training 1 epoch...

49.87483334541321 seconds used......

1 training finished! average train loss: 0.6748223990550515

total clf loss: 119.15362836420536 total adv loss: 75.17996249347925

evaluating...

Confusion matrix, without normalization

[[ 218 43 33]

[ 783 2861 177]

[ 45 9 788]]

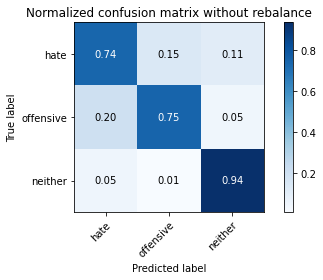
Normalized confusion matrix

[[0.7414966 0.1462585 0.1122449 ]

[0.20492018 0.74875687 0.04632295]

[0.05344418 0.01068884 0.93586698]]





precision recall f1-score support

hate 0.21 0.74 0.33 294

offensive 0.98 0.75 0.85 3821

neither 0.79 0.94 0.86 842

accuracy 0.78 4957

macro avg 0.66 0.81 0.68 4957

weighted avg 0.90 0.78 0.82 4957

average eval\_loss: 0.502119435387994, accuracy: 0.7801089368569699

training 2 epoch...

49.721786975860596 seconds used......

2 training finished! average train loss: 0.385849335824537

total clf loss: 68.55870373547077 total adv loss: 51.569349467754364

evaluating...

Confusion matrix, without normalization

[[ 226 48 20]

[ 687 2978 156]

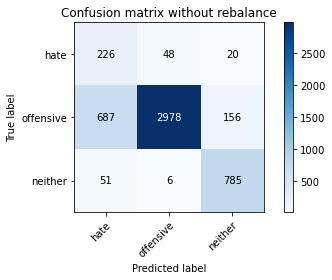
[ 51 6 785]]

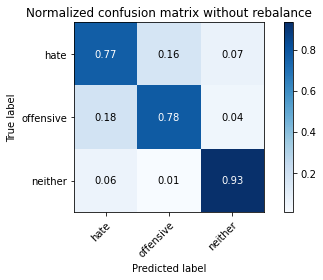
Normalized confusion matrix

[[0.76870748 0.16326531 0.06802721]

[0.17979586 0.77937713 0.04082701]

[0.06057007 0.00712589 0.93230404]]





precision recall f1-score support

hate 0.23 0.77 0.36 294

offensive 0.98 0.78 0.87 3821

neither 0.82 0.93 0.87 842

accuracy 0.80 4957

macro avg 0.68 0.83 0.70 4957

weighted avg 0.91 0.80 0.84 4957

average eval\_loss: 0.4694533568897074, accuracy: 0.8047205971353641

training 3 epoch...

49.719783544540405 seconds used......

3 training finished! average train loss: 0.24608881812956598

total clf loss: 43.99354284629226 total adv loss: 38.247099148109555

evaluating...

Confusion matrix, without normalization

[[ 217 61 16]

[ 582 3137 102]

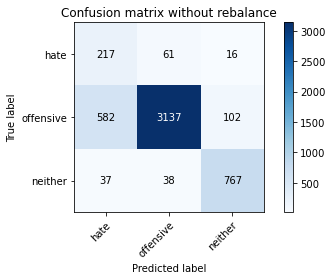
[ 37 38 767]]

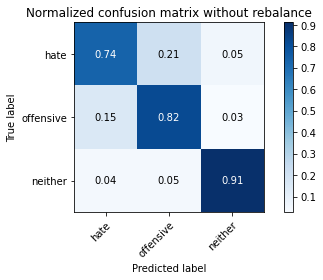
Normalized confusion matrix

[[0.73809524 0.20748299 0.05442177]

[0.15231615 0.82098927 0.02669458]

[0.04394299 0.04513064 0.91092637]]





precision recall f1-score support

hate 0.26 0.74 0.38 294

offensive 0.97 0.82 0.89 3821

neither 0.87 0.91 0.89 842

accuracy 0.83 4957

macro avg 0.70 0.82 0.72 4957

weighted avg 0.91 0.83 0.86 4957

average eval\_loss: 0.42829944884344456, accuracy: 0.8313496066169054

training 4 epoch...

49.68657970428467 seconds used......

4 training finished! average train loss: 0.13076966393967732

total clf loss: 23.602212123572826 total adv loss: 24.811990131624043

evaluating...

Confusion matrix, without normalization

[[ 236 45 13]

[1072 2646 103]

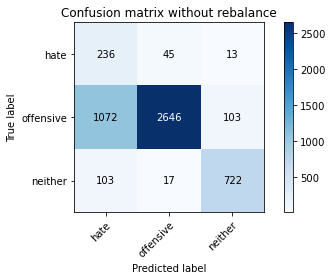
[ 103 17 722]]

Normalized confusion matrix

[[0.80272109 0.15306122 0.04421769]

[0.28055483 0.69248888 0.02695629]

[0.12232779 0.02019002 0.85748219]]





precision recall f1-score support

hate 0.17 0.80 0.28 294

offensive 0.98 0.69 0.81 3821

neither 0.86 0.86 0.86 842

accuracy 0.73 4957

macro avg 0.67 0.78 0.65 4957

weighted avg 0.91 0.73 0.79 4957

average eval\_loss: 0.8035423751589993, accuracy: 0.7270526528142022

training 5 epoch...

49.70374631881714 seconds used......

5 training finished! average train loss: 0.0748570696763934

total clf loss: 13.51234128838405 total adv loss: 14.235647554509342

evaluating...

Confusion matrix, without normalization

[[ 206 63 25]

[ 641 3022 158]

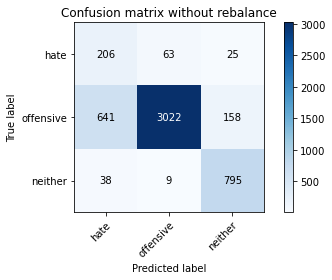
[ 38 9 795]]

Normalized confusion matrix

[[0.70068027 0.21428571 0.08503401]

[0.16775713 0.79089244 0.04135043]

[0.04513064 0.01068884 0.94418052]]





precision recall f1-score support

hate 0.23 0.70 0.35 294

offensive 0.98 0.79 0.87 3821

neither 0.81 0.94 0.87 842

accuracy 0.81 4957

macro avg 0.67 0.81 0.70 4957

weighted avg 0.90 0.81 0.84 4957

average eval\_loss: 0.6994969806936194, accuracy: 0.8115795844260641

training 6 epoch...

49.73668956756592 seconds used......

6 training finished! average train loss: 0.03213081409411276

total clf loss: 5.808864649850875 total adv loss: 6.2899097839253955

evaluating...

Confusion matrix, without normalization

[[ 228 52 14]

[ 925 2801 95]

[ 75 26 741]]

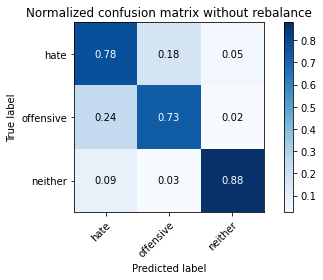
Normalized confusion matrix

[[0.7755102 0.17687075 0.04761905]

[0.24208322 0.73305417 0.0248626 ]

[0.08907363 0.03087886 0.88004751]]





precision recall f1-score support

hate 0.19 0.78 0.30 294

offensive 0.97 0.73 0.84 3821

neither 0.87 0.88 0.88 842

accuracy 0.76 4957

macro avg 0.68 0.80 0.67 4957

weighted avg 0.91 0.76 0.81 4957

average eval\_loss: 1.0558294272820106, accuracy: 0.7605406495864434

training 7 epoch...

49.723801612854004 seconds used......

7 training finished! average train loss: 0.02109831769042512

total clf loss: 3.850758232874796 total adv loss: 4.858918162935879

evaluating...

Confusion matrix, without normalization

[[ 223 52 19]

[ 777 2931 113]

[ 47 20 775]]

Normalized confusion matrix

[[0.7585034 0.17687075 0.06462585]

[0.20334991 0.76707668 0.02957341]

[0.05581948 0.02375297 0.92042755]]





precision recall f1-score support

hate 0.21 0.76 0.33 294

offensive 0.98 0.77 0.86 3821

neither 0.85 0.92 0.89 842

accuracy 0.79 4957

macro avg 0.68 0.82 0.69 4957

weighted avg 0.91 0.79 0.83 4957

average eval\_loss: 1.0264779333458958, accuracy: 0.7926165019164817

training 8 epoch...

49.759581089019775 seconds used......

8 training finished! average train loss: 0.03079022906365946

total clf loss: 5.495359623222612 total adv loss: 4.604609288202482

evaluating...

Confusion matrix, without normalization

[[ 200 74 20]

[ 562 3156 103]

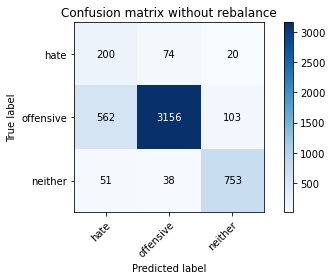
[ 51 38 753]]

Normalized confusion matrix

[[0.68027211 0.25170068 0.06802721]

[0.14708192 0.82596179 0.02695629]

[0.06057007 0.04513064 0.89429929]]





precision recall f1-score support

hate 0.25 0.68 0.36 294

offensive 0.97 0.83 0.89 3821

neither 0.86 0.89 0.88 842

accuracy 0.83 4957

macro avg 0.69 0.80 0.71 4957

weighted avg 0.91 0.83 0.86 4957

average eval\_loss: 0.7975277364030555, accuracy: 0.8289287875731289

training 9 epoch...

49.716275453567505 seconds used......

9 training finished! average train loss: 0.02012538632557711

total clf loss: 3.6490693412488326 total adv loss: 4.152565282565774

evaluating...

Confusion matrix, without normalization

[[ 210 67 17]

[ 561 3143 117]

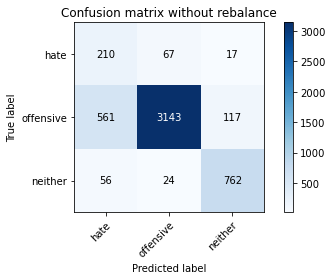
[ 56 24 762]]

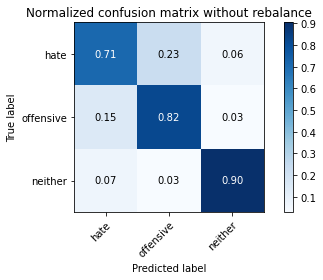
Normalized confusion matrix

[[0.71428571 0.22789116 0.05782313]

[0.1468202 0.82255954 0.03062026]

[0.06650831 0.02850356 0.90498812]]





precision recall f1-score support

hate 0.25 0.71 0.37 294

offensive 0.97 0.82 0.89 3821

neither 0.85 0.90 0.88 842

accuracy 0.83 4957

macro avg 0.69 0.81 0.71 4957

weighted avg 0.91 0.83 0.86 4957

average eval\_loss: 0.9080938801494232, accuracy: 0.8301391970950172

training 10 epoch...

49.689444065093994 seconds used......

10 training finished! average train loss: 0.006997783438487071

total clf loss: 1.2872496791387675 total adv loss: 1.8125741769763408

evaluating...

Confusion matrix, without normalization

[[ 215 55 24]

[ 603 3069 149]

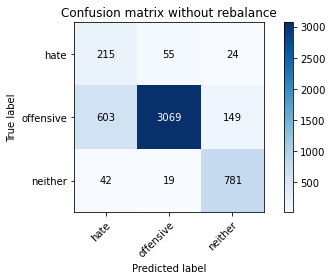
[ 42 19 781]]

Normalized confusion matrix

[[0.73129252 0.18707483 0.08163265]

[0.15781209 0.80319288 0.03899503]

[0.04988124 0.02256532 0.92755344]]





precision recall f1-score support

hate 0.25 0.73 0.37 294

offensive 0.98 0.80 0.88 3821

neither 0.82 0.93 0.87 842

accuracy 0.82 4957

macro avg 0.68 0.82 0.71 4957

weighted avg 0.91 0.82 0.85 4957

average eval\_loss: 1.1166281245832872, accuracy: 0.8200524510792818