ALPHA=2

1 run...

training 1 epoch...

393.4865062236786 seconds used......

1 training finished! average train loss: 1.3068338410579863

total clf loss: 273.67567414045334 total adv loss: 152.62585892155766

evaluating...

Confusion matrix, without normalization

[[ 665 169 188]

[ 844 4393 170]

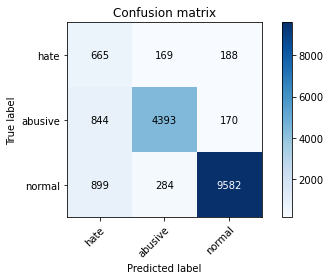
[ 899 284 9582]]

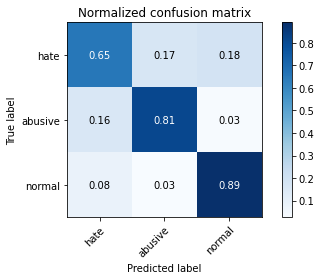
Normalized confusion matrix

[[0.65068493 0.16536204 0.18395303]

[0.15609395 0.81246532 0.03144072]

[0.08351138 0.02638179 0.89010683]]





precision recall f1-score support

hate 0.28 0.65 0.39 1022

abusive 0.91 0.81 0.86 5407

normal 0.96 0.89 0.93 10765

accuracy 0.85 17194

macro avg 0.72 0.78 0.72 17194

weighted avg 0.91 0.85 0.87 17194

average eval\_loss: 0.42060575732138267, accuracy: 0.8514598115621729

training 2 epoch...

393.05399346351624 seconds used......

2 training finished! average train loss: 0.9906326872769651

total clf loss: 205.78111590445042 total adv loss: 116.53458232060075

evaluating...

Confusion matrix, without normalization

[[ 685 202 135]

[ 693 4594 120]

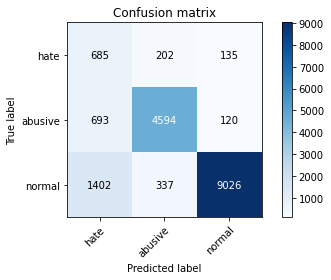
[1402 337 9026]]

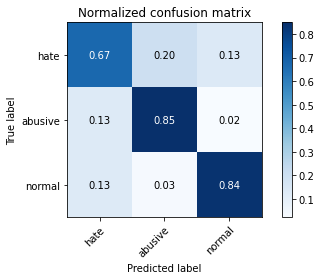
Normalized confusion matrix

[[0.6702544 0.19765166 0.13209393]

[0.12816719 0.84963936 0.02219345]

[0.13023688 0.03130516 0.83845797]]





precision recall f1-score support

hate 0.25 0.67 0.36 1022

abusive 0.89 0.85 0.87 5407

normal 0.97 0.84 0.90 10765

accuracy 0.83 17194

macro avg 0.70 0.79 0.71 17194

weighted avg 0.90 0.83 0.86 17194

average eval\_loss: 0.46974127880016037, accuracy: 0.8319762707921368

training 3 epoch...

393.1046304702759 seconds used......

3 training finished! average train loss: 0.7909740462192832

total clf loss: 164.1182005777955 total adv loss: 93.14165138918906

evaluating...

Confusion matrix, without normalization

[[ 645 242 135]

[ 615 4680 112]

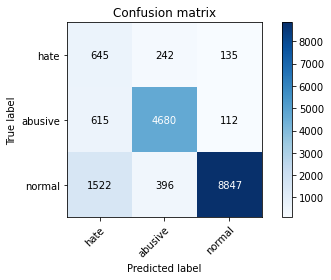
[1522 396 8847]]

Normalized confusion matrix

[[0.63111546 0.23679061 0.13209393]

[0.11374145 0.86554466 0.02071389]

[0.14138412 0.03678588 0.82183 ]]





precision recall f1-score support

hate 0.23 0.63 0.34 1022

abusive 0.88 0.87 0.87 5407

normal 0.97 0.82 0.89 10765

accuracy 0.82 17194

macro avg 0.69 0.77 0.70 17194

weighted avg 0.90 0.82 0.85 17194

average eval\_loss: 0.49355458780083544, accuracy: 0.8242410143073166

training 4 epoch...

393.34812903404236 seconds used......

4 training finished! average train loss: 0.5603131220316511

total clf loss: 116.95284532383084 total adv loss: 65.63293410930783

evaluating...

Confusion matrix, without normalization

[[ 612 212 198]

[ 758 4459 190]

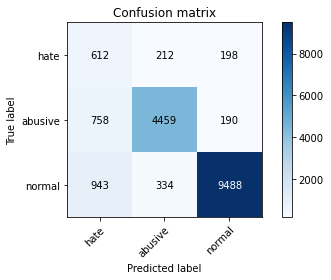
[ 943 334 9488]]

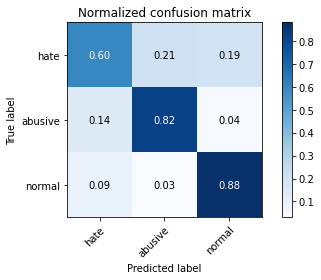
Normalized confusion matrix

[[0.59882583 0.2074364 0.19373777]

[0.14018864 0.82467172 0.03513963]

[0.0875987 0.03102647 0.88137483]]





precision recall f1-score support

hate 0.26 0.60 0.37 1022

abusive 0.89 0.82 0.86 5407

normal 0.96 0.88 0.92 10765

accuracy 0.85 17194

macro avg 0.71 0.77 0.71 17194

weighted avg 0.90 0.85 0.87 17194

average eval\_loss: 0.5133884077783414, accuracy: 0.8467488658834477

training 5 epoch...

393.5265734195709 seconds used......

5 training finished! average train loss: 0.3961260926155016

total clf loss: 80.50527003407478 total adv loss: 47.489294386003166

evaluating...

Confusion matrix, without normalization

[[ 613 232 177]

[ 739 4509 159]

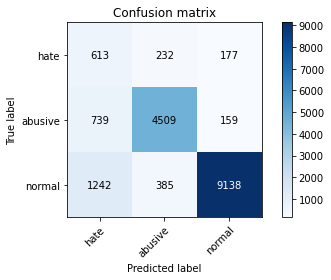
[1242 385 9138]]

Normalized confusion matrix

[[0.59980431 0.22700587 0.17318982]

[0.13667468 0.83391899 0.02940633]

[0.1153739 0.03576405 0.84886205]]





precision recall f1-score support

hate 0.24 0.60 0.34 1022

abusive 0.88 0.83 0.86 5407

normal 0.96 0.85 0.90 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: 0.6044512805853819, accuracy: 0.8293590787484006

2 run...

training 1 epoch...

393.6161615848541 seconds used......

1 training finished! average train loss: 1.303970227190388

total clf loss: 269.24355782568455 total adv loss: 154.20762553066015

evaluating...

Confusion matrix, without normalization

[[ 661 171 190]

[ 794 4431 182]

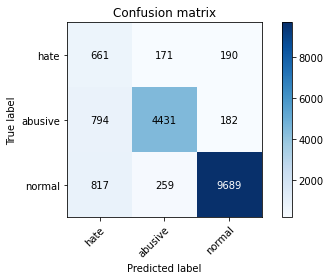
[ 817 259 9689]]

Normalized confusion matrix

[[0.64677104 0.16731898 0.18590998]

[0.14684668 0.81949325 0.03366007]

[0.0758941 0.02405945 0.90004645]]





precision recall f1-score support

hate 0.29 0.65 0.40 1022

abusive 0.91 0.82 0.86 5407

normal 0.96 0.90 0.93 10765

accuracy 0.86 17194

macro avg 0.72 0.79 0.73 17194

weighted avg 0.91 0.86 0.88 17194

average eval\_loss: 0.42724492787275203, accuracy: 0.8596603466325462

training 2 epoch...

393.71425557136536 seconds used......

2 training finished! average train loss: 0.9865104404381113

total clf loss: 204.37958320230246 total adv loss: 116.32227051071823

evaluating...

Confusion matrix, without normalization

[[ 660 191 171]

[ 677 4557 173]

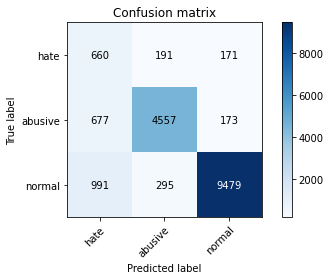
[ 991 295 9479]]

Normalized confusion matrix

[[0.64579256 0.18688845 0.16731898]

[0.12520806 0.84279638 0.03199556]

[0.09205759 0.02740362 0.88053878]]





precision recall f1-score support

hate 0.28 0.65 0.39 1022

abusive 0.90 0.84 0.87 5407

normal 0.96 0.88 0.92 10765

accuracy 0.85 17194

macro avg 0.72 0.79 0.73 17194

weighted avg 0.91 0.85 0.87 17194

average eval\_loss: 0.41030074220670515, accuracy: 0.8547167616610446

training 3 epoch...

394.0521385669708 seconds used......

3 training finished! average train loss: 0.7888935370114264

total clf loss: 163.11872059851885 total adv loss: 93.18055782467127

evaluating...

Confusion matrix, without normalization

[[ 756 118 148]

[1353 3948 106]

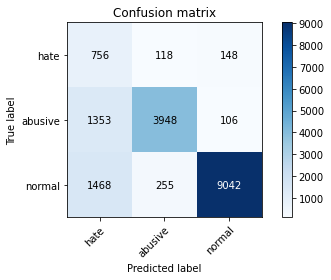
[1468 255 9042]]

Normalized confusion matrix

[[0.73972603 0.11545988 0.14481409]

[0.25023118 0.7301646 0.01960422]

[0.13636786 0.02368788 0.83994426]]





precision recall f1-score support

hate 0.21 0.74 0.33 1022

abusive 0.91 0.73 0.81 5407

normal 0.97 0.84 0.90 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: 0.5459050942645517, accuracy: 0.799464929626614

training 4 epoch...

393.82051277160645 seconds used......

4 training finished! average train loss: 0.5664833770900898

total clf loss: 116.73306479491293 total adv loss: 67.10953536070883

evaluating...

Confusion matrix, without normalization

[[ 644 203 175]

[ 827 4435 145]

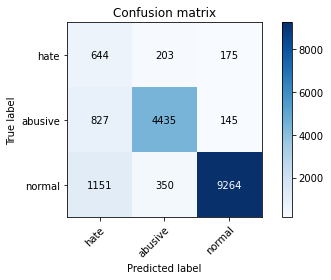
[1151 350 9264]]

Normalized confusion matrix

[[0.63013699 0.19863014 0.17123288]

[0.15294988 0.82023303 0.02681709]

[0.10692058 0.03251277 0.86056665]]





precision recall f1-score support

hate 0.25 0.63 0.35 1022

abusive 0.89 0.82 0.85 5407

normal 0.97 0.86 0.91 10765

accuracy 0.83 17194

macro avg 0.70 0.77 0.71 17194

weighted avg 0.90 0.83 0.86 17194

average eval\_loss: 0.5489515068702573, accuracy: 0.834186344073514

training 5 epoch...

393.69079303741455 seconds used......

5 training finished! average train loss: 0.3859297838566602

total clf loss: 80.05466471146792 total adv loss: 45.456114684056956

evaluating...

Confusion matrix, without normalization

[[ 666 218 138]

[ 885 4405 117]

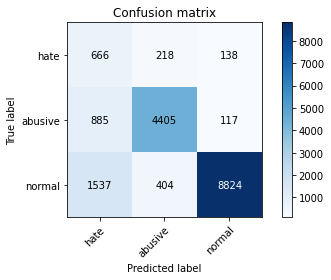
[1537 404 8824]]

Normalized confusion matrix

[[0.65166341 0.21330724 0.13502935]

[0.16367672 0.81468467 0.02163862]

[0.14277752 0.03752903 0.81969345]]





precision recall f1-score support

hate 0.22 0.65 0.32 1022

abusive 0.88 0.81 0.84 5407

normal 0.97 0.82 0.89 10765

accuracy 0.81 17194

macro avg 0.69 0.76 0.69 17194

weighted avg 0.90 0.81 0.84 17194

average eval\_loss: 0.6924574258065848, accuracy: 0.8081307432825404

3 run...

training 1 epoch...

393.7657353878021 seconds used......

1 training finished! average train loss: 1.3046624838767686

total clf loss: 272.94138400256634 total adv loss: 152.51204862445593

evaluating...

Confusion matrix, without normalization

[[ 670 185 167]

[ 763 4509 135]

[1017 330 9418]]

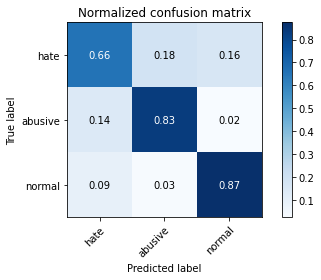
Normalized confusion matrix

[[0.6555773 0.18101761 0.16340509]

[0.14111337 0.83391899 0.02496763]

[0.09447283 0.0306549 0.87487227]]





precision recall f1-score support

hate 0.27 0.66 0.39 1022

abusive 0.90 0.83 0.86 5407

normal 0.97 0.87 0.92 10765

accuracy 0.85 17194

macro avg 0.71 0.79 0.72 17194

weighted avg 0.91 0.85 0.87 17194

average eval\_loss: 0.42996436832256096, accuracy: 0.8489589391648249

training 2 epoch...

393.8607325553894 seconds used......

2 training finished! average train loss: 0.9788047577134643

total clf loss: 203.57877610623837 total adv loss: 115.01586547121406

evaluating...

Confusion matrix, without normalization

[[ 699 167 156]

[ 866 4383 158]

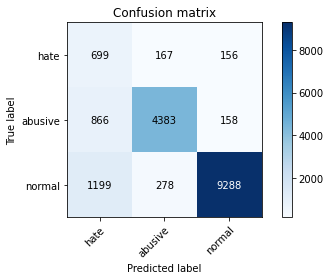
[1199 278 9288]]

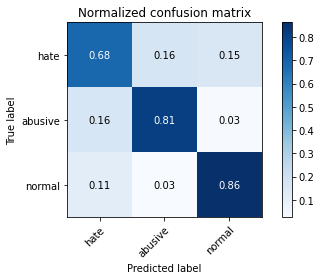
Normalized confusion matrix

[[0.68395303 0.16340509 0.15264188]

[0.16016275 0.81061587 0.02922138]

[0.11137947 0.02582443 0.8627961 ]]





precision recall f1-score support

hate 0.25 0.68 0.37 1022

abusive 0.91 0.81 0.86 5407

normal 0.97 0.86 0.91 10765

accuracy 0.84 17194

macro avg 0.71 0.79 0.71 17194

weighted avg 0.91 0.84 0.86 17194

average eval\_loss: 0.4613111641517905, accuracy: 0.8357566592997557

training 3 epoch...

393.65116596221924 seconds used......

3 training finished! average train loss: 0.7869191690102119

total clf loss: 163.17691856622696 total adv loss: 92.71413715742528

evaluating...

Confusion matrix, without normalization

[[ 591 239 192]

[ 582 4639 186]

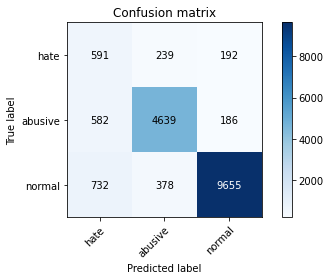
[ 732 378 9655]]

Normalized confusion matrix

[[0.57827789 0.23385519 0.18786693]

[0.10763825 0.8579619 0.03439985]

[0.06799814 0.03511379 0.89688806]]





precision recall f1-score support

hate 0.31 0.58 0.40 1022

abusive 0.88 0.86 0.87 5407

normal 0.96 0.90 0.93 10765

accuracy 0.87 17194

macro avg 0.72 0.78 0.73 17194

weighted avg 0.90 0.87 0.88 17194

average eval\_loss: 0.3898510461456554, accuracy: 0.8657089682447365

training 4 epoch...

393.5042805671692 seconds used......

4 training finished! average train loss: 0.5871979649058853

total clf loss: 120.90689724311233 total adv loss: 69.6109003808815

evaluating...

Confusion matrix, without normalization

[[ 686 179 157]

[ 895 4327 185]

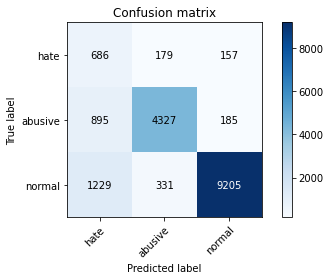
[1229 331 9205]]

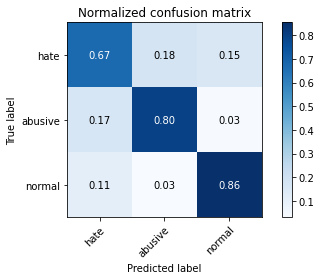
Normalized confusion matrix

[[0.67123288 0.17514677 0.15362035]

[0.16552617 0.80025892 0.03421491]

[0.11416628 0.03074779 0.85508593]]





precision recall f1-score support

hate 0.24 0.67 0.36 1022

abusive 0.89 0.80 0.84 5407

normal 0.96 0.86 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: 0.5295490073603253, accuracy: 0.8269163661742468

training 5 epoch...

393.3324439525604 seconds used......

5 training finished! average train loss: 0.40975769795056927

total clf loss: 85.08435366116464 total adv loss: 48.219152927049436

evaluating...

Confusion matrix, without normalization

[[ 665 227 130]

[ 829 4483 95]

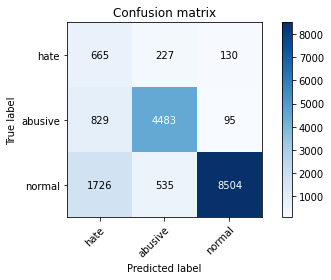
[1726 535 8504]]

Normalized confusion matrix

[[0.65068493 0.2221135 0.12720157]

[0.15331977 0.82911041 0.01756982]

[0.16033442 0.0496981 0.78996749]]





precision recall f1-score support

hate 0.21 0.65 0.31 1022

abusive 0.85 0.83 0.84 5407

normal 0.97 0.79 0.87 10765

accuracy 0.79 17194

macro avg 0.68 0.76 0.68 17194

weighted avg 0.89 0.79 0.83 17194

average eval\_loss: 0.7413836378300952, accuracy: 0.793997906246365

4 run...

training 1 epoch...

393.73865246772766 seconds used......

1 training finished! average train loss: 1.3020735132774972

total clf loss: 271.9543520063162 total adv loss: 152.43210773915052

evaluating...

Confusion matrix, without normalization

[[ 713 191 118]

[ 789 4535 83]

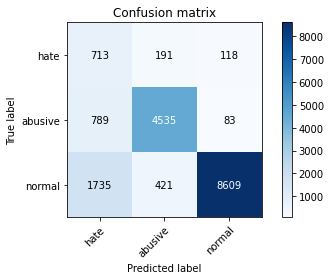
[1735 421 8609]]

Normalized confusion matrix

[[0.69765166 0.18688845 0.11545988]

[0.14592195 0.83872758 0.01535047]

[0.16117046 0.03910822 0.79972132]]





precision recall f1-score support

hate 0.22 0.70 0.33 1022

abusive 0.88 0.84 0.86 5407

normal 0.98 0.80 0.88 10765

accuracy 0.81 17194

macro avg 0.69 0.78 0.69 17194

weighted avg 0.90 0.81 0.84 17194

average eval\_loss: 0.5213474239894124, accuracy: 0.8059206700011632

training 2 epoch...

393.8894000053406 seconds used......

2 training finished! average train loss: 0.9800585955717763

total clf loss: 205.5166223347187 total adv loss: 114.32466785795987

evaluating...

Confusion matrix, without normalization

[[ 710 159 153]

[1010 4263 134]

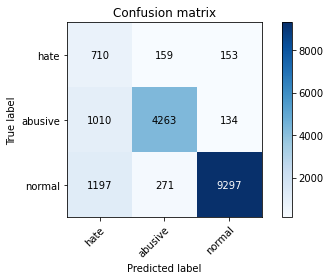
[1197 271 9297]]

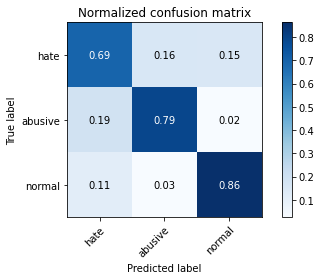
Normalized confusion matrix

[[0.69471624 0.1555773 0.14970646]

[0.1867949 0.78842242 0.02478269]

[0.11119368 0.02517418 0.86363214]]





precision recall f1-score support

hate 0.24 0.69 0.36 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 0.78 0.71 17194

weighted avg 0.91 0.83 0.86 17194

average eval\_loss: 0.46013119184347084, accuracy: 0.829940676980342

training 3 epoch...

393.98332023620605 seconds used......

3 training finished! average train loss: 0.7907083500386899

total clf loss: 165.1203519180417 total adv loss: 92.58172391774133

evaluating...

Confusion matrix, without normalization

[[ 688 213 121]

[ 797 4515 95]

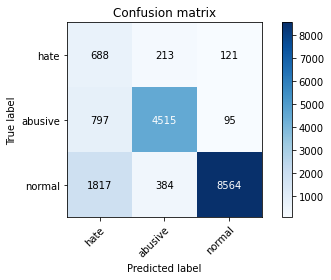
[1817 384 8564]]

Normalized confusion matrix

[[0.67318982 0.20841487 0.1183953 ]

[0.14740152 0.83502867 0.01756982]

[0.16878774 0.03567116 0.79554111]]





precision recall f1-score support

hate 0.21 0.67 0.32 1022

abusive 0.88 0.84 0.86 5407

normal 0.98 0.80 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: 0.5628777375661357, accuracy: 0.8006862859136908

training 4 epoch...

394.0145118236542 seconds used......

4 training finished! average train loss: 0.5680558249085267

total clf loss: 118.05163776129484 total adv loss: 66.79854638315737

evaluating...

Confusion matrix, without normalization

[[ 700 206 116]

[ 911 4397 99]

[2027 419 8319]]

Normalized confusion matrix

[[0.68493151 0.20156556 0.11350294]

[0.1684853 0.8132051 0.0183096 ]

[0.1882954 0.03892243 0.77278216]]





precision recall f1-score support

hate 0.19 0.68 0.30 1022

abusive 0.88 0.81 0.84 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: 0.6716963535044775, accuracy: 0.7802721879725486

training 5 epoch...

393.61537432670593 seconds used......

5 training finished! average train loss: 0.39155545588120244

total clf loss: 81.06257168017328 total adv loss: 46.19824788850383

evaluating...

Confusion matrix, without normalization

[[ 609 234 179]

[ 745 4505 157]

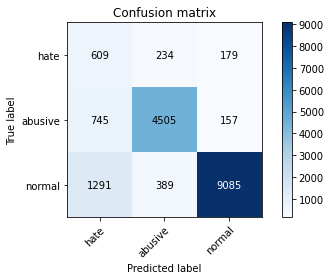
[1291 389 9085]]

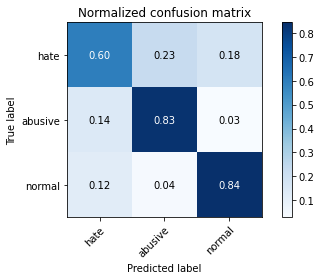
Normalized confusion matrix

[[0.59589041 0.22896282 0.17514677]

[0.13778435 0.83317921 0.02903643]

[0.11992569 0.03613562 0.84393869]]





precision recall f1-score support

hate 0.23 0.60 0.33 1022

abusive 0.88 0.83 0.86 5407

normal 0.96 0.84 0.90 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: 0.6004643317940104, accuracy: 0.8258113295335582

5 run...

training 1 epoch...

393.43156814575195 seconds used......

1 training finished! average train loss: 1.3058634816120225

total clf loss: 271.8538406342268 total adv loss: 153.3218417391181

evaluating...

Confusion matrix, without normalization

[[ 761 147 114]

[1103 4200 104]

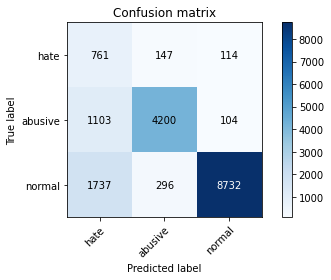
[1737 296 8732]]

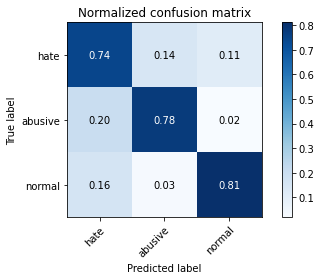
Normalized confusion matrix

[[0.7446184 0.14383562 0.11154599]

[0.20399482 0.77677085 0.01923433]

[0.16135625 0.02749652 0.81114724]]





precision recall f1-score support

hate 0.21 0.74 0.33 1022

abusive 0.90 0.78 0.84 5407

normal 0.98 0.81 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: 0.5260085258026456, accuracy: 0.7963824589973246

training 2 epoch...

393.30487084388733 seconds used......

2 training finished! average train loss: 0.9791100771580539

total clf loss: 204.3521105349064 total adv loss: 114.69682668708265

evaluating...

Confusion matrix, without normalization

[[ 619 264 139]

[ 510 4787 110]

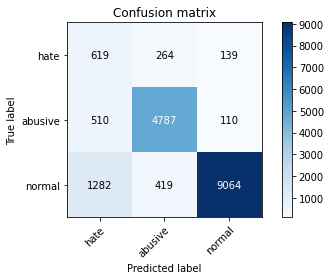
[1282 419 9064]]

Normalized confusion matrix

[[0.60567515 0.25831703 0.13600783]

[0.09432217 0.88533383 0.020344 ]

[0.11908964 0.03892243 0.84198792]]





precision recall f1-score support

hate 0.26 0.61 0.36 1022

abusive 0.88 0.89 0.88 5407

normal 0.97 0.84 0.90 10765

accuracy 0.84 17194

macro avg 0.70 0.78 0.71 17194

weighted avg 0.90 0.84 0.86 17194

average eval\_loss: 0.4486920606431573, accuracy: 0.8415726416191694

training 3 epoch...

393.3911578655243 seconds used......

3 training finished! average train loss: 0.7852126038370498

total clf loss: 163.05072959512472 total adv loss: 92.39922716841102

evaluating...

Confusion matrix, without normalization

[[ 597 273 152]

[ 607 4650 150]

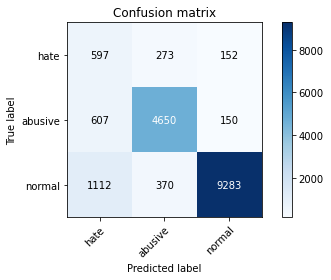
[1112 370 9283]]

Normalized confusion matrix

[[0.58414873 0.26712329 0.14872798]

[0.11226188 0.8599963 0.02774182]

[0.10329772 0.03437065 0.86233163]]





precision recall f1-score support

hate 0.26 0.58 0.36 1022

abusive 0.88 0.86 0.87 5407

normal 0.97 0.86 0.91 10765

accuracy 0.85 17194

macro avg 0.70 0.77 0.71 17194

weighted avg 0.90 0.85 0.87 17194

average eval\_loss: 0.4503563243895769, accuracy: 0.8450622310108177

training 4 epoch...

393.52234148979187 seconds used......

4 training finished! average train loss: 0.5492315547494533

total clf loss: 115.7210193015635 total adv loss: 63.79427986918017

evaluating...

Confusion matrix, without normalization

[[ 603 235 184]

[ 717 4538 152]

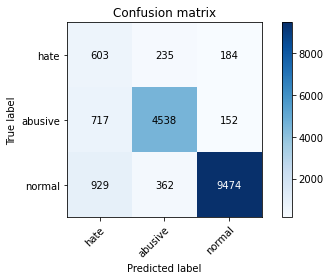
[ 929 362 9474]]

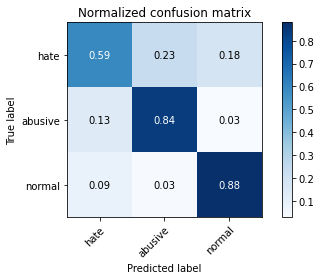
Normalized confusion matrix

[[0.59001957 0.22994129 0.18003914]

[0.13260588 0.83928241 0.02811171]

[0.08629819 0.0336275 0.88007431]]





precision recall f1-score support

hate 0.27 0.59 0.37 1022

abusive 0.88 0.84 0.86 5407

normal 0.97 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: 0.5200219492363028, accuracy: 0.8500058159823194

training 5 epoch...

393.51817417144775 seconds used......

5 training finished! average train loss: 0.38652295835576866

total clf loss: 80.24714324157685 total adv loss: 45.49126367573626

evaluating...

Confusion matrix, without normalization

[[ 698 211 113]

[ 836 4441 130]

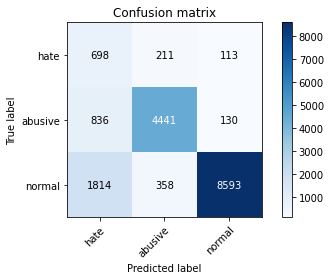
[1814 358 8593]]

Normalized confusion matrix

[[0.68297456 0.20645793 0.11056751]

[0.15461439 0.8213427 0.02404291]

[0.16850906 0.03325592 0.79823502]]





precision recall f1-score support

hate 0.21 0.68 0.32 1022

abusive 0.89 0.82 0.85 5407

normal 0.97 0.80 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: 0.6948874756953743, accuracy: 0.798650692101896