training 1 epoch...

52.159801721572876 seconds used......

1 training finished! average train loss: 0.6695682402939824

total clf loss: 118.25235909223557 total adv loss: 75.12380108237267

evaluating...

Confusion matrix, without normalization

[[ 193 72 29]

[ 454 3256 111]

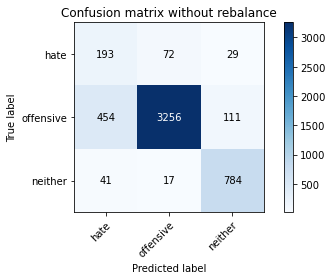
[ 41 17 784]]

Normalized confusion matrix

[[0.65646259 0.24489796 0.09863946]

[0.11881706 0.85213295 0.02904999]

[0.04869359 0.02019002 0.93111639]]





precision recall f1-score support

hate 0.28 0.66 0.39 294

offensive 0.97 0.85 0.91 3821

neither 0.85 0.93 0.89 842

accuracy 0.85 4957

macro avg 0.70 0.81 0.73 4957

weighted avg 0.91 0.85 0.87 4957

average eval\_loss: 0.37895000188220895, accuracy: 0.8539439176921525

training 2 epoch...

52.129839181900024 seconds used......

2 training finished! average train loss: 0.38803567671984956

total clf loss: 68.93808858841658 total adv loss: 51.67975728958845

evaluating...

Confusion matrix, without normalization

[[ 249 35 10]

[1016 2743 62]

[ 98 24 720]]

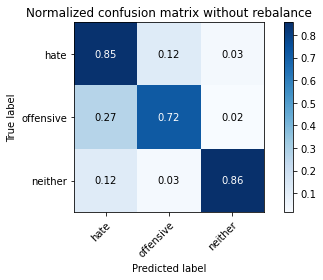
Normalized confusion matrix

[[0.84693878 0.11904762 0.03401361]

[0.26589898 0.7178749 0.01622612]

[0.11638955 0.02850356 0.85510689]]





precision recall f1-score support

hate 0.18 0.85 0.30 294

offensive 0.98 0.72 0.83 3821

neither 0.91 0.86 0.88 842

accuracy 0.75 4957

macro avg 0.69 0.81 0.67 4957

weighted avg 0.92 0.75 0.81 4957

average eval\_loss: 0.5898055970788963, accuracy: 0.7488400242081904

training 3 epoch...

52.112181663513184 seconds used......

3 training finished! average train loss: 0.2695326800104122

total clf loss: 48.16131425276399 total adv loss: 41.42451928369701

evaluating...

Confusion matrix, without normalization

[[ 197 76 21]

[ 413 3293 115]

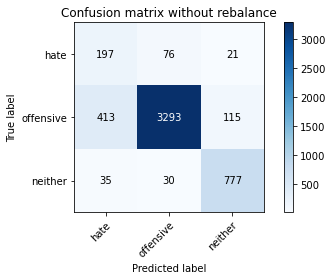
[ 35 30 777]]

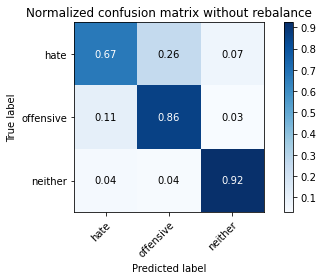
Normalized confusion matrix

[[0.67006803 0.2585034 0.07142857]

[0.10808689 0.86181628 0.03009683]

[0.0415677 0.03562945 0.92280285]]





precision recall f1-score support

hate 0.31 0.67 0.42 294

offensive 0.97 0.86 0.91 3821

neither 0.85 0.92 0.89 842

accuracy 0.86 4957

macro avg 0.71 0.82 0.74 4957

weighted avg 0.91 0.86 0.88 4957

average eval\_loss: 0.38541478979130905, accuracy: 0.8608029049828525

training 4 epoch...

52.107786655426025 seconds used......

4 training finished! average train loss: 0.15731187270442296

total clf loss: 28.253485575318336 total adv loss: 27.063107697293162

evaluating...

Confusion matrix, without normalization

[[ 223 53 18]

[ 663 3049 109]

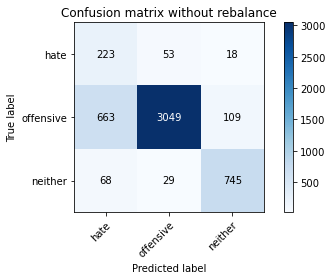
[ 68 29 745]]

Normalized confusion matrix

[[0.7585034 0.18027211 0.06122449]

[0.17351479 0.79795865 0.02852656]

[0.0807601 0.03444181 0.8847981 ]]





precision recall f1-score support

hate 0.23 0.76 0.36 294

offensive 0.97 0.80 0.88 3821

neither 0.85 0.88 0.87 842

accuracy 0.81 4957

macro avg 0.69 0.81 0.70 4957

weighted avg 0.91 0.81 0.84 4957

average eval\_loss: 0.556540418498879, accuracy: 0.8103691749041759

training 5 epoch...

52.13081479072571 seconds used......

5 training finished! average train loss: 0.08978587079913034

total clf loss: 16.189108482096344 total adv loss: 16.714491399936378

evaluating...

Confusion matrix, without normalization

[[ 224 51 19]

[ 890 2832 99]

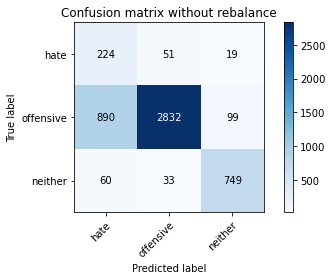
[ 60 33 749]]

Normalized confusion matrix

[[0.76190476 0.17346939 0.06462585]

[0.23292332 0.74116723 0.02590945]

[0.07125891 0.0391924 0.88954869]]





precision recall f1-score support

hate 0.19 0.76 0.31 294

offensive 0.97 0.74 0.84 3821

neither 0.86 0.89 0.88 842

accuracy 0.77 4957

macro avg 0.68 0.80 0.67 4957

weighted avg 0.91 0.77 0.82 4957

average eval\_loss: 0.7808423396951009, accuracy: 0.7676013717974581

training 6 epoch...

52.32032322883606 seconds used......

6 training finished! average train loss: 0.04215005330049722

total clf loss: 7.60661827219883 total adv loss: 7.979183689865749

evaluating...

Confusion matrix, without normalization

[[ 210 65 19]

[ 585 3121 115]

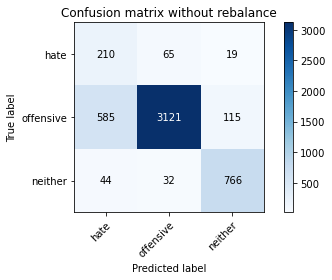
[ 44 32 766]]

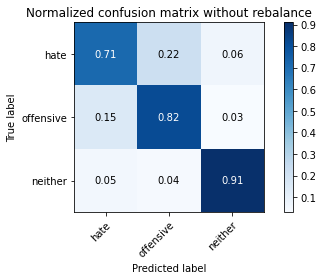
Normalized confusion matrix

[[0.71428571 0.22108844 0.06462585]

[0.15310128 0.81680188 0.03009683]

[0.05225653 0.03800475 0.90973872]]





precision recall f1-score support

hate 0.25 0.71 0.37 294

offensive 0.97 0.82 0.89 3821

neither 0.85 0.91 0.88 842

accuracy 0.83 4957

macro avg 0.69 0.81 0.71 4957

weighted avg 0.91 0.83 0.85 4957

average eval\_loss: 0.7201758963346947, accuracy: 0.8265079685293524

training 7 epoch...

52.18972373008728 seconds used......

7 training finished! average train loss: 0.03209095524299933

total clf loss: 5.802832291054074 total adv loss: 6.305578809173312

evaluating...

Confusion matrix, without normalization

[[ 228 50 16]

[ 876 2838 107]

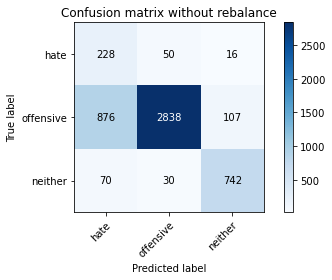
[ 70 30 742]]

Normalized confusion matrix

[[0.7755102 0.17006803 0.05442177]

[0.22925936 0.7427375 0.02800314]

[0.08313539 0.03562945 0.88123515]]





precision recall f1-score support

hate 0.19 0.78 0.31 294

offensive 0.97 0.74 0.84 3821

neither 0.86 0.88 0.87 842

accuracy 0.77 4957

macro avg 0.67 0.80 0.67 4957

weighted avg 0.91 0.77 0.82 4957

average eval\_loss: 1.1045433446960222, accuracy: 0.7682065765584023

training 8 epoch...

52.26701641082764 seconds used......

8 training finished! average train loss: 0.02738294621118704

total clf loss: 4.957764858350856 total adv loss: 5.505620868032565

evaluating...

Confusion matrix, without normalization

[[ 202 68 24]

[ 610 3086 125]

[ 41 32 769]]

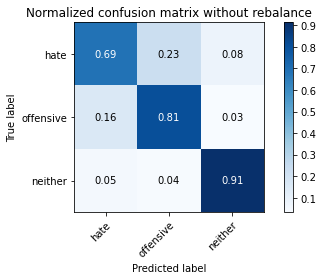
Normalized confusion matrix

[[0.68707483 0.23129252 0.08163265]

[0.15964407 0.80764198 0.03271395]

[0.04869359 0.03800475 0.91330166]]





precision recall f1-score support

hate 0.24 0.69 0.35 294

offensive 0.97 0.81 0.88 3821

neither 0.84 0.91 0.87 842

accuracy 0.82 4957

macro avg 0.68 0.80 0.70 4957

weighted avg 0.90 0.82 0.85 4957

average eval\_loss: 0.815923257782537, accuracy: 0.8184385717167642

training 9 epoch...

52.217323780059814 seconds used......

9 training finished! average train loss: 0.018212049908300165

total clf loss: 3.2841942303057294 total adv loss: 3.398674110037973

evaluating...

Confusion matrix, without normalization

[[ 209 68 17]

[ 638 3095 88]

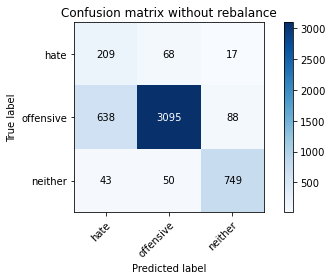
[ 43 50 749]]

Normalized confusion matrix

[[0.71088435 0.23129252 0.05782313]

[0.166972 0.80999738 0.02303062]

[0.05106888 0.05938242 0.88954869]]





precision recall f1-score support

hate 0.23 0.71 0.35 294

offensive 0.96 0.81 0.88 3821

neither 0.88 0.89 0.88 842

accuracy 0.82 4957

macro avg 0.69 0.80 0.71 4957

weighted avg 0.91 0.82 0.85 4957

average eval\_loss: 0.9084680149080081, accuracy: 0.8176316320355054

training 10 epoch...

52.29458689689636 seconds used......

10 training finished! average train loss: 0.01779440759098408

total clf loss: 3.1893360542308073 total adv loss: 2.929847372317454

evaluating...

Confusion matrix, without normalization

[[ 239 35 20]

[ 933 2752 136]

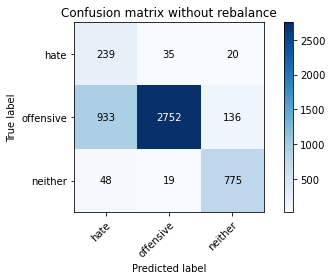
[ 48 19 775]]

Normalized confusion matrix

[[0.81292517 0.11904762 0.06802721]

[0.24417692 0.72023031 0.03559278]

[0.05700713 0.02256532 0.92042755]]





precision recall f1-score support

hate 0.20 0.81 0.32 294

offensive 0.98 0.72 0.83 3821

neither 0.83 0.92 0.87 842

accuracy 0.76 4957

macro avg 0.67 0.82 0.67 4957

weighted avg 0.91 0.76 0.81 4957

average eval\_loss: 1.4040623077251497, accuracy: 0.7597337099051846