusive 0.87 0.87 0.87 5407

normal 0.97 0.88 0.92 10765

accuracy 0.86 17194

macro avg 0.71 0.79 0.73 17194

weighted avg 0.90 0.86 0.87 17194

average eval\_loss: 0.4286386078873346, accuracy: 0.8597185064557403

training 2 epoch...

245.09509921073914 seconds used......

2 training finished! average train loss: 0.9872107675328481

total clf loss: 204.227981492877 total adv loss: 116.5531941242516

evaluating...

Confusion matrix, without normalization

[[ 685 173 164]

[ 897 4368 142]

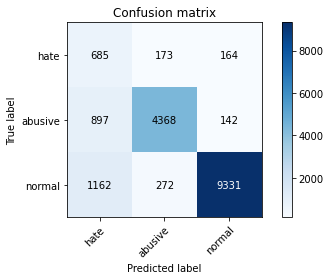
[1162 272 9331]]

Normalized confusion matrix

[[0.6702544 0.16927593 0.16046967]

[0.16589606 0.80784169 0.02626225]

[0.10794241 0.02526707 0.86679052]]





precision recall f1-score support

hate 0.25 0.67 0.36 1022

abusive 0.91 0.81 0.85 5407

normal 0.97 0.87 0.91 10765

accuracy 0.84 17194

macro avg 0.71 0.78 0.71 17194

weighted avg 0.91 0.84 0.86 17194

average eval\_loss: 0.4607370041484056, accuracy: 0.8365708968244736

training 3 epoch...

244.5757713317871 seconds used......

3 training finished! average train loss: 0.766220608965807

total clf loss: 158.26400282979012 total adv loss: 90.58586357161403

evaluating...

Confusion matrix, without normalization

[[ 716 192 114]

[ 839 4471 97]

[1843 358 8564]]

Normalized confusion matrix

[[0.70058708 0.18786693 0.11154599]

[0.15516923 0.82689107 0.01793971]

[0.17120297 0.03325592 0.79554111]]

244.67862057685852 seconds used......

4 training finished! average train loss: 0.5368703755447073

total clf loss: 112.53894441016018 total adv loss: 62.64731593965553

evaluating...

Confusion matrix, without normalization

[[ 684 198 140]

[ 846 4424 137]

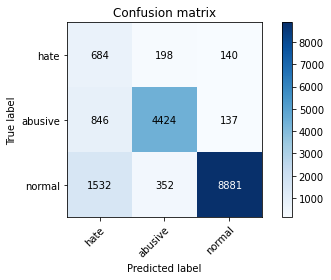
[1532 352 8881]]

Normalized confusion matrix

[[0.66927593 0.19373777 0.1369863 ]

[0.15646384 0.81819863 0.02533753]

[0.14231305 0.03269856 0.82498839]]





precision recall f1-score support

hate 0.22 0.67 0.33 1022

abusive 0.89 0.82 0.85 5407

normal 0.97 0.82 0.89 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: 0.5737404373691006, accuracy: 0.8135977666627894

training 5 epoch...

244.94830131530762 seconds used......

5 training finished! average train loss: 0.37549365914803057

total clf loss: 77.8179714679718 total adv loss: 44.26286010222975

evaluating...

Confusion matrix, without normalization

[[ 623 251 148]

[ 660 4638 109]

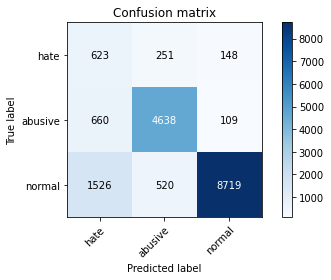
[1526 520 8719]]

Normalized confusion matrix

[[0.60958904 0.24559687 0.14481409]

[0.12206399 0.85777696 0.02015905]

[0.14175569 0.04830469 0.80993962]]





precision recall f1-score support

hate 0.22 0.61 0.33 1022

abusive 0.86 0.86 0.86 5407

normal 0.97 0.81 0.88 10765

accuracy 0.81 17194

macro avg 0.68 0.76 0.69 17194

weighted avg 0.89 0.81 0.84 17194

average eval\_loss: 0.6730797988167683, accuracy: 0.8130743282540421

5 run...

training 1 epoch...

5 run...

training 1 epoch...

247.12394332885742 seconds used......

1 training finished! average train loss: 1.29930539163456

total clf loss: 269.65899254381657 total adv loss: 152.96664863824844

evaluating...

Confusion matrix, without normalization

[[ 597 228 197]

[ 573 4705 129]

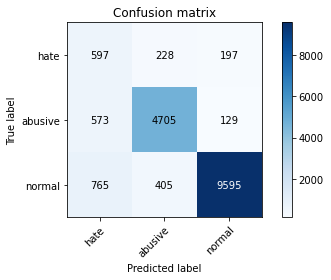
[ 765 405 9595]]

Normalized confusion matrix

[[0.58414873 0.22309198 0.1927593 ]

[0.10597374 0.8701683 0.02385796]

[0.07106363 0.03762192 0.89131444]]





precision recall f1-score support

hate 0.31 0.58 0.40 1022

abusive 0.88 0.87 0.88 5407

normal 0.97 0.89 0.93 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: 0.4199236385437638, accuracy: 0.8664068861230662

training 2 epoch...

246.34732151031494 seconds used......

2 training finished! average train loss: 0.9849004506661176

total clf loss: 204.27564583718777 total adv loss: 116.01762676052749

evaluating...

Confusion matrix, without normalization

[[ 625 258 139]

[ 478 4793 136]

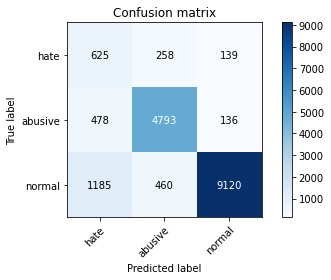
[1185 460 9120]]

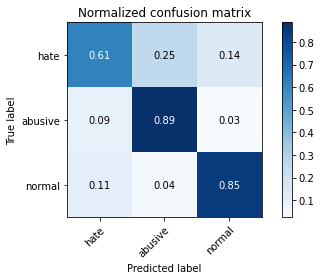
Normalized confusion matrix

[[0.61154599 0.25244618 0.13600783]

[0.08840392 0.8864435 0.02515258]

[0.11007896 0.04273107 0.84718997]]





precision recall f1-score support

hate 0.27 0.61 0.38 1022

abusive 0.87 0.89 0.88 5407

normal 0.97 0.85 0.90 10765

accuracy 0.85 17194

macro avg 0.70 0.78 0.72 17194

weighted avg 0.90 0.85 0.87 17194

average eval\_loss: 0.44011212375101655, accuracy: 0.8455275095963708

training 3 epoch...

246.331218957901 seconds used......

3 training finished! average train loss: 0.7767704998397397

total clf loss: 160.44396313279867 total adv loss: 91.83268428687006

evaluating...

Confusion matrix, without normalization

[[ 742 169 111]

[1022 4261 124]

[1767 307 8691]]

Normalized confusion matrix

[[0.7260274 0.16536204 0.10861057]

[0.18901424 0.78805252 0.02293323]

[0.16414306 0.02851835 0.8073386 ]]





precision recall f1-score support

hate 0.21 0.73 0.33 1022

abusive 0.90 0.79 0.84 5407

normal 0.97 0.81 0.88 10765

accuracy 0.80 17194

macro avg 0.69 0.77 0.68 17194

weighted avg 0.90 0.80 0.84 17194

average eval\_loss: 0.5734861059071019, accuracy: 0.7964406188205188

training 4 epoch...

246.53232884407043 seconds used......

4 training finished! average train loss: 0.5608405026840572

total clf loss: 117.04953736439347 total adv loss: 65.701402655337

evaluating...

Confusion matrix, without normalization

[[ 672 204 146]

[ 787 4490 130]

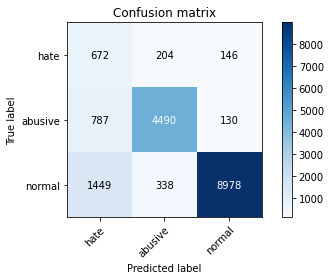
[1449 338 8978]]

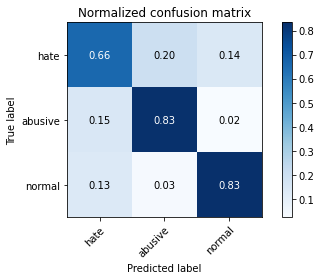
Normalized confusion matrix

[[0.65753425 0.19960861 0.14285714]

[0.14555206 0.83040503 0.02404291]

[0.13460288 0.03139805 0.83399907]]





precision recall f1-score support

hate 0.23 0.66 0.34 1022

abusive 0.89 0.83 0.86 5407

normal 0.97 0.83 0.90 10765

accuracy 0.82 17194

macro avg 0.70 0.77 0.70 17194

weighted avg 0.90 0.82 0.85 17194

average eval\_loss: 0.5782913272377364, accuracy: 0.8223798999651041

training 5 epoch...

246.83566188812256 seconds used......

5 training finished! average train loss: 0.3940809165163315

total clf loss: 81.18620286881924 total adv loss: 46.695821779896505

evaluating...

Confusion matrix, without normalization

[[ 657 237 128]

[ 716 4561 130]

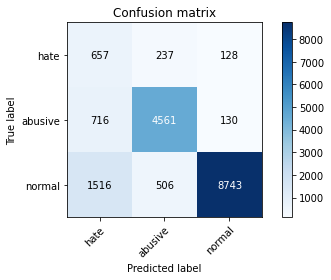
[1516 506 8743]]

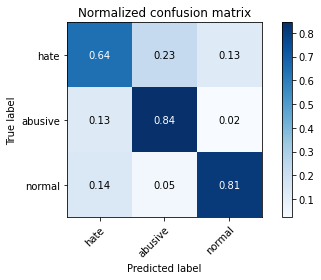
Normalized confusion matrix

[[0.64285714 0.23189824 0.12524462]

[0.13242094 0.84353616 0.02404291]

[0.14082675 0.04700418 0.81216907]]





precision recall f1-score support

hate 0.23 0.64 0.34 1022

abusive 0.86 0.84 0.85 5407

normal 0.97 0.81 0.88 10765

accuracy 0.81 17194

macro avg 0.69 0.77 0.69 17194

weighted avg 0.89 0.81 0.84 17194

average eval\_loss: 0.6614437442186267, accuracy: 0.8119692916133535