weights = [1.0, 0.5, 1.0]

1 run

training 1 epoch...

238.1450777053833 seconds used......

1 training finished! average train loss: 0.5984911773664957

evaluating...

Confusion matrix, without normalization

[[ 704 113 205]

[1210 4002 195]

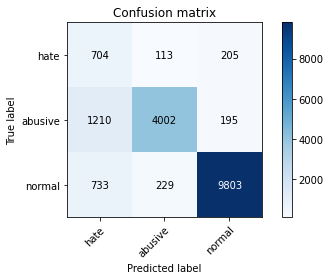
[ 733 229 9803]]

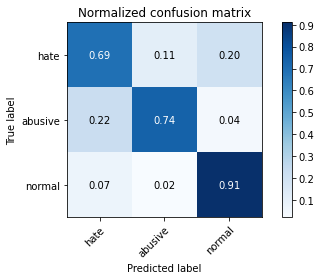
Normalized confusion matrix

[[0.6888454 0.11056751 0.20058708]

[0.22378398 0.74015166 0.03606436]

[0.06809104 0.02127264 0.91063632]]





precision recall f1-score support

hate 0.27 0.69 0.38 1022

abusive 0.92 0.74 0.82 5407

normal 0.96 0.91 0.94 10765

accuracy 0.84 17194

macro avg 0.72 0.78 0.71 17194

weighted avg 0.91 0.84 0.87 17194

average eval\_loss: 0.4253830050139926, accuracy: 0.8438408747237408

training 2 epoch...

237.59506607055664 seconds used......

2 training finished! average train loss: 0.4365563109752556

evaluating...

Confusion matrix, without normalization

[[ 679 192 151]

[ 841 4438 128]

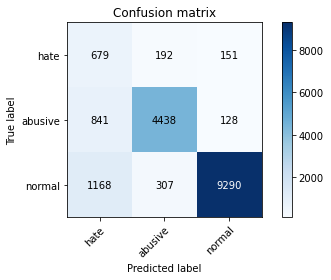
[1168 307 9290]]

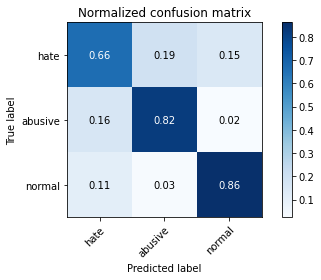
Normalized confusion matrix

[[0.66438356 0.18786693 0.14774951]

[0.15553912 0.82078787 0.02367302]

[0.10849977 0.02851835 0.86298189]]





precision recall f1-score support

hate 0.25 0.66 0.37 1022

abusive 0.90 0.82 0.86 5407

normal 0.97 0.86 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: 0.44026333532880907, accuracy: 0.8379085727579388

training 3 epoch...

237.61918592453003 seconds used......

3 training finished! average train loss: 0.3199227697323192

evaluating...

Confusion matrix, without normalization

[[ 619 197 206]

[ 751 4388 268]

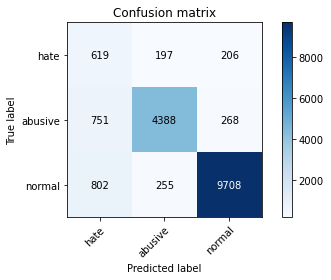
[ 802 255 9708]]

Normalized confusion matrix

[[0.60567515 0.1927593 0.20156556]

[0.13889403 0.8115406 0.04956538]

[0.0745007 0.02368788 0.90181143]]





precision recall f1-score support

hate 0.28 0.61 0.39 1022

abusive 0.91 0.81 0.86 5407

normal 0.95 0.90 0.93 10765

accuracy 0.86 17194

macro avg 0.72 0.77 0.72 17194

weighted avg 0.90 0.86 0.87 17194

average eval\_loss: 0.39621463256649847, accuracy: 0.8558217983017332

training 4 epoch...

237.54479885101318 seconds used......

4 training finished! average train loss: 0.20897861835851625

evaluating...

Confusion matrix, without normalization

[[ 669 177 176]

[ 947 4254 206]

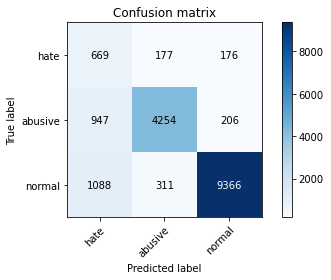
[1088 311 9366]]

Normalized confusion matrix

[[0.65459883 0.17318982 0.17221135]

[0.17514333 0.78675791 0.03809876]

[0.10106828 0.02888992 0.8700418 ]]





precision recall f1-score support

hate 0.25 0.65 0.36 1022

abusive 0.90 0.79 0.84 5407

normal 0.96 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: 0.5930314701732768, accuracy: 0.8310457136210306

training 5 epoch...

237.40467715263367 seconds used......

5 training finished! average train loss: 0.14888120259202461

evaluating...

Confusion matrix, without normalization

[[ 648 200 174]

[ 894 4316 197]

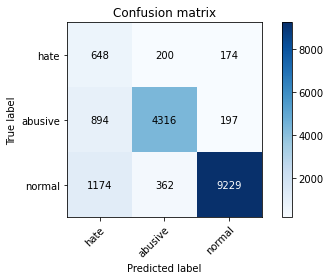
[1174 362 9229]]

Normalized confusion matrix

[[0.63405088 0.19569472 0.1702544 ]

[0.16534122 0.79822452 0.03643425]

[0.10905713 0.0336275 0.85731537]]





precision recall f1-score support

hate 0.24 0.63 0.35 1022

abusive 0.88 0.80 0.84 5407

normal 0.96 0.86 0.91 10765

accuracy 0.83 17194

macro avg 0.69 0.76 0.70 17194

weighted avg 0.89 0.83 0.85 17194

average eval\_loss: 0.6364033984897528, accuracy: 0.8254623705943934

2 run...

training 1 epoch...

237.3417558670044 seconds used......

1 training finished! average train loss: 0.6176036211155338

evaluating...

Confusion matrix, without normalization

[[ 562 154 306]

[ 834 4255 318]

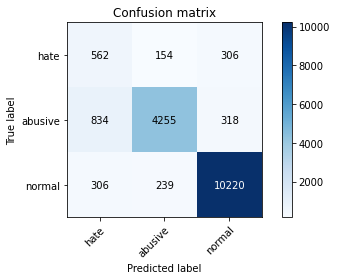
[ 306 239 10220]]

Normalized confusion matrix

[[0.54990215 0.15068493 0.29941292]

[0.1542445 0.78694285 0.05881265]

[0.02842545 0.02220158 0.94937297]]





precision recall f1-score support

hate 0.33 0.55 0.41 1022

abusive 0.92 0.79 0.85 5407

normal 0.94 0.95 0.95 10765

accuracy 0.87 17194

macro avg 0.73 0.76 0.73 17194

weighted avg 0.90 0.87 0.88 17194

average eval\_loss: 0.3205982637093511, accuracy: 0.8745492613702455

training 2 epoch...

237.39474296569824 seconds used......

2 training finished! average train loss: 0.4499659133784927

evaluating...

Confusion matrix, without normalization

[[ 616 163 243]

[ 823 4306 278]

[ 515 239 10011]]

Normalized confusion matrix

[[0.60273973 0.15949119 0.23776908]

[0.1522101 0.79637507 0.05141483]

[0.04784022 0.02220158 0.9299582 ]]





precision recall f1-score support

hate 0.32 0.60 0.41 1022

abusive 0.91 0.80 0.85 5407

normal 0.95 0.93 0.94 10765

accuracy 0.87 17194

macro avg 0.73 0.78 0.74 17194

weighted avg 0.90 0.87 0.88 17194

average eval\_loss: 0.3480990322003531, accuracy: 0.8685006397580551

training 3 epoch...

237.61274695396423 seconds used......

3 training finished! average train loss: 0.33428343650348835

evaluating...

Confusion matrix, without normalization

[[ 745 125 152]

[1220 4023 164]

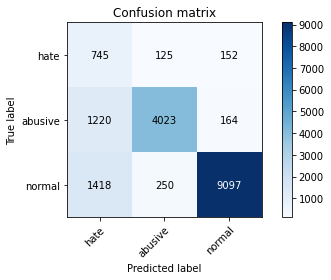
[1418 250 9097]]

Normalized confusion matrix

[[0.72896282 0.1223092 0.14872798]

[0.22563344 0.74403551 0.03033105]

[0.13172318 0.02322341 0.84505341]]





precision recall f1-score support

hate 0.22 0.73 0.34 1022

abusive 0.91 0.74 0.82 5407

normal 0.97 0.85 0.90 10765

accuracy 0.81 17194

macro avg 0.70 0.77 0.69 17194

weighted avg 0.91 0.81 0.84 17194

average eval\_loss: 0.5241760392521703, accuracy: 0.8063859485867163

training 4 epoch...

237.58592295646667 seconds used......

4 training finished! average train loss: 0.21262837164776724

evaluating...

Confusion matrix, without normalization

[[ 644 218 160]

[ 752 4485 170]

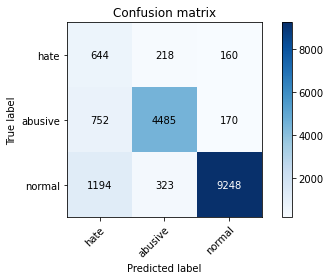
[1194 323 9248]]

Normalized confusion matrix

[[0.63013699 0.21330724 0.15655577]

[0.13907897 0.8294803 0.03144072]

[0.110915 0.03000464 0.85908035]]





precision recall f1-score support

hate 0.25 0.63 0.36 1022

abusive 0.89 0.83 0.86 5407

normal 0.97 0.86 0.91 10765

accuracy 0.84 17194

macro avg 0.70 0.77 0.71 17194

weighted avg 0.90 0.84 0.86 17194

average eval\_loss: 0.5543177532443647, accuracy: 0.8361637780621147

training 5 epoch...

237.64038634300232 seconds used......

5 training finished! average train loss: 0.14547992628341033

evaluating...

Confusion matrix, without normalization

[[ 674 195 153]

[ 964 4281 162]

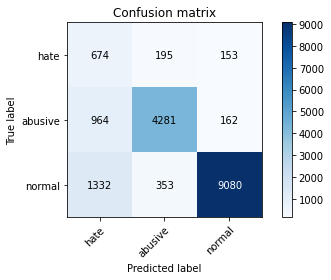
[1332 353 9080]]

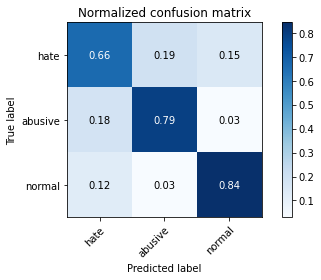
Normalized confusion matrix

[[0.65949119 0.19080235 0.14970646]

[0.17828741 0.79175143 0.02996116]

[0.12373432 0.03279145 0.84347422]]





precision recall f1-score support

hate 0.23 0.66 0.34 1022

abusive 0.89 0.79 0.84 5407

normal 0.97 0.84 0.90 10765

accuracy 0.82 17194

macro avg 0.69 0.76 0.69 17194

weighted avg 0.90 0.82 0.85 17194

average eval\_loss: 0.7091320495900892, accuracy: 0.8162731185297196

3 run...

training 1 epoch...

237.39137649536133 seconds used......

1 training finished! average train loss: 0.6079132820628835

evaluating...

Confusion matrix, without normalization

[[ 747 154 121]

[ 986 4290 131]

[1590 282 8893]]

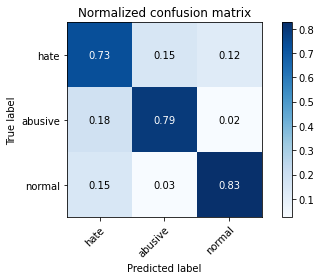
Normalized confusion matrix

[[0.73091977 0.15068493 0.1183953 ]

[0.1823562 0.79341594 0.02422785]

[0.14770088 0.02619601 0.82610311]]





precision recall f1-score support

hate 0.22 0.73 0.34 1022

abusive 0.91 0.79 0.85 5407

normal 0.97 0.83 0.89 10765

accuracy 0.81 17194

macro avg 0.70 0.78 0.69 17194

weighted avg 0.91 0.81 0.85 17194

average eval\_loss: 0.5015061507044837, accuracy: 0.8101663370943353

training 2 epoch...

237.579909324646 seconds used......

2 training finished! average train loss: 0.43898275921764546

evaluating...

Confusion matrix, without normalization

[[ 746 91 185]

[1405 3782 220]

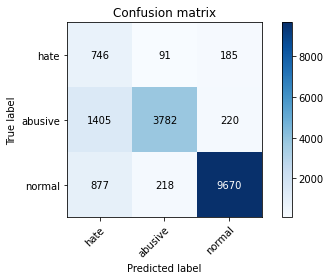
[ 877 218 9670]]

Normalized confusion matrix

[[0.72994129 0.0890411 0.18101761]

[0.25984834 0.69946366 0.040688 ]

[0.08146772 0.02025081 0.89828147]]





precision recall f1-score support

hate 0.25 0.73 0.37 1022

abusive 0.92 0.70 0.80 5407

normal 0.96 0.90 0.93 10765

accuracy 0.83 17194

macro avg 0.71 0.78 0.70 17194

weighted avg 0.91 0.83 0.85 17194

average eval\_loss: 0.42607861122679574, accuracy: 0.825753169710364

training 3 epoch...

237.46844172477722 seconds used......

3 training finished! average train loss: 0.31677164444300443

evaluating...

Confusion matrix, without normalization

[[ 759 104 159]

[1418 3843 146]

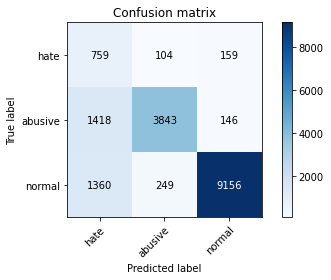
[1360 249 9156]]

Normalized confusion matrix

[[0.74266145 0.10176125 0.1555773 ]

[0.26225264 0.71074533 0.02700203]

[0.12633535 0.02313052 0.85053414]]





precision recall f1-score support

hate 0.21 0.74 0.33 1022

abusive 0.92 0.71 0.80 5407

normal 0.97 0.85 0.91 10765

accuracy 0.80 17194

macro avg 0.70 0.77 0.68 17194

weighted avg 0.91 0.80 0.84 17194

average eval\_loss: 0.5362758970797755, accuracy: 0.8001628475049436

training 4 epoch...

237.6150176525116 seconds used......

4 training finished! average train loss: 0.21143787725918986

evaluating...

Confusion matrix, without normalization

[[ 675 167 180]

[1032 4221 154]

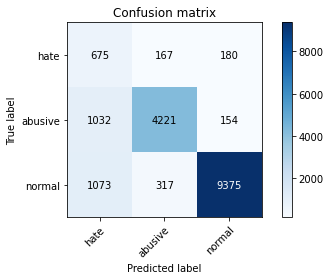
[1073 317 9375]]

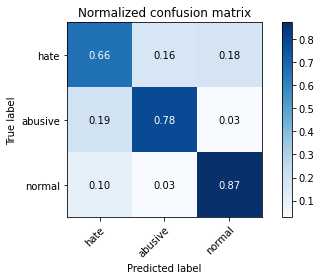
Normalized confusion matrix

[[0.66046967 0.16340509 0.17612524]

[0.1908637 0.78065471 0.0284816 ]

[0.09967487 0.02944728 0.87087784]]





precision recall f1-score support

hate 0.24 0.66 0.36 1022

abusive 0.90 0.78 0.83 5407

normal 0.97 0.87 0.92 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: 0.5563784475319175, accuracy: 0.8299988368035361

training 5 epoch...

237.58400225639343 seconds used......

5 training finished! average train loss: 0.1417068738624891

evaluating...

Confusion matrix, without normalization

[[ 699 170 153]

[1154 4111 142]

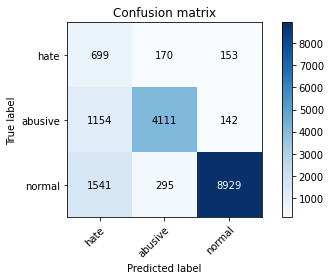
[1541 295 8929]]

Normalized confusion matrix

[[0.68395303 0.16634051 0.14970646]

[0.21342704 0.76031071 0.02626225]

[0.14314909 0.02740362 0.82944728]]





precision recall f1-score support

hate 0.21 0.68 0.32 1022

abusive 0.90 0.76 0.82 5407

normal 0.97 0.83 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: 0.7685274884234681, accuracy: 0.799057810864255

4 run...

training 1 epoch...

237.58600044250488 seconds used......

1 training finished! average train loss: 0.6094984826263941

evaluating...

Confusion matrix, without normalization

[[ 688 96 238]

[1508 3685 214]

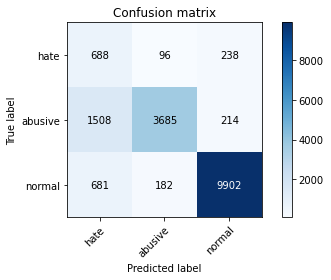
[ 681 182 9902]]

Normalized confusion matrix

[[0.67318982 0.09393346 0.23287671]

[0.27889773 0.68152395 0.03957832]

[0.06326057 0.01690664 0.91983279]]





precision recall f1-score support

hate 0.24 0.67 0.35 1022

abusive 0.93 0.68 0.79 5407

normal 0.96 0.92 0.94 10765

accuracy 0.83 17194

macro avg 0.71 0.76 0.69 17194

weighted avg 0.91 0.83 0.86 17194

average eval\_loss: 0.3898277277083591, accuracy: 0.8302314760963126

training 2 epoch...

237.91907000541687 seconds used......

2 training finished! average train loss: 0.4379085454560026

evaluating...

Confusion matrix, without normalization

[[ 710 136 176]

[1118 4130 159]

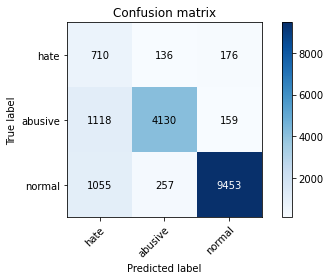
[1055 257 9453]]

Normalized confusion matrix

[[0.69471624 0.13307241 0.17221135]

[0.206769 0.76382467 0.02940633]

[0.09800279 0.02387366 0.87812355]]





precision recall f1-score support

hate 0.25 0.69 0.36 1022

abusive 0.91 0.76 0.83 5407

normal 0.97 0.88 0.92 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: 0.43303454024500626, accuracy: 0.8312783529138071

training 3 epoch...

237.79856777191162 seconds used......

3 training finished! average train loss: 0.32017373627484785

evaluating...

Confusion matrix, without normalization

[[ 711 121 190]

[1257 3932 218]

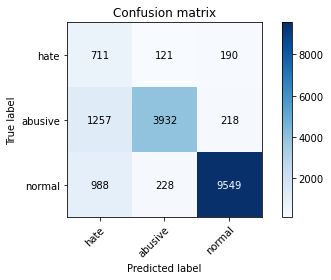
[ 988 228 9549]]

Normalized confusion matrix

[[0.69569472 0.1183953 0.18590998]

[0.23247642 0.72720547 0.04031811]

[0.09177891 0.02117975 0.88704134]]





precision recall f1-score support

hate 0.24 0.70 0.36 1022

abusive 0.92 0.73 0.81 5407

normal 0.96 0.89 0.92 10765

accuracy 0.83 17194

macro avg 0.71 0.77 0.70 17194

weighted avg 0.90 0.83 0.85 17194

average eval\_loss: 0.4565731484139728, accuracy: 0.8254042107711993

training 4 epoch...

4 run...

Downloading: 100%

411/411 [00:00<00:00, 7.86kB/s]

Downloading: 100%

263M/263M [00:08<00:00, 33.1MB/s]

training 1 epoch...

238.03816843032837 seconds used......

1 training finished! average train loss: 0.6059396283082984

evaluating...

Confusion matrix, without normalization

[[ 806 109 107]

[1318 3973 116]

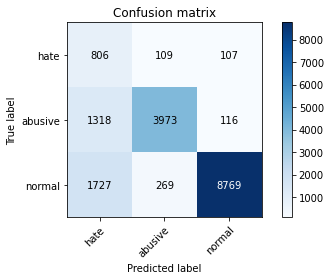
[1727 269 8769]]

Normalized confusion matrix

[[0.78864971 0.10665362 0.10469667]

[0.24375809 0.73478824 0.02145367]

[0.16042731 0.02498839 0.8145843 ]]





precision recall f1-score support

hate 0.21 0.79 0.33 1022

abusive 0.91 0.73 0.81 5407

normal 0.98 0.81 0.89 10765

accuracy 0.79 17194

macro avg 0.70 0.78 0.68 17194

weighted avg 0.91 0.79 0.83 17194

average eval\_loss: 0.5211548132432061, accuracy: 0.7879492846341747

training 2 epoch...

237.41164994239807 seconds used......

2 training finished! average train loss: 0.4389080432481206

evaluating...

Confusion matrix, without normalization

[[ 782 105 135]

[1367 3865 175]

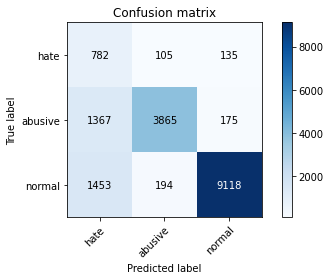
[1453 194 9118]]

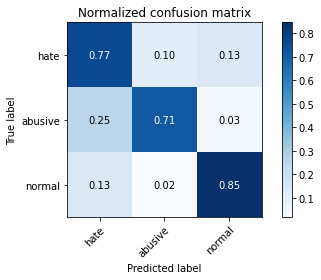
Normalized confusion matrix

[[0.76516634 0.10273973 0.13209393]

[0.25282042 0.71481413 0.03236545]

[0.13497445 0.01802137 0.84700418]]





precision recall f1-score support

hate 0.22 0.77 0.34 1022

abusive 0.93 0.71 0.81 5407

normal 0.97 0.85 0.90 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: 0.47071282004201137, accuracy: 0.8005699662673026

training 3 epoch...

237.26673340797424 seconds used......

3 training finished! average train loss: 0.31885421208818665

evaluating...

Confusion matrix, without normalization

[[ 623 256 143]

[ 577 4721 109]

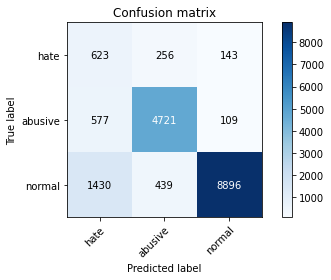
[1430 439 8896]]

Normalized confusion matrix

[[0.60958904 0.25048924 0.13992172]

[0.10671352 0.87312743 0.02015905]

[0.1328379 0.04078031 0.82638179]]





precision recall f1-score support

hate 0.24 0.61 0.34 1022

abusive 0.87 0.87 0.87 5407

normal 0.97 0.83 0.89 10765

accuracy 0.83 17194

macro avg 0.69 0.77 0.70 17194

weighted avg 0.90 0.83 0.85 17194

average eval\_loss: 0.5420009925528321, accuracy: 0.8281958822845179

training 4 epoch...

237.24717450141907 seconds used......

4 training finished! average train loss: 0.2155814655545603

evaluating...

Confusion matrix, without normalization

[[ 801 109 112]

[1649 3640 118]

[2317 222 8226]]

Normalized confusion matrix

[[0.78375734 0.10665362 0.10958904]

[0.30497503 0.67320141 0.02182356]

[0.21523456 0.02062239 0.76414306]]





precision recall f1-score support

hate 0.17 0.78 0.28 1022

abusive 0.92 0.67 0.78 5407

normal 0.97 0.76 0.86 10765

accuracy 0.74 17194

macro avg 0.69 0.74 0.64 17194

weighted avg 0.91 0.74 0.80 17194

average eval\_loss: 0.8470681400753037, accuracy: 0.7367104804001395

training 5 epoch...

237.2761950492859 seconds used......

5 training finished! average train loss: 0.14056257434550917

evaluating...

Confusion matrix, without normalization

[[ 725 138 159]

[1268 3937 202]

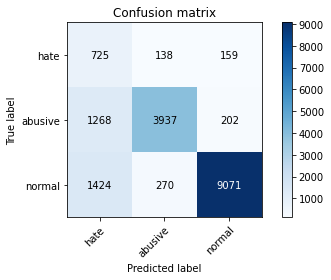
[1424 270 9071]]

Normalized confusion matrix

[[0.70939335 0.13502935 0.1555773 ]

[0.23451082 0.7281302 0.03735898]

[0.13228054 0.02508128 0.84263818]]





precision recall f1-score support

hate 0.21 0.71 0.33 1022

abusive 0.91 0.73 0.81 5407

normal 0.96 0.84 0.90 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: 0.7947734420101137, accuracy: 0.7987088519250901

5 run...

training 1 epoch...

237.44762778282166 seconds used......

1 training finished! average train loss: 0.6114453173854281

evaluating...

Confusion matrix, without normalization

[[ 669 162 191]

[ 891 4328 188]

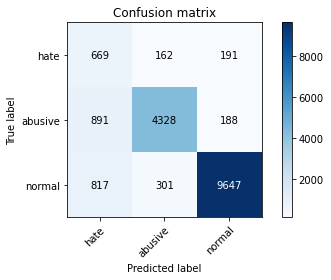
[ 817 301 9647]]

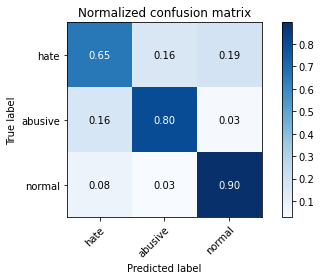
Normalized confusion matrix

[[0.65459883 0.15851272 0.18688845]

[0.16478639 0.80044387 0.03476974]

[0.0758941 0.02796098 0.89614491]]





precision recall f1-score support

hate 0.28 0.65 0.39 1022

abusive 0.90 0.80 0.85 5407

normal 0.96 0.90 0.93 10765

accuracy 0.85 17194

macro avg 0.72 0.78 0.72 17194

weighted avg 0.90 0.85 0.87 17194

average eval\_loss: 0.4060275583679593, accuracy: 0.8516924508549494

training 2 epoch...

237.43760538101196 seconds used......

2 training finished! average train loss: 0.4418298236807098

evaluating...

Confusion matrix, without normalization

[[ 668 134 220]

[1060 4110 237]

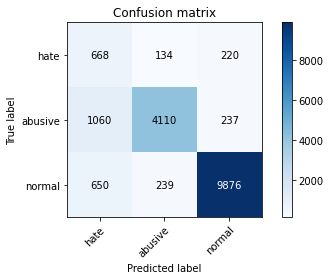
[ 650 239 9876]]

Normalized confusion matrix

[[0.65362035 0.13111546 0.21526419]

[0.19604217 0.76012576 0.04383207]

[0.06038086 0.02220158 0.91741756]]





precision recall f1-score support

hate 0.28 0.65 0.39 1022

abusive 0.92 0.76 0.83 5407

normal 0.96 0.92 0.94 10765

accuracy 0.85 17194

macro avg 0.72 0.78 0.72 17194

weighted avg 0.90 0.85 0.87 17194

average eval\_loss: 0.3684399254013633, accuracy: 0.8522740490868908

training 3 epoch...

237.38072276115417 seconds used......

3 training finished! average train loss: 0.3269817762278676

evaluating...

Confusion matrix, without normalization

[[ 776 102 144]

[1485 3763 159]

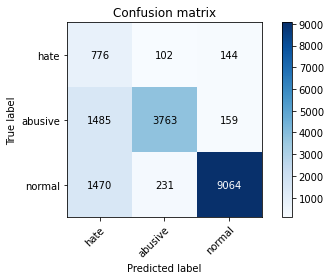
[1470 231 9064]]

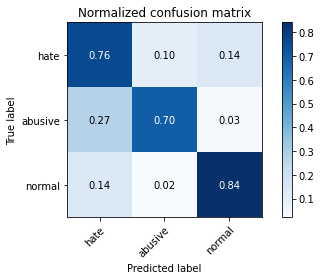
Normalized confusion matrix

[[0.7592955 0.09980431 0.1409002 ]

[0.27464398 0.69594969 0.02940633]

[0.13655365 0.02145843 0.84198792]]





precision recall f1-score support

hate 0.21 0.76 0.33 1022

abusive 0.92 0.70 0.79 5407

normal 0.97 0.84 0.90 10765

accuracy 0.79 17194

macro avg 0.70 0.77 0.67 17194

weighted avg 0.91 0.79 0.83 17194

average eval\_loss: 0.5704801017158599, accuracy: 0.7911480749098523

training 4 epoch...

237.42322444915771 seconds used......

4 training finished! average train loss: 0.21762936328102583

evaluating...

Confusion matrix, without normalization

[[ 726 133 163]

[1205 4002 200]

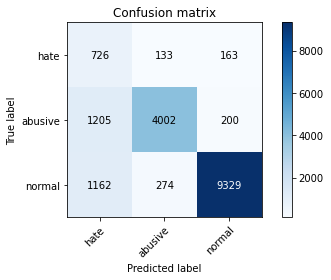
[1162 274 9329]]

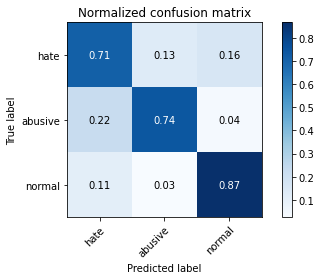
Normalized confusion matrix

[[0.71037182 0.13013699 0.15949119]

[0.22285926 0.74015166 0.03698909]

[0.10794241 0.02545286 0.86660474]]





precision recall f1-score support

hate 0.23 0.71 0.35 1022

abusive 0.91 0.74 0.82 5407

normal 0.96 0.87 0.91 10765

accuracy 0.82 17194

macro avg 0.70 0.77 0.69 17194

weighted avg 0.90 0.82 0.85 17194

average eval\_loss: 0.5986525290688977, accuracy: 0.8175526346399907

training 5 epoch...