1 run......

Downloading: 100%

464/464 [00:00<00:00, 16.5kB/s]

Downloading: 100%

270M/270M [00:08<00:00, 30.5MB/s]

training 1 epoch...

9.77463674545288 seconds used......

1 training finished! average train loss: 1.0443979303042095

total clf loss: 38.8395174741745 total adv loss: 24.823840379714966

evaluating...

Confusion matrix, without normalization

[[ 0 31 10]

[ 0 29 66]

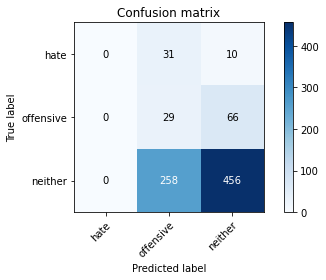
[ 0 258 456]]

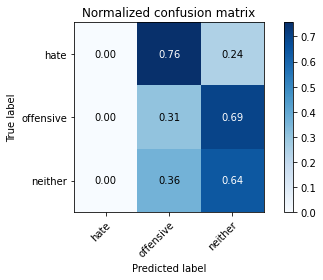
Normalized confusion matrix

[[0. 0.75609756 0.24390244]

[0. 0.30526316 0.69473684]

[0. 0.36134454 0.63865546]]





/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.09 0.31 0.14 95

neither 0.86 0.64 0.73 714

accuracy 0.57 850

macro avg 0.32 0.31 0.29 850

weighted avg 0.73 0.57 0.63 850

average eval\_loss: 0.9107906922351482, accuracy: 0.5705882352941176

training 2 epoch...

8.97422170639038 seconds used......

2 training finished! average train loss: 0.9135121438238356

total clf loss: 34.026667416095734 total adv loss: 22.804607331752777

evaluating...

Confusion matrix, without normalization

[[ 0 25 16]

[ 0 24 71]

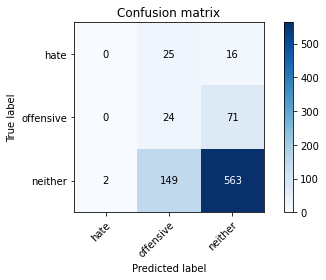
[ 2 149 563]]

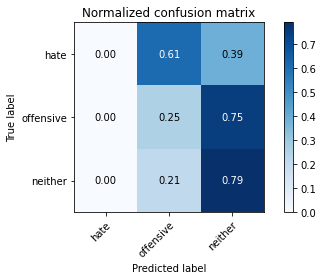
Normalized confusion matrix

[[0. 0.6097561 0.3902439 ]

[0. 0.25263158 0.74736842]

[0.00280112 0.20868347 0.78851541]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.12 0.25 0.16 95

neither 0.87 0.79 0.83 714

accuracy 0.69 850

macro avg 0.33 0.35 0.33 850

weighted avg 0.74 0.69 0.71 850

average eval\_loss: 0.8005456709584524, accuracy: 0.6905882352941176

training 3 epoch...

8.948679685592651 seconds used......

3 training finished! average train loss: 0.7948501391543282

total clf loss: 29.640138924121857 total adv loss: 20.51068076491356

evaluating...

Confusion matrix, without normalization

[[ 2 14 25]

[ 0 12 83]

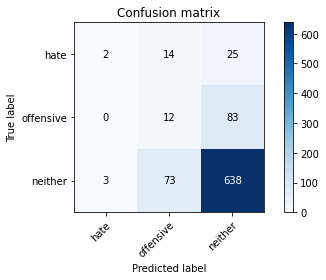
[ 3 73 638]]

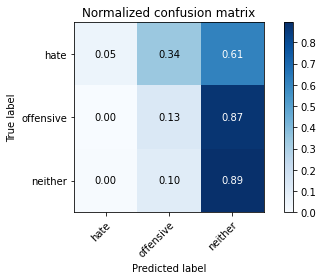
Normalized confusion matrix

[[0.04878049 0.34146341 0.6097561 ]

[0. 0.12631579 0.87368421]

[0.00420168 0.1022409 0.89355742]]





precision recall f1-score support

hate 0.40 0.05 0.09 41

offensive 0.12 0.13 0.12 95

neither 0.86 0.89 0.87 714

accuracy 0.77 850

macro avg 0.46 0.36 0.36 850

weighted avg 0.75 0.77 0.75 850

average eval\_loss: 0.6331580287322055, accuracy: 0.7670588235294118

training 4 epoch...

8.956411123275757 seconds used......

4 training finished! average train loss: 0.6653994537062116

total clf loss: 24.96354803442955 total adv loss: 20.183356285095215

evaluating...

Confusion matrix, without normalization

[[ 5 18 18]

[ 2 52 41]

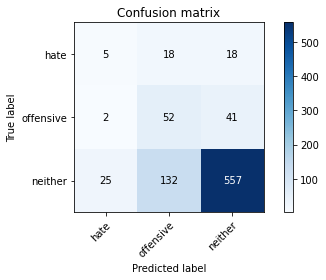
[ 25 132 557]]

Normalized confusion matrix

[[0.12195122 0.43902439 0.43902439]

[0.02105263 0.54736842 0.43157895]

[0.03501401 0.18487395 0.78011204]]





precision recall f1-score support

hate 0.16 0.12 0.14 41

offensive 0.26 0.55 0.35 95

neither 0.90 0.78 0.84 714

accuracy 0.72 850

macro avg 0.44 0.48 0.44 850

weighted avg 0.80 0.72 0.75 850

average eval\_loss: 0.6959676334504471, accuracy: 0.7223529411764706

training 5 epoch...

8.94086766242981 seconds used......

5 training finished! average train loss: 0.43989016198449665

total clf loss: 16.364782199263573 total adv loss: 10.574724704027176

evaluating...

Confusion matrix, without normalization

[[ 10 22 9]

[ 6 67 22]

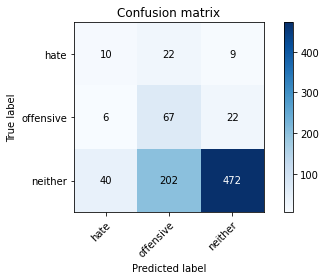
[ 40 202 472]]

Normalized confusion matrix

[[0.24390244 0.53658537 0.2195122 ]

[0.06315789 0.70526316 0.23157895]

[0.05602241 0.28291317 0.66106443]]





precision recall f1-score support

hate 0.18 0.24 0.21 41

offensive 0.23 0.71 0.35 95

neither 0.94 0.66 0.78 714

accuracy 0.65 850

macro avg 0.45 0.54 0.44 850

weighted avg 0.82 0.65 0.70 850

average eval\_loss: 0.8560256615800913, accuracy: 0.6458823529411765

training 6 epoch...

8.872042417526245 seconds used......

6 training finished! average train loss: 0.2632604551812013

total clf loss: 9.785439170897007 total adv loss: 6.161255741491914

evaluating...

Confusion matrix, without normalization

[[ 14 12 15]

[ 6 47 42]

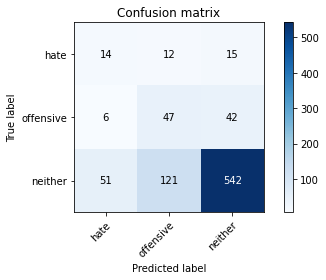
[ 51 121 542]]

Normalized confusion matrix

[[0.34146341 0.29268293 0.36585366]

[0.06315789 0.49473684 0.44210526]

[0.07142857 0.16946779 0.75910364]]





precision recall f1-score support

hate 0.20 0.34 0.25 41

offensive 0.26 0.49 0.34 95

neither 0.90 0.76 0.83 714

accuracy 0.71 850

macro avg 0.45 0.53 0.47 850

weighted avg 0.80 0.71 0.74 850

average eval\_loss: 0.7598447361961007, accuracy: 0.7094117647058824

training 7 epoch...

8.950706005096436 seconds used......

7 training finished! average train loss: 0.16141009537710083

total clf loss: 6.08597806841135 total adv loss: 5.504293534904718

evaluating...

Confusion matrix, without normalization

[[ 17 13 11]

[ 11 51 33]

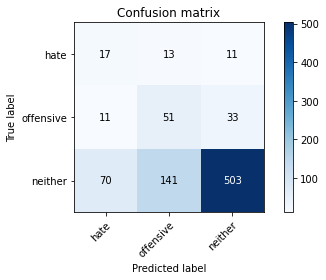
[ 70 141 503]]

Normalized confusion matrix

[[0.41463415 0.31707317 0.26829268]

[0.11578947 0.53684211 0.34736842]

[0.09803922 0.19747899 0.70448179]]





precision recall f1-score support

hate 0.17 0.41 0.24 41

offensive 0.25 0.54 0.34 95

neither 0.92 0.70 0.80 714

accuracy 0.67 850

macro avg 0.45 0.55 0.46 850

weighted avg 0.81 0.67 0.72 850

average eval\_loss: 0.931997504626769, accuracy: 0.6717647058823529

training 8 epoch...

8.982971906661987 seconds used......

8 training finished! average train loss: 0.0965655800472531

total clf loss: 3.6286821644753218 total adv loss: 3.0464259227737784

evaluating...

Confusion matrix, without normalization

[[ 11 11 19]

[ 7 45 43]

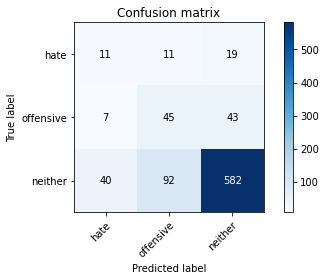
[ 40 92 582]]

Normalized confusion matrix

[[0.26829268 0.26829268 0.46341463]

[0.07368421 0.47368421 0.45263158]

[0.05602241 0.12885154 0.81512605]]





precision recall f1-score support

hate 0.19 0.27 0.22 41

offensive 0.30 0.47 0.37 95

neither 0.90 0.82 0.86 714

accuracy 0.75 850

macro avg 0.47 0.52 0.48 850

weighted avg 0.80 0.75 0.77 850

average eval\_loss: 0.7501801767408155, accuracy: 0.7505882352941177

training 9 epoch...

8.99692440032959 seconds used......

9 training finished! average train loss: 0.026961950365350477

total clf loss: 0.9973305249586701 total adv loss: 0.5340063312905841

evaluating...

Confusion matrix, without normalization

[[ 17 12 12]

[ 13 58 24]

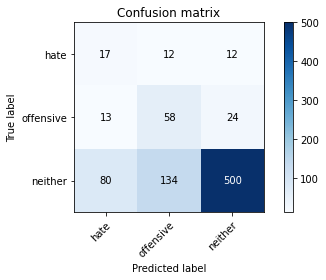
[ 80 134 500]]

Normalized confusion matrix

[[0.41463415 0.29268293 0.29268293]

[0.13684211 0.61052632 0.25263158]

[0.11204482 0.18767507 0.70028011]]





precision recall f1-score support

hate 0.15 0.41 0.23 41

offensive 0.28 0.61 0.39 95

neither 0.93 0.70 0.80 714

accuracy 0.68 850

macro avg 0.46 0.58 0.47 850

weighted avg 0.82 0.68 0.73 850

average eval\_loss: 1.0058352589325676, accuracy: 0.6764705882352942

training 10 epoch...

8.959996700286865 seconds used......

10 training finished! average train loss: 0.012213475924606124

total clf loss: 0.4425118239596486 total adv loss: 0.056533785435021855

evaluating...

Confusion matrix, without normalization

[[ 15 13 13]

[ 8 60 27]

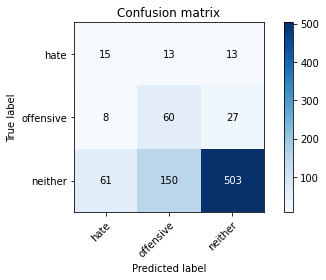
[ 61 150 503]]

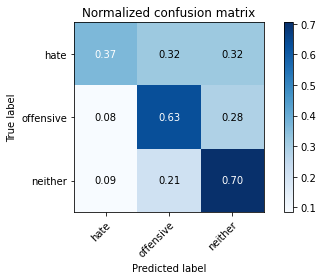
Normalized confusion matrix

[[0.36585366 0.31707317 0.31707317]

[0.08421053 0.63157895 0.28421053]

[0.08543417 0.21008403 0.70448179]]





precision recall f1-score support

hate 0.18 0.37 0.24 41

offensive 0.27 0.63 0.38 95

neither 0.93 0.70 0.80 714

accuracy 0.68 850

macro avg 0.46 0.57 0.47 850

weighted avg 0.82 0.68 0.73 850

average eval\_loss: 1.066766189064744, accuracy: 0.68

training 11 epoch...

8.96493124961853 seconds used......

11 training finished! average train loss: 0.007156881600773583

total clf loss: 0.25808204896748066 total adv loss: 0.008686274302817765

evaluating...

Confusion matrix, without normalization

[[ 17 11 13]

[ 11 52 32]

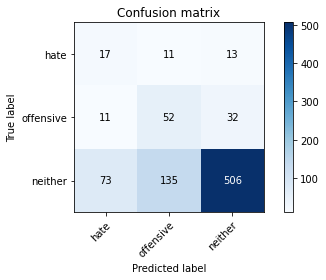
[ 73 135 506]]

Normalized confusion matrix

[[0.41463415 0.26829268 0.31707317]

[0.11578947 0.54736842 0.33684211]

[0.1022409 0.18907563 0.70868347]]





precision recall f1-score support

hate 0.17 0.41 0.24 41

offensive 0.26 0.55 0.35 95

neither 0.92 0.71 0.80 714

accuracy 0.68 850

macro avg 0.45 0.56 0.46 850

weighted avg 0.81 0.68 0.72 850

average eval\_loss: 1.1145549189450956, accuracy: 0.6764705882352942

training 12 epoch...

8.938297986984253 seconds used......

12 training finished! average train loss: 0.004603984109255382

total clf loss: 0.16590624197851866 total adv loss: 0.0032562853739364073

evaluating...

Confusion matrix, without normalization

[[ 15 16 10]

[ 9 62 24]

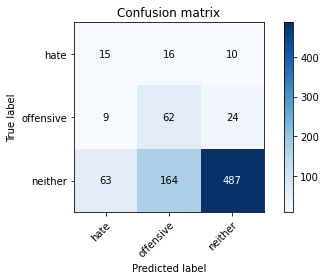
[ 63 164 487]]

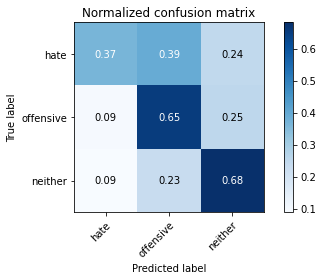
Normalized confusion matrix

[[0.36585366 0.3902439 0.24390244]

[0.09473684 0.65263158 0.25263158]

[0.08823529 0.22969188 0.68207283]]





precision recall f1-score support

hate 0.17 0.37 0.23 41

offensive 0.26 0.65 0.37 95

neither 0.93 0.68 0.79 714

accuracy 0.66 850

macro avg 0.45 0.57 0.46 850

weighted avg 0.82 0.66 0.71 850

average eval\_loss: 1.2633902664339647, accuracy: 0.6635294117647059

training 13 epoch...

8.92268419265747 seconds used......

13 training finished! average train loss: 0.003649560060390892

total clf loss: 0.13152203970821574 total adv loss: 0.002757551936838354

evaluating...

Confusion matrix, without normalization

[[ 15 12 14]

[ 9 54 32]

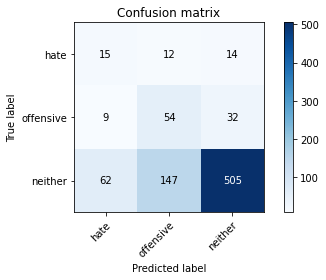
[ 62 147 505]]

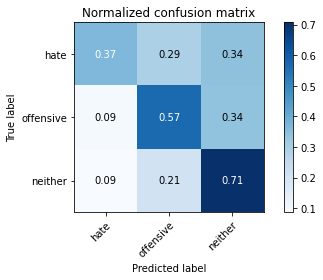
Normalized confusion matrix

[[0.36585366 0.29268293 0.34146341]

[0.09473684 0.56842105 0.33684211]

[0.08683473 0.20588235 0.70728291]]





precision recall f1-score support

hate 0.17 0.37 0.24 41

offensive 0.25 0.57 0.35 95

neither 0.92 0.71 0.80 714

accuracy 0.68 850

macro avg 0.45 0.55 0.46 850

weighted avg 0.81 0.68 0.72 850

average eval\_loss: 1.1877558890264481, accuracy: 0.6752941176470588

training 14 epoch...

8.908011436462402 seconds used......

14 training finished! average train loss: 0.002929576973353202

total clf loss: 0.10552651190664619 total adv loss: 0.0012348240870778682

evaluating...

Confusion matrix, without normalization

[[ 12 13 16]

[ 9 54 32]

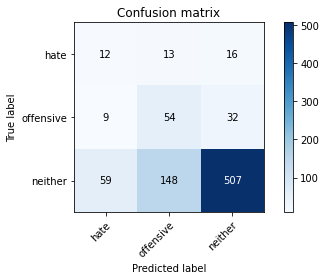
[ 59 148 507]]

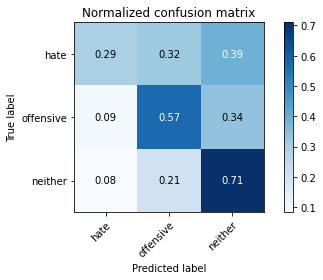
Normalized confusion matrix

[[0.29268293 0.31707317 0.3902439 ]

[0.09473684 0.56842105 0.33684211]

[0.08263305 0.20728291 0.71008403]]





precision recall f1-score support

hate 0.15 0.29 0.20 41

offensive 0.25 0.57 0.35 95

neither 0.91 0.71 0.80 714

accuracy 0.67 850

macro avg 0.44 0.52 0.45 850

weighted avg 0.80 0.67 0.72 850

average eval\_loss: 1.2094311540253286, accuracy: 0.6741176470588235

training 15 epoch...

8.972638368606567 seconds used......

15 training finished! average train loss: 0.0022775745827756408

total clf loss: 0.08202194617479108 total adv loss: 0.0005852232286542858

evaluating...

Confusion matrix, without normalization

[[ 14 14 13]

[ 9 59 27]

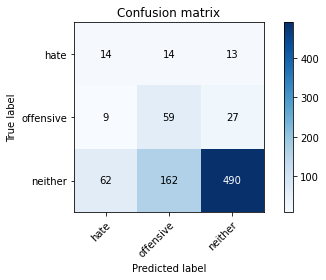
[ 62 162 490]]

Normalized confusion matrix

[[0.34146341 0.34146341 0.31707317]

[0.09473684 0.62105263 0.28421053]

[0.08683473 0.22689076 0.68627451]]





precision recall f1-score support

hate 0.16 0.34 0.22 41

offensive 0.25 0.62 0.36 95

neither 0.92 0.69 0.79 714

accuracy 0.66 850

macro avg 0.45 0.55 0.46 850

weighted avg 0.81 0.66 0.71 850

average eval\_loss: 1.3174042493635572, accuracy: 0.6623529411764706

training 16 epoch...

8.995550870895386 seconds used......

16 training finished! average train loss: 0.0024559391216219715

total clf loss: 0.08846426347736269 total adv loss: 0.0010091025067140436

evaluating...

Confusion matrix, without normalization

[[ 13 14 14]

[ 9 55 31]

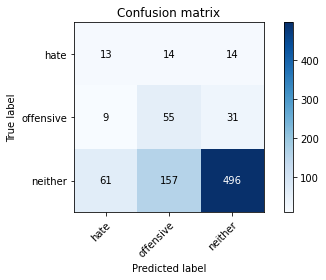
[ 61 157 496]]

Normalized confusion matrix

[[0.31707317 0.34146341 0.34146341]

[0.09473684 0.57894737 0.32631579]

[0.08543417 0.21988796 0.69467787]]





precision recall f1-score support

hate 0.16 0.32 0.21 41

offensive 0.24 0.58 0.34 95

neither 0.92 0.69 0.79 714

accuracy 0.66 850

macro avg 0.44 0.53 0.45 850

weighted avg 0.80 0.66 0.71 850

average eval\_loss: 1.2941894297634229, accuracy: 0.6635294117647059

training 17 epoch...

8.976407051086426 seconds used......

17 training finished! average train loss: 0.001929238747430241

total clf loss: 0.06947743595810607 total adv loss: 0.00049682979647514

evaluating...

Confusion matrix, without normalization

[[ 11 15 15]

[ 8 58 29]

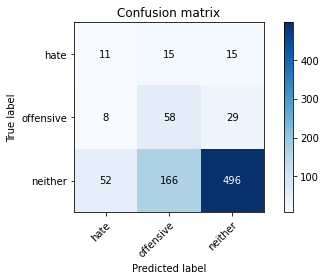
[ 52 166 496]]

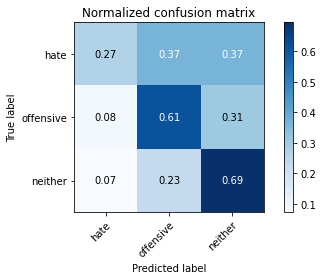
Normalized confusion matrix

[[0.26829268 0.36585366 0.36585366]

[0.08421053 0.61052632 0.30526316]

[0.07282913 0.232493 0.69467787]]





precision recall f1-score support

hate 0.15 0.27 0.20 41

offensive 0.24 0.61 0.35 95

neither 0.92 0.69 0.79 714

accuracy 0.66 850

macro avg 0.44 0.52 0.44 850

weighted avg 0.81 0.66 0.71 850

average eval\_loss: 1.3327233558222868, accuracy: 0.6647058823529411

training 18 epoch...

8.97204875946045 seconds used......

18 training finished! average train loss: 0.0015753261961637894

total clf loss: 0.05673396593192592 total adv loss: 0.0004444670146313001

evaluating...

Confusion matrix, without normalization

[[ 12 13 16]

[ 9 54 32]

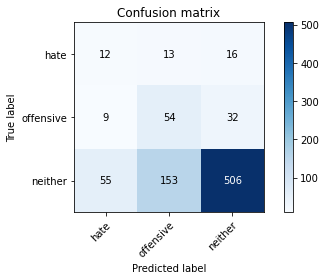
[ 55 153 506]]

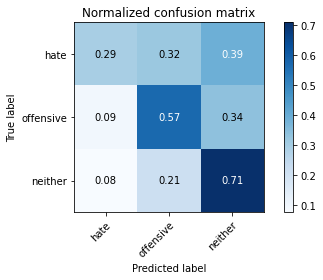
Normalized confusion matrix

[[0.29268293 0.31707317 0.3902439 ]

[0.09473684 0.56842105 0.33684211]

[0.07703081 0.21428571 0.70868347]]





precision recall f1-score support

hate 0.16 0.29 0.21 41

offensive 0.25 0.57 0.34 95

neither 0.91 0.71 0.80 714

accuracy 0.67 850

macro avg 0.44 0.52 0.45 850

weighted avg 0.80 0.67 0.72 850

average eval\_loss: 1.313151829066513, accuracy: 0.6729411764705883

training 19 epoch...

8.980314254760742 seconds used......

19 training finished! average train loss: 0.0013406014113570563

total clf loss: 0.04827196724363603 total adv loss: 0.0002063277742649916

evaluating...

Confusion matrix, without normalization

[[ 13 13 15]

[ 9 54 32]

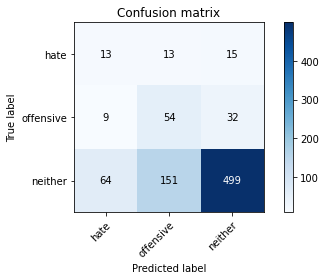
[ 64 151 499]]

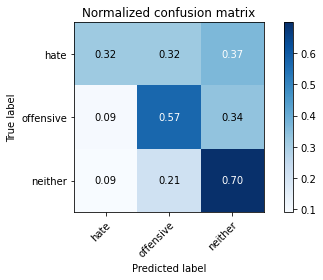
Normalized confusion matrix

[[0.31707317 0.31707317 0.36585366]

[0.09473684 0.56842105 0.33684211]

[0.08963585 0.21148459 0.69887955]]





precision recall f1-score support

hate 0.15 0.32 0.20 41

offensive 0.25 0.57 0.35 95

neither 0.91 0.70 0.79 714

accuracy 0.67 850

macro avg 0.44 0.53 0.45 850

weighted avg 0.80 0.67 0.71 850

average eval\_loss: 1.358160970398214, accuracy: 0.6658823529411765

training 20 epoch...

8.939250469207764 seconds used......

20 training finished! average train loss: 0.0012677944840915087

total clf loss: 0.045650331769138575 total adv loss: 0.00019459740059346586

evaluating...

Confusion matrix, without normalization

[[ 13 13 15]

[ 9 54 32]

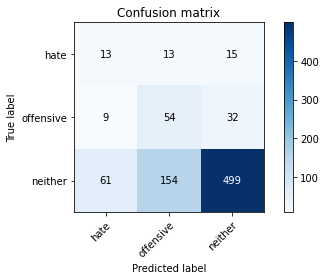
[ 61 154 499]]

Normalized confusion matrix

[[0.31707317 0.31707317 0.36585366]

[0.09473684 0.56842105 0.33684211]

[0.08543417 0.21568627 0.69887955]]





precision recall f1-score support

hate 0.16 0.32 0.21 41

offensive 0.24 0.57 0.34 95

neither 0.91 0.70 0.79 714

accuracy 0.67 850

macro avg 0.44 0.53 0.45 850

weighted avg 0.80 0.67 0.71 850

average eval\_loss: 1.3640875613308772, accuracy: 0.6658823529411765

2 run......

training 1 epoch...

9.01229453086853 seconds used......

1 training finished! average train loss: 1.0146093964576721

total clf loss: 37.75280958414078 total adv loss: 24.537421941757202

evaluating...

Confusion matrix, without normalization

[[ 0 31 10]

[ 0 20 75]

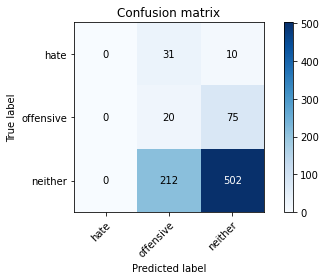
[ 0 212 502]]

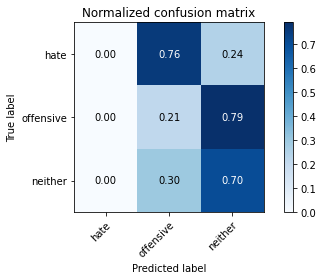
Normalized confusion matrix

[[0. 0.75609756 0.24390244]

[0. 0.21052632 0.78947368]

[0. 0.29691877 0.70308123]]





/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.08 0.21 0.11 95

neither 0.86 0.70 0.77 714

accuracy 0.61 850

macro avg 0.31 0.30 0.29 850

weighted avg 0.73 0.61 0.66 850

average eval\_loss: 0.820935094079306, accuracy: 0.6141176470588235

training 2 epoch...

8.910428047180176 seconds used......

2 training finished! average train loss: 0.8897926857074102

total clf loss: 33.136600345373154 total adv loss: 22.081277310848236

evaluating...

Confusion matrix, without normalization

[[ 0 25 16]

[ 0 36 59]

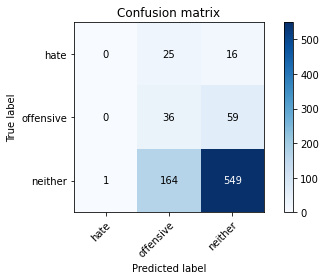
[ 1 164 549]]

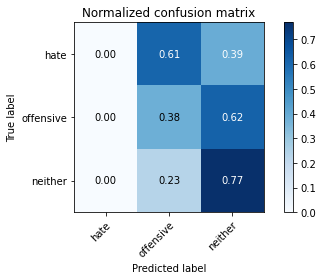
Normalized confusion matrix

[[0. 0.6097561 0.3902439 ]

[0. 0.37894737 0.62105263]

[0.00140056 0.22969188 0.76890756]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.16 0.38 0.23 95

neither 0.88 0.77 0.82 714

accuracy 0.69 850

macro avg 0.35 0.38 0.35 850

weighted avg 0.76 0.69 0.71 850

average eval\_loss: 0.7801905021417973, accuracy: 0.6882352941176471

training 3 epoch...

9.001877546310425 seconds used......

3 training finished! average train loss: 0.7923093934853872

total clf loss: 29.547801077365875 total adv loss: 20.493252009153366

evaluating...

Confusion matrix, without normalization

[[ 1 18 22]

[ 1 44 50]

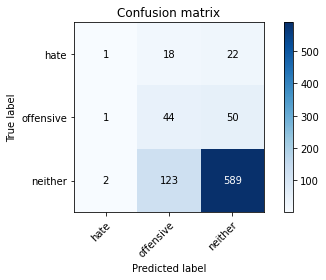
[ 2 123 589]]

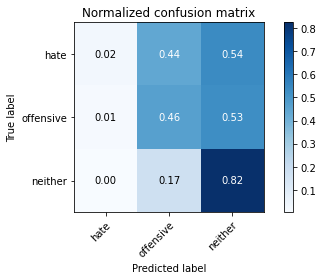
Normalized confusion matrix

[[0.02439024 0.43902439 0.53658537]

[0.01052632 0.46315789 0.52631579]

[0.00280112 0.17226891 0.82492997]]





precision recall f1-score support

hate 0.25 0.02 0.04 41

offensive 0.24 0.46 0.31 95

neither 0.89 0.82 0.86 714

accuracy 0.75 850

macro avg 0.46 0.44 0.41 850

weighted avg 0.79 0.75 0.76 850

average eval\_loss: 0.6622377152359763, accuracy: 0.7458823529411764

training 4 epoch...

8.952674150466919 seconds used......

4 training finished! average train loss: 0.6107284708155526

total clf loss: 22.749538749456406 total adv loss: 15.266278952360153

evaluating...

Confusion matrix, without normalization

[[ 13 14 14]

[ 16 54 25]

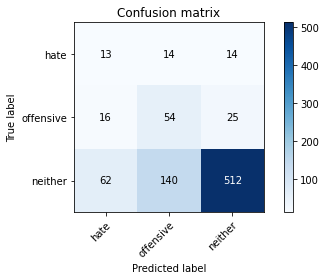
[ 62 140 512]]

Normalized confusion matrix

[[0.31707317 0.34146341 0.34146341]

[0.16842105 0.56842105 0.26315789]

[0.08683473 0.19607843 0.71708683]]





precision recall f1-score support

hate 0.14 0.32 0.20 41

offensive 0.26 0.57 0.36 95

neither 0.93 0.72 0.81 714

accuracy 0.68 850

macro avg 0.44 0.53 0.45 850

weighted avg 0.82 0.68 0.73 850

average eval\_loss: 0.7535028455908909, accuracy: 0.6811764705882353

training 5 epoch...

8.939473152160645 seconds used......

5 training finished! average train loss: 0.4218013895054658

total clf loss: 15.668171718716621 total adv loss: 9.66643351316452

evaluating...

Confusion matrix, without normalization

[[ 12 14 15]

[ 10 59 26]

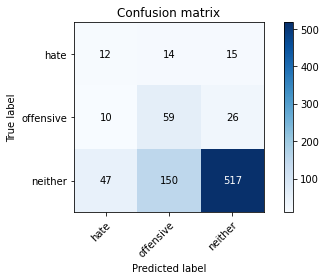
[ 47 150 517]]

Normalized confusion matrix

[[0.29268293 0.34146341 0.36585366]

[0.10526316 0.62105263 0.27368421]

[0.06582633 0.21008403 0.72408964]]





precision recall f1-score support

hate 0.17 0.29 0.22 41

offensive 0.26 0.62 0.37 95

neither 0.93 0.72 0.81 714

accuracy 0.69 850

macro avg 0.46 0.55 0.47 850

weighted avg 0.82 0.69 0.73 850

average eval\_loss: 0.7440318851623424, accuracy: 0.691764705882353

training 6 epoch...

8.967307806015015 seconds used......

6 training finished! average train loss: 0.23970773195226988

total clf loss: 8.884254090487957 total adv loss: 5.095514779444784

evaluating...

Confusion matrix, without normalization

[[ 3 27 11]

[ 2 75 18]

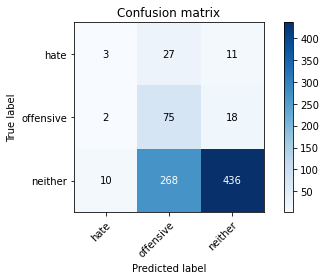
[ 10 268 436]]

Normalized confusion matrix

[[0.07317073 0.65853659 0.26829268]

[0.02105263 0.78947368 0.18947368]

[0.0140056 0.37535014 0.61064426]]





precision recall f1-score support

hate 0.20 0.07 0.11 41

offensive 0.20 0.79 0.32 95

neither 0.94 0.61 0.74 714

accuracy 0.60 850

macro avg 0.45 0.49 0.39 850

weighted avg 0.82 0.60 0.66 850

average eval\_loss: 1.096241690131814, accuracy: 0.6047058823529412

training 7 epoch...

8.935689926147461 seconds used......

7 training finished! average train loss: 0.12429017708119419

total clf loss: 4.543098896741867 total adv loss: 1.3730502520920709

evaluating...

Confusion matrix, without normalization

[[ 12 14 15]

[ 9 64 22]

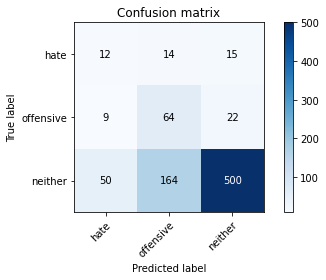
[ 50 164 500]]

Normalized confusion matrix

[[0.29268293 0.34146341 0.36585366]

[0.09473684 0.67368421 0.23157895]

[0.07002801 0.22969188 0.70028011]]





precision recall f1-score support

hate 0.17 0.29 0.21 41

offensive 0.26 0.67 0.38 95

neither 0.93 0.70 0.80 714

accuracy 0.68 850

macro avg 0.45 0.56 0.46 850

weighted avg 0.82 0.68 0.72 850

average eval\_loss: 0.9620315887705359, accuracy: 0.6776470588235294

training 8 epoch...

8.90807580947876 seconds used......

8 training finished! average train loss: 0.04188235378306773

total clf loss: 1.5149238361045718 total adv loss: 0.1431819586432539

evaluating...

Confusion matrix, without normalization

[[ 14 13 14]

[ 16 58 21]

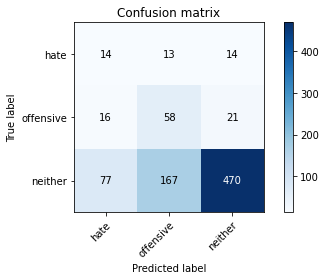
[ 77 167 470]]

Normalized confusion matrix

[[0.34146341 0.31707317 0.34146341]

[0.16842105 0.61052632 0.22105263]

[0.10784314 0.23389356 0.65826331]]





precision recall f1-score support

hate 0.13 0.34 0.19 41

offensive 0.24 0.61 0.35 95

neither 0.93 0.66 0.77 714

accuracy 0.64 850

macro avg 0.44 0.54 0.44 850

weighted avg 0.82 0.64 0.70 850

average eval\_loss: 1.2116608620989462, accuracy: 0.6376470588235295

training 9 epoch...

8.903123617172241 seconds used......

9 training finished! average train loss: 0.016864275857288804

total clf loss: 0.6113126347772777 total adv loss: 0.08397412481281208

evaluating...

Confusion matrix, without normalization

[[ 10 16 15]

[ 11 66 18]

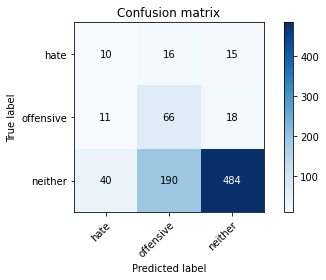
[ 40 190 484]]

Normalized confusion matrix

[[0.24390244 0.3902439 0.36585366]

[0.11578947 0.69473684 0.18947368]

[0.05602241 0.26610644 0.67787115]]





precision recall f1-score support

hate 0.16 0.24 0.20 41

offensive 0.24 0.69 0.36 95

neither 0.94 0.68 0.79 714

accuracy 0.66 850

macro avg 0.45 0.54 0.45 850

weighted avg 0.82 0.66 0.71 850

average eval\_loss: 1.3049531599051905, accuracy: 0.6588235294117647

training 10 epoch...

8.889697551727295 seconds used......

10 training finished! average train loss: 0.013248064723383222

total clf loss: 0.48066229932010174 total adv loss: 0.07463936645945068

evaluating...

Confusion matrix, without normalization

[[ 9 19 13]

[ 3 75 17]

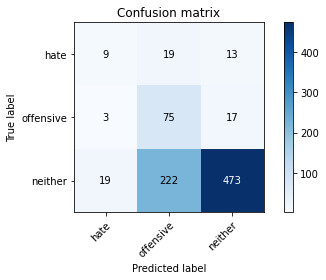
[ 19 222 473]]

Normalized confusion matrix

[[0.2195122 0.46341463 0.31707317]

[0.03157895 0.78947368 0.17894737]

[0.02661064 0.31092437 0.66246499]]





precision recall f1-score support

hate 0.29 0.22 0.25 41

offensive 0.24 0.79 0.36 95

neither 0.94 0.66 0.78 714

accuracy 0.66 850

macro avg 0.49 0.56 0.46 850

weighted avg 0.83 0.66 0.71 850

average eval\_loss: 1.441319541778242, accuracy: 0.6552941176470588

training 11 epoch...

8.988553047180176 seconds used......

11 training finished! average train loss: 0.011761012732879154

total clf loss: 0.42474057199433446 total adv loss: 0.026882312764428207

evaluating...

Confusion matrix, without normalization

[[ 14 11 16]

[ 19 52 24]

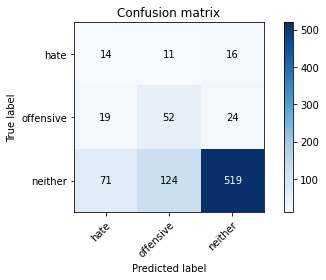
[ 71 124 519]]

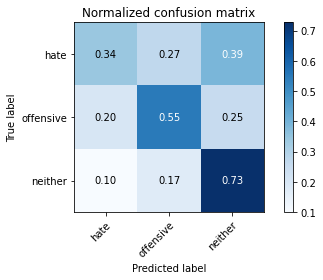
Normalized confusion matrix

[[0.34146341 0.26829268 0.3902439 ]

[0.2 0.54736842 0.25263158]

[0.09943978 0.17366947 0.72689076]]





precision recall f1-score support

hate 0.13 0.34 0.19 41

offensive 0.28 0.55 0.37 95

neither 0.93 0.73 0.82 714

accuracy 0.69 850

macro avg 0.45 0.54 0.46 850

weighted avg 0.82 0.69 0.74 850

average eval\_loss: 1.1957834315063909, accuracy: 0.6882352941176471

training 12 epoch...

8.943063259124756 seconds used......

12 training finished! average train loss: 0.006152845893262161

total clf loss: 0.22199884883593768 total adv loss: 0.009927891443567205

evaluating...

Confusion matrix, without normalization

[[ 12 15 14]

[ 13 67 15]

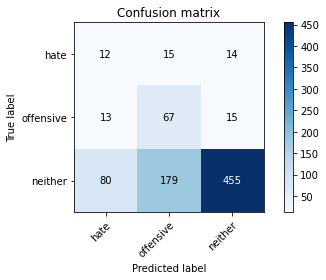
[ 80 179 455]]

Normalized confusion matrix

[[0.29268293 0.36585366 0.34146341]

[0.13684211 0.70526316 0.15789474]

[0.11204482 0.25070028 0.6372549 ]]





precision recall f1-score support

hate 0.11 0.29 0.16 41

offensive 0.26 0.71 0.38 95

neither 0.94 0.64 0.76 714

accuracy 0.63 850

macro avg 0.44 0.55 0.43 850

weighted avg 0.82 0.63 0.69 850

average eval\_loss: 1.5914302884254519, accuracy: 0.6282352941176471

training 13 epoch...

8.924218654632568 seconds used......

13 training finished! average train loss: 0.0028769084763351

total clf loss: 0.10374097968451679 total adv loss: 0.0034454932413154893

evaluating...

Confusion matrix, without normalization

[[ 9 17 15]

[ 4 69 22]

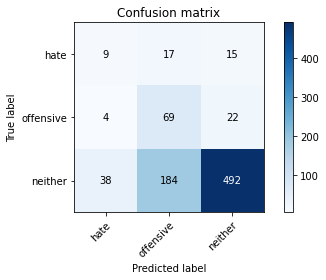
[ 38 184 492]]

Normalized confusion matrix

[[0.2195122 0.41463415 0.36585366]

[0.04210526 0.72631579 0.23157895]

[0.05322129 0.25770308 0.68907563]]





precision recall f1-score support

hate 0.18 0.22 0.20 41

offensive 0.26 0.73 0.38 95

neither 0.93 0.69 0.79 714

accuracy 0.67 850

macro avg 0.45 0.54 0.46 850

weighted avg 0.82 0.67 0.72 850

average eval\_loss: 1.4439324786491303, accuracy: 0.6705882352941176

training 14 epoch...

8.971091747283936 seconds used......

14 training finished! average train loss: 0.0023071046404462927

total clf loss: 0.08313076361082494 total adv loss: 0.0014999180534687184

evaluating...

Confusion matrix, without normalization

[[ 10 15 16]

[ 7 60 28]

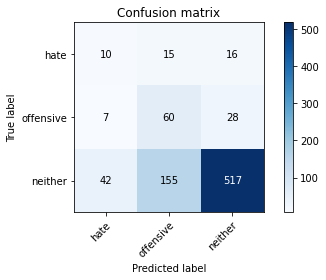
[ 42 155 517]]

Normalized confusion matrix

[[0.24390244 0.36585366 0.3902439 ]

[0.07368421 0.63157895 0.29473684]

[0.05882353 0.21708683 0.72408964]]





precision recall f1-score support

hate 0.17 0.24 0.20 41

offensive 0.26 0.63 0.37 95

neither 0.92 0.72 0.81 714

accuracy 0.69 850

macro avg 0.45 0.53 0.46 850

weighted avg 0.81 0.69 0.73 850

average eval\_loss: 1.2964315060959306, accuracy: 0.6905882352941176

training 15 epoch...

8.947010278701782 seconds used......

15 training finished! average train loss: 0.0020610151049267086

total clf loss: 0.0742549387505278 total adv loss: 0.001167893509773421

evaluating...

Confusion matrix, without normalization

[[ 11 14 16]

[ 9 60 26]

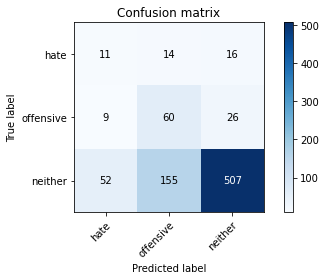
[ 52 155 507]]

Normalized confusion matrix

[[0.26829268 0.34146341 0.3902439 ]

[0.09473684 0.63157895 0.27368421]

[0.07282913 0.21708683 0.71008403]]





precision recall f1-score support

hate 0.15 0.27 0.19 41

offensive 0.26 0.63 0.37 95

neither 0.92 0.71 0.80 714

accuracy 0.68 850

macro avg 0.45 0.54 0.46 850

weighted avg 0.81 0.68 0.73 850

average eval\_loss: 1.3884425392905033, accuracy: 0.68

training 16 epoch...

8.945465087890625 seconds used......

16 training finished! average train loss: 0.0016335551352110794

total clf loss: 0.05884068162413314 total adv loss: 0.0006539381884067552

evaluating...

Confusion matrix, without normalization

[[ 11 14 16]

[ 10 60 25]

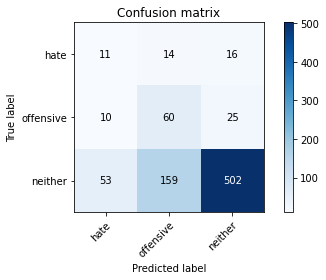
[ 53 159 502]]

Normalized confusion matrix

[[0.26829268 0.34146341 0.3902439 ]

[0.10526316 0.63157895 0.26315789]

[0.07422969 0.22268908 0.70308123]]





precision recall f1-score support

hate 0.15 0.27 0.19 41

offensive 0.26 0.63 0.37 95

neither 0.92 0.70 0.80 714

accuracy 0.67 850

macro avg 0.44 0.53 0.45 850

weighted avg 0.81 0.67 0.72 850

average eval\_loss: 1.475861047811503, accuracy: 0.6741176470588235

training 17 epoch...

8.905315637588501 seconds used......

17 training finished! average train loss: 0.0015857628895901144

total clf loss: 0.057162772631272674 total adv loss: 0.0015061798205806554

evaluating...

Confusion matrix, without normalization

[[ 10 15 16]

[ 8 63 24]

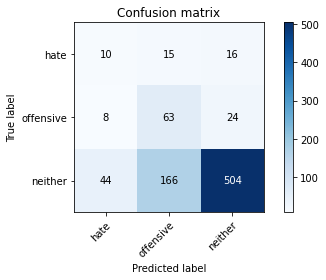
[ 44 166 504]]

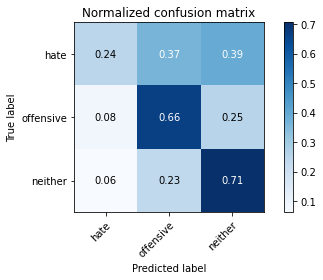
Normalized confusion matrix

[[0.24390244 0.36585366 0.3902439 ]

[0.08421053 0.66315789 0.25263158]

[0.06162465 0.232493 0.70588235]]





precision recall f1-score support

hate 0.16 0.24 0.19 41

offensive 0.26 0.66 0.37 95

neither 0.93 0.71 0.80 714

accuracy 0.68 850

macro avg 0.45 0.54 0.46 850

weighted avg 0.81 0.68 0.72 850

average eval\_loss: 1.504653572264604, accuracy: 0.6788235294117647

training 18 epoch...

8.896197080612183 seconds used......

18 training finished! average train loss: 0.0013109150564155749

total clf loss: 0.0472264050913509 total adv loss: 0.0006692576951081719

evaluating...

Confusion matrix, without normalization

[[ 10 15 16]

[ 8 61 26]

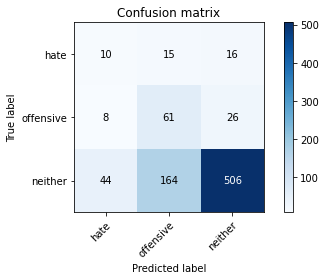
[ 44 164 506]]

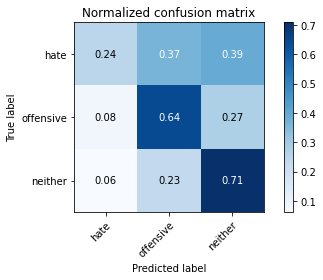
Normalized confusion matrix

[[0.24390244 0.36585366 0.3902439 ]

[0.08421053 0.64210526 0.27368421]

[0.06162465 0.22969188 0.70868347]]





precision recall f1-score support

hate 0.16 0.24 0.19 41

offensive 0.25 0.64 0.36 95

neither 0.92 0.71 0.80 714

accuracy 0.68 850

macro avg 0.45 0.53 0.45 850

weighted avg 0.81 0.68 0.72 850

average eval\_loss: 1.4910771581966993, accuracy: 0.6788235294117647

training 19 epoch...

8.984904527664185 seconds used......

19 training finished! average train loss: 0.0011123155782115646

total clf loss: 0.04005870217224583 total adv loss: 0.0003068265340857579

evaluating...

Confusion matrix, without normalization

[[ 9 16 16]

[ 8 62 25]

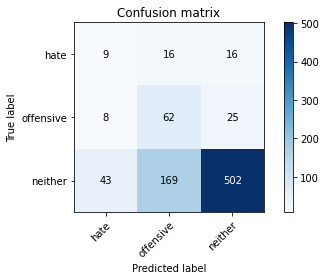
[ 43 169 502]]

Normalized confusion matrix

[[0.2195122 0.3902439 0.3902439 ]

[0.08421053 0.65263158 0.26315789]

[0.06022409 0.23669468 0.70308123]]





precision recall f1-score support

hate 0.15 0.22 0.18 41

offensive 0.25 0.65 0.36 95

neither 0.92 0.70 0.80 714

accuracy 0.67 850

macro avg 0.44 0.53 0.45 850

weighted avg 0.81 0.67 0.72 850

average eval\_loss: 1.5408048942674297, accuracy: 0.6741176470588235

training 20 epoch...

8.922697305679321 seconds used......

20 training finished! average train loss: 0.0010340671271680752

total clf loss: 0.037238541641272604 total adv loss: 0.0002424975381245531

evaluating...

Confusion matrix, without normalization

[[ 9 16 16]

[ 7 66 22]

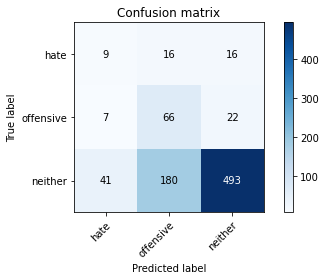
[ 41 180 493]]

Normalized confusion matrix

[[0.2195122 0.3902439 0.3902439 ]

[0.07368421 0.69473684 0.23157895]

[0.05742297 0.25210084 0.69047619]]





precision recall f1-score support

hate 0.16 0.22 0.18 41

offensive 0.25 0.69 0.37 95

neither 0.93 0.69 0.79 714

accuracy 0.67 850

macro avg 0.45 0.53 0.45 850

weighted avg 0.82 0.67 0.72 850

average eval\_loss: 1.6026675633106078, accuracy: 0.668235294117647

3 run......

training 1 epoch...

9.013583421707153 seconds used......

1 training finished! average train loss: 1.0495071427689657

total clf loss: 38.983895778656006 total adv loss: 24.032775580883026

evaluating...

Confusion matrix, without normalization

[[ 0 25 16]

[ 0 14 81]

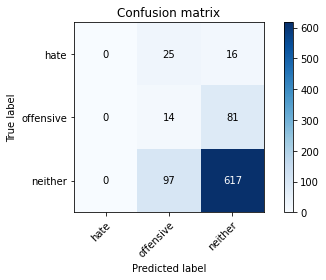
[ 0 97 617]]

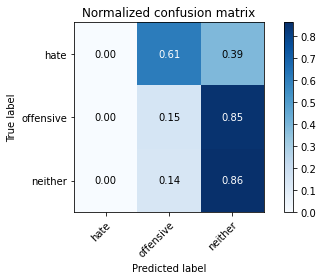
Normalized confusion matrix

[[0. 0.6097561 0.3902439 ]

[0. 0.14736842 0.85263158]

[0. 0.13585434 0.86414566]]





/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.10 0.15 0.12 95

neither 0.86 0.86 0.86 714

accuracy 0.74 850

macro avg 0.32 0.34 0.33 850

weighted avg 0.74 0.74 0.74 850

average eval\_loss: 0.7613911649515462, accuracy: 0.7423529411764705

training 2 epoch...

8.909711599349976 seconds used......

2 training finished! average train loss: 0.8906061848004659

total clf loss: 33.16070258617401 total adv loss: 21.977601140737534

evaluating...

Confusion matrix, without normalization

[[ 1 20 20]

[ 0 14 81]

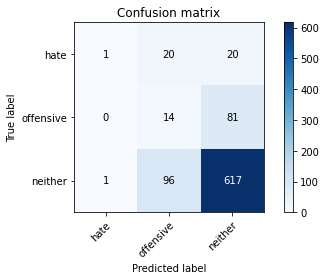
[ 1 96 617]]

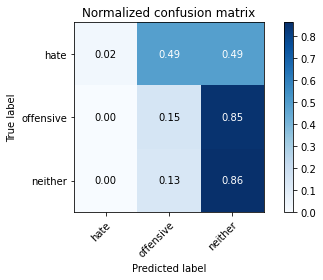
Normalized confusion matrix

[[0.02439024 0.48780488 0.48780488]

[0. 0.14736842 0.85263158]

[0.00140056 0.13445378 0.86414566]]





precision recall f1-score support

hate 0.50 0.02 0.05 41

offensive 0.11 0.15 0.12 95

neither 0.86 0.86 0.86 714

accuracy 0.74 850

macro avg 0.49 0.35 0.34 850

weighted avg 0.76 0.74 0.74 850

average eval\_loss: 0.7550587141236593, accuracy: 0.7435294117647059

training 3 epoch...

8.913017988204956 seconds used......

3 training finished! average train loss: 0.7732924024264017

total clf loss: 28.812710613012314 total adv loss: 19.483684569597244

evaluating...

Confusion matrix, without normalization

[[ 1 34 6]

[ 0 81 14]

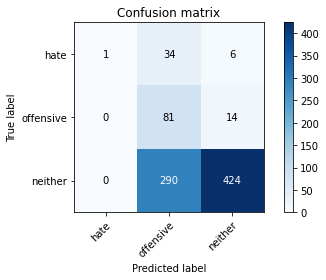
[ 0 290 424]]

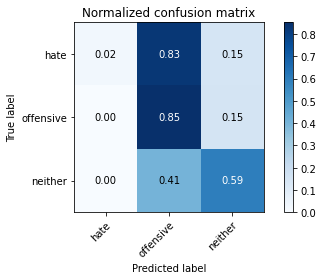
Normalized confusion matrix

[[0.02439024 0.82926829 0.14634146]

[0. 0.85263158 0.14736842]

[0. 0.40616246 0.59383754]]





precision recall f1-score support

hate 1.00 0.02 0.05 41

offensive 0.20 0.85 0.32 95

neither 0.95 0.59 0.73 714

accuracy 0.60 850

macro avg 0.72 0.49 0.37 850

weighted avg 0.87 0.60 0.65 850

average eval\_loss: 0.8456629095382469, accuracy: 0.5952941176470589

training 4 epoch...

8.942427635192871 seconds used......

4 training finished! average train loss: 0.6031463435954518

total clf loss: 22.436147302389145 total adv loss: 14.457580000162125

evaluating...

Confusion matrix, without normalization

[[ 8 16 17]

[ 8 61 26]

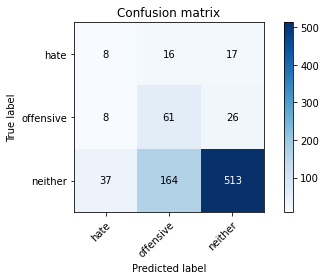
[ 37 164 513]]

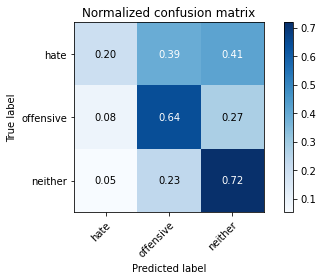
Normalized confusion matrix

[[0.19512195 0.3902439 0.41463415]

[0.08421053 0.64210526 0.27368421]

[0.05182073 0.22969188 0.71848739]]





precision recall f1-score support

hate 0.15 0.20 0.17 41

offensive 0.25 0.64 0.36 95

neither 0.92 0.72 0.81 714

accuracy 0.68 850

macro avg 0.44 0.52 0.45 850

weighted avg 0.81 0.68 0.73 850

average eval\_loss: 0.7296679493819558, accuracy: 0.6847058823529412

training 5 epoch...

8.872147560119629 seconds used......

5 training finished! average train loss: 0.4087040225664775

total clf loss: 15.23022748529911 total adv loss: 10.337651789188385

evaluating...

Confusion matrix, without normalization

[[ 12 13 16]

[ 13 50 32]

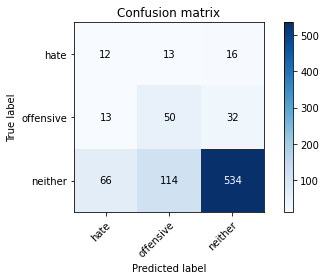
[ 66 114 534]]

Normalized confusion matrix

[[0.29268293 0.31707317 0.3902439 ]

[0.13684211 0.52631579 0.33684211]

[0.09243697 0.15966387 0.74789916]]





precision recall f1-score support

hate 0.13 0.29 0.18 41

offensive 0.28 0.53 0.37 95

neither 0.92 0.75 0.82 714

accuracy 0.70 850

macro avg 0.44 0.52 0.46 850

weighted avg 0.81 0.70 0.74 850

average eval\_loss: 0.7431706555274337, accuracy: 0.7011764705882353

training 6 epoch...

8.909358978271484 seconds used......

6 training finished! average train loss: 0.2328501045703888

total clf loss: 8.658659417182207 total adv loss: 5.521111832000315

evaluating...

Confusion matrix, without normalization

[[ 11 17 13]

[ 12 53 30]

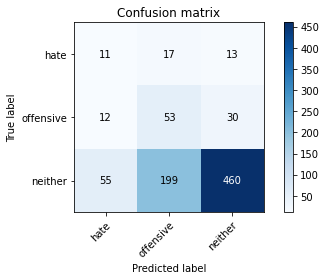
[ 55 199 460]]

Normalized confusion matrix

[[0.26829268 0.41463415 0.31707317]

[0.12631579 0.55789474 0.31578947]

[0.07703081 0.27871148 0.6442577 ]]





precision recall f1-score support

hate 0.14 0.27 0.18 41

offensive 0.20 0.56 0.29 95

neither 0.91 0.64 0.76 714

accuracy 0.62 850

macro avg 0.42 0.49 0.41 850

weighted avg 0.80 0.62 0.68 850

average eval\_loss: 1.0290185303189034, accuracy: 0.6164705882352941

training 7 epoch...

8.937682151794434 seconds used......

7 training finished! average train loss: 0.09897242222602169

total clf loss: 3.641544731333852 total adv loss: 1.5707507426850498

evaluating...

Confusion matrix, without normalization

[[ 10 23 8]

[ 7 78 10]

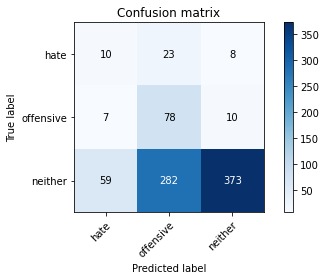
[ 59 282 373]]

Normalized confusion matrix

[[0.24390244 0.56097561 0.19512195]

[0.07368421 0.82105263 0.10526316]

[0.08263305 0.39495798 0.52240896]]





precision recall f1-score support

hate 0.13 0.24 0.17 41

offensive 0.20 0.82 0.33 95

neither 0.95 0.52 0.68 714

accuracy 0.54 850

macro avg 0.43 0.53 0.39 850

weighted avg 0.83 0.54 0.61 850

average eval\_loss: 1.5874407148465168, accuracy: 0.5423529411764706

training 8 epoch...

8.914092540740967 seconds used......

8 training finished! average train loss: 0.046987166783461966

total clf loss: 1.7357761589810252 total adv loss: 0.8847631701282808

evaluating...

Confusion matrix, without normalization

[[ 10 19 12]

[ 8 75 12]

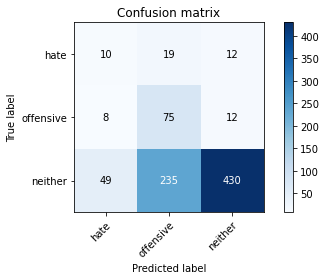
[ 49 235 430]]

Normalized confusion matrix

[[0.24390244 0.46341463 0.29268293]

[0.08421053 0.78947368 0.12631579]

[0.06862745 0.32913165 0.6022409 ]]





precision recall f1-score support

hate 0.15 0.24 0.19 41

offensive 0.23 0.79 0.35 95

neither 0.95 0.60 0.74 714

accuracy 0.61 850

macro avg 0.44 0.55 0.43 850

weighted avg 0.83 0.61 0.67 850

average eval\_loss: 1.4391989317441056, accuracy: 0.6058823529411764

training 9 epoch...

8.943803548812866 seconds used......

9 training finished! average train loss: 0.02601419938986914

total clf loss: 0.9783417452126741 total adv loss: 0.8366113854281139

evaluating...

Confusion matrix, without normalization

[[ 9 20 12]

[ 7 65 23]

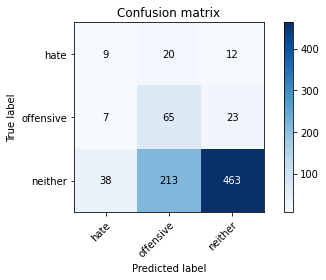
[ 38 213 463]]

Normalized confusion matrix

[[0.2195122 0.48780488 0.29268293]

[0.07368421 0.68421053 0.24210526]

[0.05322129 0.29831933 0.64845938]]





precision recall f1-score support

hate 0.17 0.22 0.19 41

offensive 0.22 0.68 0.33 95

neither 0.93 0.65 0.76 714

accuracy 0.63 850

macro avg 0.44 0.52 0.43 850

weighted avg 0.81 0.63 0.69 850

average eval\_loss: 1.3850578350491476, accuracy: 0.6317647058823529

training 10 epoch...

8.91066598892212 seconds used......

10 training finished! average train loss: 0.010819358854658075

total clf loss: 0.3928463214542717 total adv loss: 0.0669880841014674

evaluating...

Confusion matrix, without normalization

[[ 9 17 15]

[ 8 66 21]

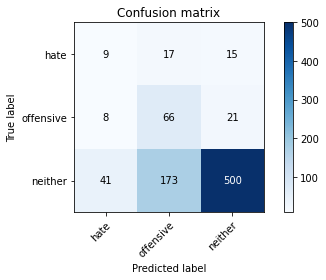
[ 41 173 500]]

Normalized confusion matrix

[[0.2195122 0.41463415 0.36585366]

[0.08421053 0.69473684 0.22105263]

[0.05742297 0.24229692 0.70028011]]





precision recall f1-score support

hate 0.16 0.22 0.18 41

offensive 0.26 0.69 0.38 95

neither 0.93 0.70 0.80 714

accuracy 0.68 850

macro avg 0.45 0.54 0.45 850

weighted avg 0.82 0.68 0.72 850

average eval\_loss: 1.254122859294808, accuracy: 0.6764705882352942

training 11 epoch...

8.886844873428345 seconds used......

11 training finished! average train loss: 0.006108253463430123

total clf loss: 0.2207372966222465 total adv loss: 0.016803451639134437

evaluating...

Confusion matrix, without normalization

[[ 10 15 16]

[ 11 69 15]

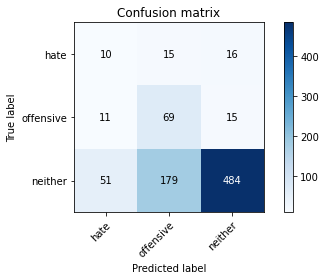
[ 51 179 484]]

Normalized confusion matrix

[[0.24390244 0.36585366 0.3902439 ]

[0.11578947 0.72631579 0.15789474]

[0.07142857 0.25070028 0.67787115]]





precision recall f1-score support

hate 0.14 0.24 0.18 41

offensive 0.26 0.73 0.39 95

neither 0.94 0.68 0.79 714

accuracy 0.66 850

macro avg 0.45 0.55 0.45 850

weighted avg 0.83 0.66 0.71 850

average eval\_loss: 1.3792306138527428, accuracy: 0.6623529411764706

training 12 epoch...

8.962741374969482 seconds used......

12 training finished! average train loss: 0.004557667497364391

total clf loss: 0.16443264728877693 total adv loss: 0.007132324642952881

evaluating...

Confusion matrix, without normalization

[[ 10 16 15]

[ 10 70 15]

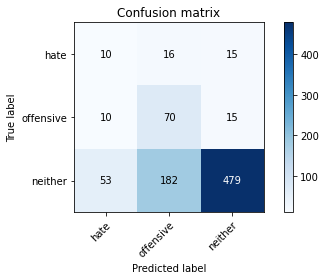
[ 53 182 479]]

Normalized confusion matrix

[[0.24390244 0.3902439 0.36585366]

[0.10526316 0.73684211 0.15789474]

[0.07422969 0.25490196 0.67086835]]





precision recall f1-score support

hate 0.14 0.24 0.18 41

offensive 0.26 0.74 0.39 95

neither 0.94 0.67 0.78 714

accuracy 0.66 850

macro avg 0.45 0.55 0.45 850

weighted avg 0.83 0.66 0.71 850

average eval\_loss: 1.460581439330654, accuracy: 0.6576470588235294

training 13 epoch...

8.913010835647583 seconds used......

13 training finished! average train loss: 0.005818891038264458

total clf loss: 0.21234507928602397 total adv loss: 0.057300034757645335

evaluating...

Confusion matrix, without normalization

[[ 9 18 14]

[ 10 70 15]

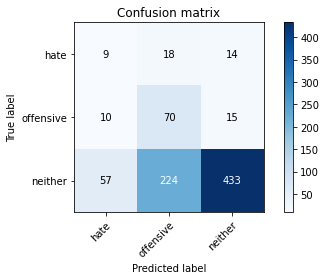
[ 57 224 433]]

Normalized confusion matrix

[[0.2195122 0.43902439 0.34146341]

[0.10526316 0.73684211 0.15789474]

[0.07983193 0.31372549 0.60644258]]





precision recall f1-score support

hate 0.12 0.22 0.15 41

offensive 0.22 0.74 0.34 95

neither 0.94 0.61 0.74 714

accuracy 0.60 850

macro avg 0.43 0.52 0.41 850

weighted avg 0.82 0.60 0.66 850

average eval\_loss: 1.8253158844877468, accuracy: 0.6023529411764705

training 14 epoch...

8.96118688583374 seconds used......

14 training finished! average train loss: 0.0051002489344682544

total clf loss: 0.18450711760669947 total adv loss: 0.017963105498893128

evaluating...

Confusion matrix, without normalization

[[ 9 18 14]

[ 7 77 11]

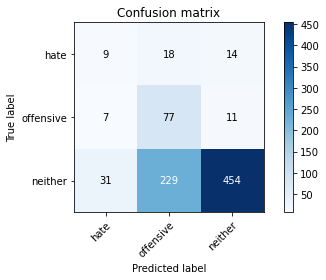
[ 31 229 454]]

Normalized confusion matrix

[[0.2195122 0.43902439 0.34146341]

[0.07368421 0.81052632 0.11578947]

[0.04341737 0.32072829 0.63585434]]





precision recall f1-score support

hate 0.19 0.22 0.20 41

offensive 0.24 0.81 0.37 95

neither 0.95 0.64 0.76 714

accuracy 0.64 850

macro avg 0.46 0.56 0.44 850

weighted avg 0.83 0.64 0.69 850

average eval\_loss: 1.7187188384691687, accuracy: 0.6352941176470588

training 15 epoch...

8.913506031036377 seconds used......

15 training finished! average train loss: 0.0022753692125358307

total clf loss: 0.082054658443667 total adv loss: 0.002827323950441496

evaluating...

Confusion matrix, without normalization

[[ 9 14 18]

[ 7 69 19]

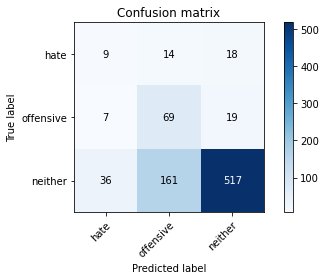
[ 36 161 517]]

Normalized confusion matrix

[[0.2195122 0.34146341 0.43902439]

[0.07368421 0.72631579 0.2 ]

[0.05042017 0.2254902 0.72408964]]





precision recall f1-score support

hate 0.17 0.22 0.19 41

offensive 0.28 0.73 0.41 95

neither 0.93 0.72 0.82 714

accuracy 0.70 850

macro avg 0.46 0.56 0.47 850

weighted avg 0.82 0.70 0.74 850

average eval\_loss: 1.3517471812259363, accuracy: 0.7

training 16 epoch...

8.905560731887817 seconds used......

16 training finished! average train loss: 0.0016860005853231996

total clf loss: 0.060748615593183786 total adv loss: 0.001051895542332204

evaluating...

Confusion matrix, without normalization

[[ 9 18 14]

[ 8 73 14]

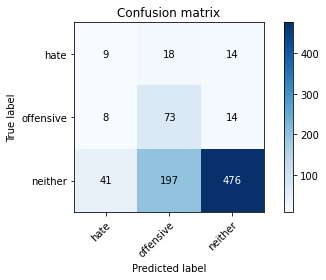
[ 41 197 476]]

Normalized confusion matrix

[[0.2195122 0.43902439 0.34146341]

[0.08421053 0.76842105 0.14736842]

[0.05742297 0.27591036 0.66666667]]





precision recall f1-score support

hate 0.16 0.22 0.18 41

offensive 0.25 0.77 0.38 95

neither 0.94 0.67 0.78 714

accuracy 0.66 850

macro avg 0.45 0.55 0.45 850

weighted avg 0.83 0.66 0.71 850

average eval\_loss: 1.6254103771508346, accuracy: 0.6564705882352941

training 17 epoch...

8.963082075119019 seconds used......

17 training finished! average train loss: 0.0016419382688278954

total clf loss: 0.059156996197998524 total adv loss: 0.0009443527595749401

evaluating...

Confusion matrix, without normalization

[[ 9 17 15]

[ 7 71 17]

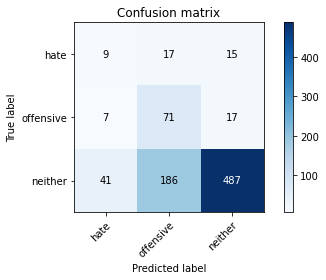
[ 41 186 487]]

Normalized confusion matrix

[[0.2195122 0.41463415 0.36585366]

[0.07368421 0.74736842 0.17894737]

[0.05742297 0.2605042 0.68207283]]





precision recall f1-score support

hate 0.16 0.22 0.18 41

offensive 0.26 0.75 0.38 95

neither 0.94 0.68 0.79 714

accuracy 0.67 850

macro avg 0.45 0.55 0.45 850

weighted avg 0.82 0.67 0.72 850

average eval\_loss: 1.569903337832562, accuracy: 0.6670588235294118

training 18 epoch...

8.88854718208313 seconds used......

18 training finished! average train loss: 0.0015593005357206697

total clf loss: 0.056178543833084404 total adv loss: 0.0008744877025037567

evaluating...

Confusion matrix, without normalization

[[ 10 15 16]

[ 9 68 18]

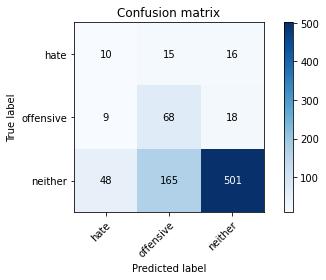
[ 48 165 501]]

Normalized confusion matrix

[[0.24390244 0.36585366 0.3902439 ]

[0.09473684 0.71578947 0.18947368]

[0.06722689 0.23109244 0.70168067]]





precision recall f1-score support

hate 0.15 0.24 0.19 41

offensive 0.27 0.72 0.40 95

neither 0.94 0.70 0.80 714

accuracy 0.68 850

macro avg 0.45 0.55 0.46 850

weighted avg 0.82 0.68 0.73 850

average eval\_loss: 1.5062015759596212, accuracy: 0.6811764705882353

training 19 epoch...

8.930191040039062 seconds used......

19 training finished! average train loss: 0.001212395493186907

total clf loss: 0.043665502715157345 total adv loss: 0.0003853025981470637

evaluating...

Confusion matrix, without normalization

[[ 9 17 15]

[ 7 71 17]

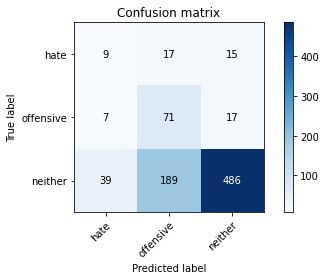
[ 39 189 486]]

Normalized confusion matrix

[[0.2195122 0.41463415 0.36585366]

[0.07368421 0.74736842 0.17894737]

[0.05462185 0.26470588 0.68067227]]





precision recall f1-score support

hate 0.16 0.22 0.19 41

offensive 0.26 0.75 0.38 95

neither 0.94 0.68 0.79 714

accuracy 0.67 850

macro avg 0.45 0.55 0.45 850

weighted avg 0.82 0.67 0.71 850

average eval\_loss: 1.5999756704888652, accuracy: 0.6658823529411765

training 20 epoch...

8.91045355796814 seconds used......

20 training finished! average train loss: 0.001176082498406888

total clf loss: 0.04235429939581081 total adv loss: 0.0003065842852265632

evaluating...

Confusion matrix, without normalization

[[ 9 16 16]

[ 9 69 17]

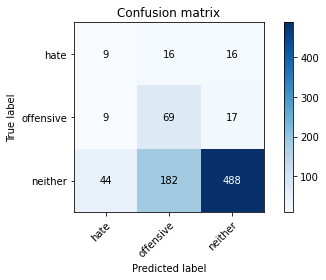
[ 44 182 488]]

Normalized confusion matrix

[[0.2195122 0.3902439 0.3902439 ]

[0.09473684 0.72631579 0.17894737]

[0.06162465 0.25490196 0.68347339]]





precision recall f1-score support

hate 0.15 0.22 0.17 41

offensive 0.26 0.73 0.38 95

neither 0.94 0.68 0.79 714

accuracy 0.67 850

macro avg 0.45 0.54 0.45 850

weighted avg 0.82 0.67 0.71 850

average eval\_loss: 1.5971033962790009, accuracy: 0.6658823529411765

4 run......

training 1 epoch...

8.923470735549927 seconds used......

1 training finished! average train loss: 1.083516110976537

total clf loss: 40.2148619890213 total adv loss: 24.16563844680786

evaluating...

Confusion matrix, without normalization

[[ 0 40 1]

[ 0 67 28]

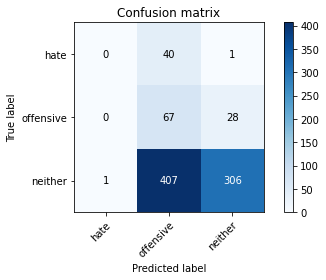
[ 1 407 306]]

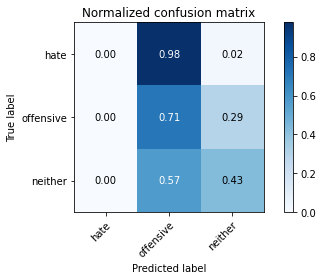
Normalized confusion matrix

[[0. 0.97560976 0.02439024]

[0. 0.70526316 0.29473684]

[0.00140056 0.57002801 0.42857143]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.13 0.71 0.22 95

neither 0.91 0.43 0.58 714

accuracy 0.44 850

macro avg 0.35 0.38 0.27 850

weighted avg 0.78 0.44 0.51 850

average eval\_loss: 0.993002267770989, accuracy: 0.4388235294117647

training 2 epoch...

8.94726848602295 seconds used......

2 training finished! average train loss: 0.9123108320766025

total clf loss: 33.99223220348358 total adv loss: 22.980841487646103

evaluating...

Confusion matrix, without normalization

[[ 0 36 5]

[ 0 77 18]

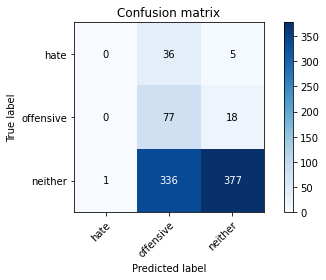
[ 1 336 377]]

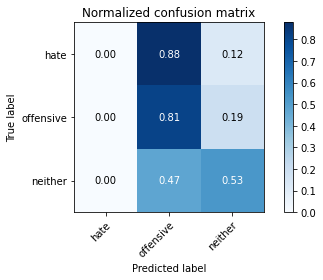
Normalized confusion matrix

[[0. 0.87804878 0.12195122]

[0. 0.81052632 0.18947368]

[0.00140056 0.47058824 0.5280112 ]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.17 0.81 0.28 95

neither 0.94 0.53 0.68 714

accuracy 0.53 850

macro avg 0.37 0.45 0.32 850

weighted avg 0.81 0.53 0.60 850

average eval\_loss: 0.8812555320734201, accuracy: 0.5341176470588235

training 3 epoch...

8.91312861442566 seconds used......

3 training finished! average train loss: 0.7966118719842699

total clf loss: 29.627075999975204 total adv loss: 18.980974048376083

evaluating...

Confusion matrix, without normalization

[[ 3 25 13]

[ 2 67 26]

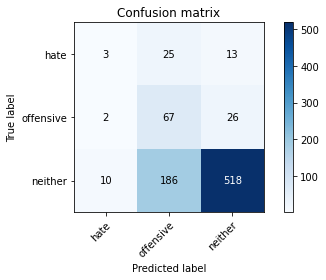
[ 10 186 518]]

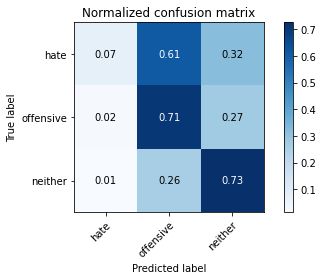
Normalized confusion matrix

[[0.07317073 0.6097561 0.31707317]

[0.02105263 0.70526316 0.27368421]

[0.0140056 0.2605042 0.7254902 ]]





precision recall f1-score support

hate 0.20 0.07 0.11 41

offensive 0.24 0.71 0.36 95

neither 0.93 0.73 0.82 714

accuracy 0.69 850

macro avg 0.46 0.50 0.43 850

weighted avg 0.82 0.69 0.73 850

average eval\_loss: 0.7308533299107884, accuracy: 0.691764705882353

training 4 epoch...

8.986391544342041 seconds used......

4 training finished! average train loss: 0.6451803868015608

total clf loss: 24.0890254676342 total adv loss: 17.25063292682171

evaluating...

Confusion matrix, without normalization

[[ 4 26 11]

[ 2 81 12]

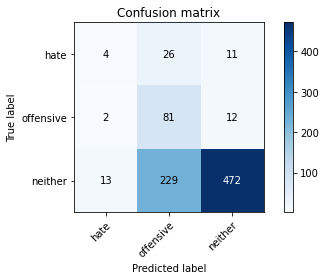
[ 13 229 472]]

Normalized confusion matrix

[[0.09756098 0.63414634 0.26829268]

[0.02105263 0.85263158 0.12631579]

[0.01820728 0.32072829 0.66106443]]





precision recall f1-score support

hate 0.21 0.10 0.13 41

offensive 0.24 0.85 0.38 95

neither 0.95 0.66 0.78 714

accuracy 0.66 850

macro avg 0.47 0.54 0.43 850

weighted avg 0.84 0.66 0.70 850

average eval\_loss: 0.7637627611673156, accuracy: 0.6552941176470588

training 5 epoch...

8.918492794036865 seconds used......

5 training finished! average train loss: 0.5007212621470293

total clf loss: 18.726106017827988 total adv loss: 14.002812415361404

evaluating...

Confusion matrix, without normalization

[[ 14 17 10]

[ 11 65 19]

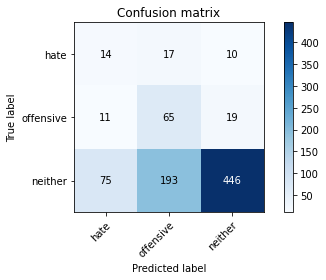
[ 75 193 446]]

Normalized confusion matrix

[[0.34146341 0.41463415 0.24390244]

[0.11578947 0.68421053 0.2 ]

[0.10504202 0.27030812 0.62464986]]





precision recall f1-score support

hate 0.14 0.34 0.20 41

offensive 0.24 0.68 0.35 95

neither 0.94 0.62 0.75 714

accuracy 0.62 850

macro avg 0.44 0.55 0.43 850

weighted avg 0.82 0.62 0.68 850

average eval\_loss: 0.8700481571776922, accuracy: 0.6176470588235294

training 6 epoch...

8.90714168548584 seconds used......

6 training finished! average train loss: 0.29640575498342514

total clf loss: 11.011507272720337 total adv loss: 6.81800140440464

evaluating...

Confusion matrix, without normalization

[[ 10 12 19]

[ 3 56 36]

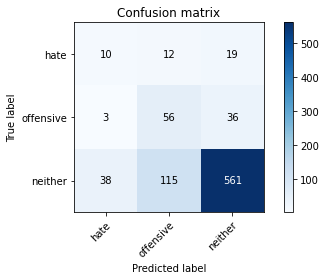
[ 38 115 561]]

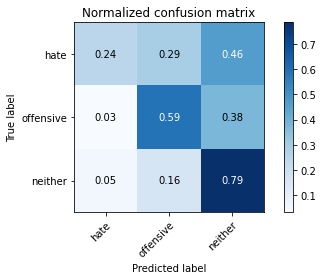
Normalized confusion matrix

[[0.24390244 0.29268293 0.46341463]

[0.03157895 0.58947368 0.37894737]

[0.05322129 0.16106443 0.78571429]]





precision recall f1-score support

hate 0.20 0.24 0.22 41

offensive 0.31 0.59 0.40 95

neither 0.91 0.79 0.84 714

accuracy 0.74 850

macro avg 0.47 0.54 0.49 850

weighted avg 0.81 0.74 0.76 850

average eval\_loss: 0.6398649886900256, accuracy: 0.7376470588235294

training 7 epoch...

8.986894607543945 seconds used......

7 training finished! average train loss: 0.16285781303627622

total clf loss: 6.01183508336544 total adv loss: 2.979076591320336

evaluating...

Confusion matrix, without normalization

[[ 20 8 13]

[ 17 47 31]

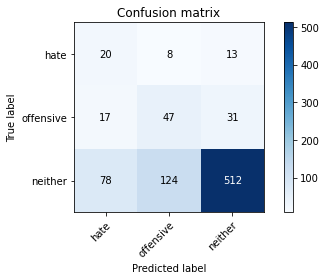
[ 78 124 512]]

Normalized confusion matrix

[[0.48780488 0.19512195 0.31707317]

[0.17894737 0.49473684 0.32631579]

[0.1092437 0.17366947 0.71708683]]





precision recall f1-score support

hate 0.17 0.49 0.26 41

offensive 0.26 0.49 0.34 95

neither 0.92 0.72 0.81 714

accuracy 0.68 850

macro avg 0.45 0.57 0.47 850

weighted avg 0.81 0.68 0.73 850

average eval\_loss: 0.8924962013180173, accuracy: 0.6811764705882353

training 8 epoch...

8.92441201210022 seconds used......

8 training finished! average train loss: 0.05651845322507951

total clf loss: 2.0596034340560436 total adv loss: 0.49878250499023125

evaluating...

Confusion matrix, without normalization

[[ 10 20 11]

[ 4 58 33]

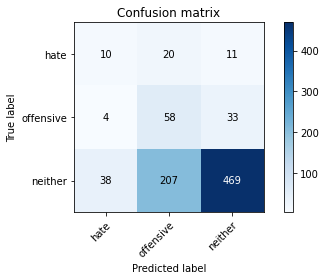
[ 38 207 469]]

Normalized confusion matrix

[[0.24390244 0.48780488 0.26829268]

[0.04210526 0.61052632 0.34736842]

[0.05322129 0.28991597 0.65686275]]





precision recall f1-score support

hate 0.19 0.24 0.22 41

offensive 0.20 0.61 0.31 95

neither 0.91 0.66 0.76 714

accuracy 0.63 850

macro avg 0.44 0.50 0.43 850

weighted avg 0.80 0.63 0.69 850

average eval\_loss: 1.2012044944648825, accuracy: 0.6317647058823529

training 9 epoch...

8.894875049591064 seconds used......

9 training finished! average train loss: 0.03151026241377824

total clf loss: 1.1433760565705597 total adv loss: 0.18013213937956607

evaluating...

Confusion matrix, without normalization

[[ 15 14 12]

[ 9 55 31]

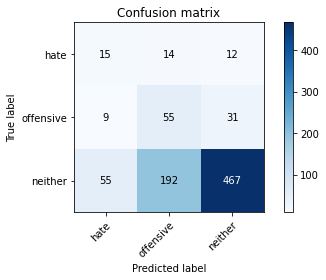
[ 55 192 467]]

Normalized confusion matrix

[[0.36585366 0.34146341 0.29268293]

[0.09473684 0.57894737 0.32631579]

[0.07703081 0.26890756 0.65406162]]





precision recall f1-score support

hate 0.19 0.37 0.25 41

offensive 0.21 0.58 0.31 95

neither 0.92 0.65 0.76 714

accuracy 0.63 850

macro avg 0.44 0.53 0.44 850

weighted avg 0.80 0.63 0.69 850

average eval\_loss: 1.2911258079653043, accuracy: 0.6317647058823529

training 10 epoch...

8.90371060371399 seconds used......

10 training finished! average train loss: 0.011761706238353832

total clf loss: 0.4242022961843759 total adv loss: 0.015617410881532123

evaluating...

Confusion matrix, without normalization

[[ 16 16 9]

[ 10 63 22]

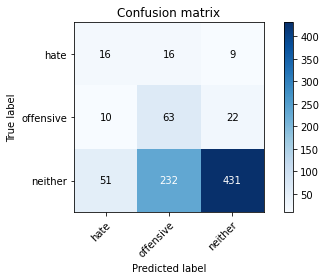
[ 51 232 431]]

Normalized confusion matrix

[[0.3902439 0.3902439 0.2195122 ]

[0.10526316 0.66315789 0.23157895]

[0.07142857 0.32492997 0.60364146]]





precision recall f1-score support

hate 0.21 0.39 0.27 41

offensive 0.20 0.66 0.31 95

neither 0.93 0.60 0.73 714

accuracy 0.60 850

macro avg 0.45 0.55 0.44 850

weighted avg 0.82 0.60 0.66 850

average eval\_loss: 1.5524418156182522, accuracy: 0.6

training 11 epoch...

8.903096199035645 seconds used......

11 training finished! average train loss: 0.00743481439874611

total clf loss: 0.27474803826771677 total adv loss: 0.14189438185167091

evaluating...

Confusion matrix, without normalization

[[ 14 16 11]

[ 9 58 28]

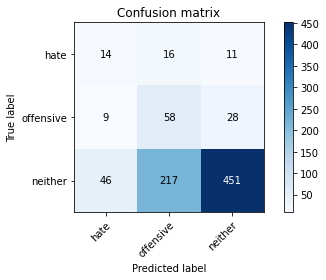
[ 46 217 451]]

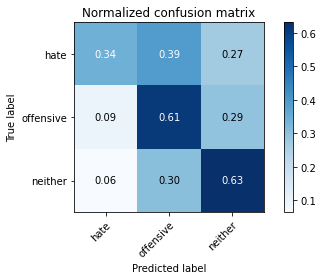
Normalized confusion matrix

[[0.34146341 0.3902439 0.26829268]

[0.09473684 0.61052632 0.29473684]

[0.06442577 0.30392157 0.63165266]]





precision recall f1-score support

hate 0.20 0.34 0.25 41

offensive 0.20 0.61 0.30 95

neither 0.92 0.63 0.75 714

accuracy 0.62 850

macro avg 0.44 0.53 0.43 850

weighted avg 0.81 0.62 0.68 850

average eval\_loss: 1.4882253840305777, accuracy: 0.6152941176470588

training 12 epoch...

8.996561765670776 seconds used......

12 training finished! average train loss: 0.005349037418555882

total clf loss: 0.19553537759929895 total adv loss: 0.059400633484074206

evaluating...

Confusion matrix, without normalization

[[ 14 18 9]

[ 9 66 20]

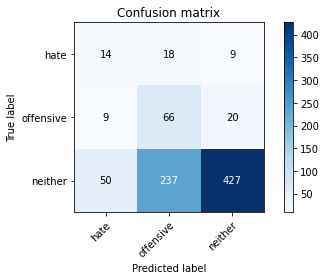
[ 50 237 427]]

Normalized confusion matrix

[[0.34146341 0.43902439 0.2195122 ]

[0.09473684 0.69473684 0.21052632]

[0.07002801 0.33193277 0.59803922]]





precision recall f1-score support

hate 0.19 0.34 0.25 41

offensive 0.21 0.69 0.32 95

neither 0.94 0.60 0.73 714

accuracy 0.60 850

macro avg 0.44 0.54 0.43 850

weighted avg 0.82 0.60 0.66 850

average eval\_loss: 1.6806532465419624, accuracy: 0.5964705882352941

training 13 epoch...

8.959847927093506 seconds used......

13 training finished! average train loss: 0.004262541886419058

total clf loss: 0.15387233602814376 total adv loss: 0.008416553391725756

evaluating...

Confusion matrix, without normalization

[[ 12 18 11]

[ 9 60 26]

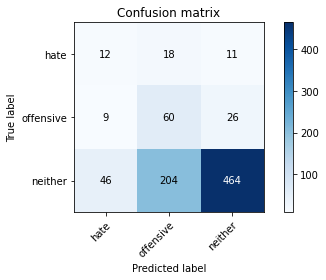
[ 46 204 464]]

Normalized confusion matrix

[[0.29268293 0.43902439 0.26829268]

[0.09473684 0.63157895 0.27368421]

[0.06442577 0.28571429 0.64985994]]





precision recall f1-score support

hate 0.18 0.29 0.22 41

offensive 0.21 0.63 0.32 95

neither 0.93 0.65 0.76 714

accuracy 0.63 850

macro avg 0.44 0.52 0.43 850

weighted avg 0.81 0.63 0.69 850

average eval\_loss: 1.5216946222895193, accuracy: 0.6305882352941177

training 14 epoch...

8.945127964019775 seconds used......

14 training finished! average train loss: 0.003199455709869249

total clf loss: 0.11538323451532051 total adv loss: 0.004056588120874949

evaluating...

Confusion matrix, without normalization

[[ 16 14 11]

[ 10 58 27]

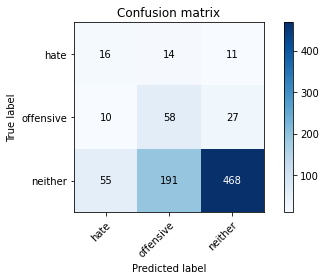
[ 55 191 468]]

Normalized confusion matrix

[[0.3902439 0.34146341 0.26829268]

[0.10526316 0.61052632 0.28421053]

[0.07703081 0.267507 0.65546218]]





precision recall f1-score support

hate 0.20 0.39 0.26 41

offensive 0.22 0.61 0.32 95

neither 0.92 0.66 0.77 714

accuracy 0.64 850

macro avg 0.45 0.55 0.45 850

weighted avg 0.81 0.64 0.69 850

average eval\_loss: 1.5342612850295683, accuracy: 0.6376470588235295

training 15 epoch...

8.93045449256897 seconds used......

15 training finished! average train loss: 0.003031829032503689

total clf loss: 0.10945028997957706 total adv loss: 0.006088895055199828

evaluating...

Confusion matrix, without normalization

[[ 15 15 11]

[ 10 56 29]

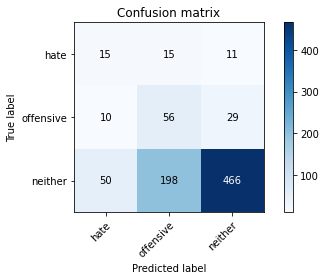
[ 50 198 466]]

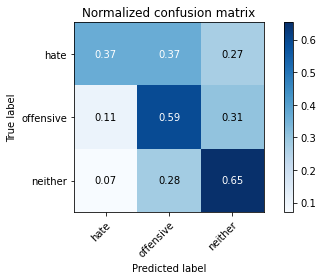
Normalized confusion matrix

[[0.36585366 0.36585366 0.26829268]

[0.10526316 0.58947368 0.30526316]

[0.07002801 0.27731092 0.65266106]]





precision recall f1-score support

hate 0.20 0.37 0.26 41

offensive 0.21 0.59 0.31 95

neither 0.92 0.65 0.76 714

accuracy 0.63 850

macro avg 0.44 0.54 0.44 850

weighted avg 0.81 0.63 0.69 850

average eval\_loss: 1.591673337805713, accuracy: 0.6317647058823529

training 16 epoch...

8.928400754928589 seconds used......

16 training finished! average train loss: 0.0021733033516082084

total clf loss: 0.07831304974388331 total adv loss: 0.0014825846783423913

evaluating...

Confusion matrix, without normalization

[[ 11 20 10]

[ 8 63 24]

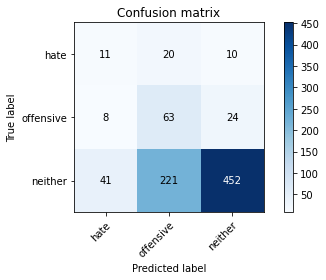
[ 41 221 452]]

Normalized confusion matrix

[[0.26829268 0.48780488 0.24390244]

[0.08421053 0.66315789 0.25263158]

[0.05742297 0.30952381 0.63305322]]





precision recall f1-score support

hate 0.18 0.27 0.22 41

offensive 0.21 0.66 0.32 95

neither 0.93 0.63 0.75 714

accuracy 0.62 850

macro avg 0.44 0.52 0.43 850

weighted avg 0.81 0.62 0.68 850

average eval\_loss: 1.7066297302574873, accuracy: 0.6188235294117647

training 17 epoch...

8.991243600845337 seconds used......

17 training finished! average train loss: 0.0022495858364790264

total clf loss: 0.08105658716522157 total adv loss: 0.001429943855100646

evaluating...

Confusion matrix, without normalization

[[ 12 19 10]

[ 8 64 23]

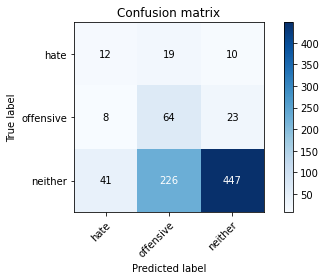
[ 41 226 447]]

Normalized confusion matrix

[[0.29268293 0.46341463 0.24390244]

[0.08421053 0.67368421 0.24210526]

[0.05742297 0.31652661 0.62605042]]





precision recall f1-score support

hate 0.20 0.29 0.24 41

offensive 0.21 0.67 0.32 95

neither 0.93 0.63 0.75 714

accuracy 0.62 850

macro avg 0.45 0.53 0.43 850

weighted avg 0.81 0.62 0.68 850

average eval\_loss: 1.7485588786197652, accuracy: 0.6152941176470588

training 18 epoch...

8.939120054244995 seconds used......

18 training finished! average train loss: 0.0015487103680950692

total clf loss: 0.05579627421684563 total adv loss: 0.0008540125490981154

evaluating...

Confusion matrix, without normalization

[[ 15 16 10]

[ 10 63 22]

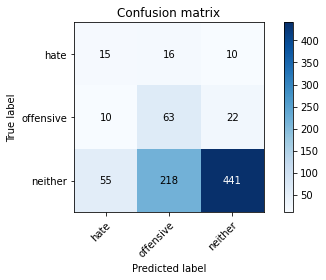
[ 55 218 441]]

Normalized confusion matrix

[[0.36585366 0.3902439 0.24390244]

[0.10526316 0.66315789 0.23157895]

[0.07703081 0.30532213 0.61764706]]





precision recall f1-score support

hate 0.19 0.37 0.25 41

offensive 0.21 0.66 0.32 95

neither 0.93 0.62 0.74 714

accuracy 0.61 850

macro avg 0.44 0.55 0.44 850

weighted avg 0.82 0.61 0.67 850

average eval\_loss: 1.7996753908715448, accuracy: 0.6105882352941177

training 19 epoch...

8.895154237747192 seconds used......

19 training finished! average train loss: 0.0016212931206812048

total clf loss: 0.05841066408902407 total adv loss: 0.0008822360364320048

evaluating...

Confusion matrix, without normalization

[[ 16 16 9]

[ 12 63 20]

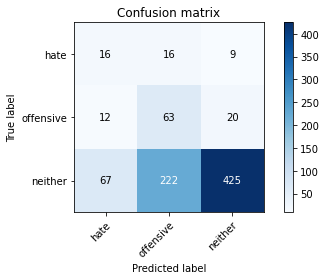
[ 67 222 425]]

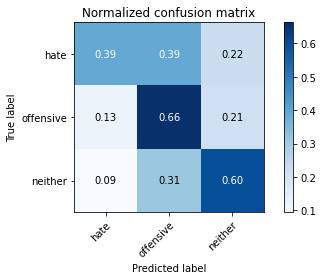
Normalized confusion matrix

[[0.3902439 0.3902439 0.2195122 ]

[0.12631579 0.66315789 0.21052632]

[0.09383754 0.31092437 0.5952381 ]]





precision recall f1-score support

hate 0.17 0.39 0.24 41

offensive 0.21 0.66 0.32 95

neither 0.94 0.60 0.73 714

accuracy 0.59 850

macro avg 0.44 0.55 0.43 850

weighted avg 0.82 0.59 0.66 850

average eval\_loss: 1.9455869351351245, accuracy: 0.5929411764705882

training 20 epoch...

8.915183067321777 seconds used......

20 training finished! average train loss: 0.0013590133951058509

total clf loss: 0.048943218134809285 total adv loss: 0.00037471599864602467

evaluating...

Confusion matrix, without normalization

[[ 15 15 11]

[ 10 54 31]

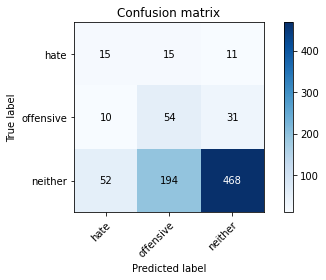
[ 52 194 468]]

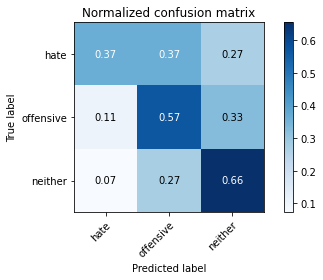
Normalized confusion matrix

[[0.36585366 0.36585366 0.26829268]

[0.10526316 0.56842105 0.32631579]

[0.07282913 0.27170868 0.65546218]]





precision recall f1-score support

hate 0.19 0.37 0.25 41

offensive 0.21 0.57 0.30 95

neither 0.92 0.66 0.76 714

accuracy 0.63 850

macro avg 0.44 0.53 0.44 850

weighted avg 0.80 0.63 0.69 850

average eval\_loss: 1.6797817823542065, accuracy: 0.6317647058823529

5 run......

training 1 epoch...

8.923255681991577 seconds used......

1 training finished! average train loss: 1.0778062757518556

total clf loss: 39.96199369430542 total adv loss: 23.21935847401619

evaluating...

Confusion matrix, without normalization

[[ 0 34 7]

[ 0 28 67]

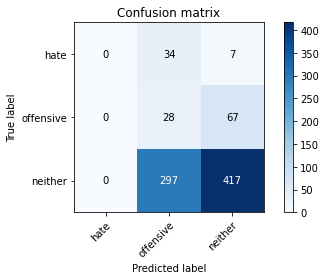
[ 0 297 417]]

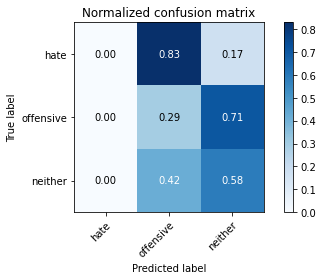
Normalized confusion matrix

[[0. 0.82926829 0.17073171]

[0. 0.29473684 0.70526316]

[0. 0.41596639 0.58403361]]





/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.08 0.29 0.12 95

neither 0.85 0.58 0.69 714

accuracy 0.52 850

macro avg 0.31 0.29 0.27 850

weighted avg 0.72 0.52 0.60 850

average eval\_loss: 0.9462103067442428, accuracy: 0.5235294117647059

training 2 epoch...

8.986348390579224 seconds used......

2 training finished! average train loss: 0.9271473437547684

total clf loss: 34.42682492733002 total adv loss: 20.990414172410965

evaluating...

Confusion matrix, without normalization

[[ 0 35 6]

[ 0 71 24]

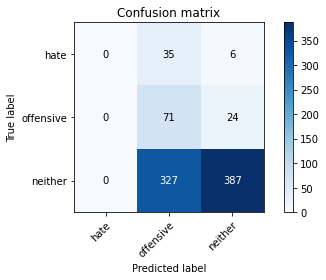
[ 0 327 387]]

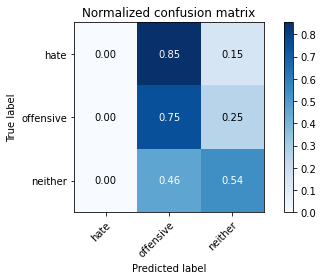
Normalized confusion matrix

[[0. 0.85365854 0.14634146]

[0. 0.74736842 0.25263158]

[0. 0.45798319 0.54201681]]





/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

/opt/conda/lib/python3.7/site-packages/sklearn/metrics/\_classification.py:1245: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.16 0.75 0.27 95

neither 0.93 0.54 0.68 714

accuracy 0.54 850

macro avg 0.36 0.43 0.32 850

weighted avg 0.80 0.54 0.60 850

average eval\_loss: 0.9231292182622954, accuracy: 0.5388235294117647

training 3 epoch...

8.956287145614624 seconds used......

3 training finished! average train loss: 0.8163481752077738

total clf loss: 30.389025688171387 total adv loss: 20.009830951690674

evaluating...

Confusion matrix, without normalization

[[ 0 36 5]

[ 3 84 8]

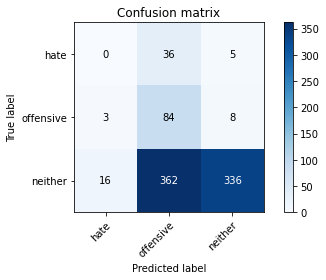
[ 16 362 336]]

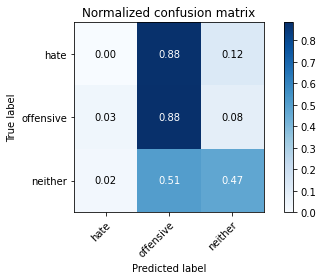
Normalized confusion matrix

[[0. 0.87804878 0.12195122]

[0.03157895 0.88421053 0.08421053]

[0.02240896 0.5070028 0.47058824]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.17 0.88 0.29 95

neither 0.96 0.47 0.63 714

accuracy 0.49 850

macro avg 0.38 0.45 0.31 850

weighted avg 0.83 0.49 0.56 850

average eval\_loss: 0.9788093414417532, accuracy: 0.49411764705882355

training 4 epoch...

8.98755693435669 seconds used......

4 training finished! average train loss: 0.6724182325932715

total clf loss: 25.234736025333405 total adv loss: 20.553596526384354

evaluating...

Confusion matrix, without normalization

[[ 4 33 4]

[ 6 82 7]

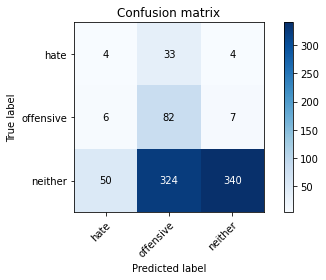
[ 50 324 340]]

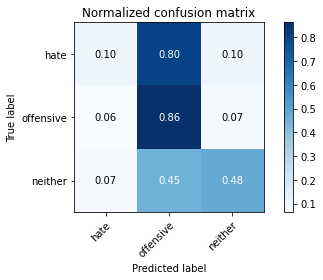
Normalized confusion matrix

[[0.09756098 0.80487805 0.09756098]

[0.06315789 0.86315789 0.07368421]

[0.07002801 0.45378151 0.47619048]]





precision recall f1-score support

hate 0.07 0.10 0.08 41

offensive 0.19 0.86 0.31 95

neither 0.97 0.48 0.64 714

accuracy 0.50 850

macro avg 0.41 0.48 0.34 850

weighted avg 0.84 0.50 0.57 850

average eval\_loss: 1.0105423715918562, accuracy: 0.5011764705882353

training 5 epoch...

8.935197830200195 seconds used......

5 training finished! average train loss: 0.4645971291595035

total clf loss: 17.258277982473373 total adv loss: 10.655626073479652

evaluating...

Confusion matrix, without normalization

[[ 4 23 14]

[ 4 80 11]

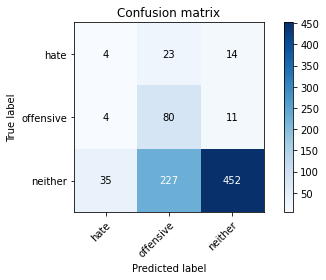
[ 35 227 452]]

Normalized confusion matrix

[[0.09756098 0.56097561 0.34146341]

[0.04210526 0.84210526 0.11578947]

[0.04901961 0.31792717 0.63305322]]





precision recall f1-score support

hate 0.09 0.10 0.10 41

offensive 0.24 0.84 0.38 95

neither 0.95 0.63 0.76 714

accuracy 0.63 850

macro avg 0.43 0.52 0.41 850

weighted avg 0.83 0.63 0.68 850

average eval\_loss: 0.843408313081708, accuracy: 0.6305882352941177

training 6 epoch...

8.922797679901123 seconds used......

6 training finished! average train loss: 0.29461781308054924

total clf loss: 10.953178480267525 total adv loss: 6.938743881881237

evaluating...

Confusion matrix, without normalization

[[ 7 16 18]

[ 8 69 18]

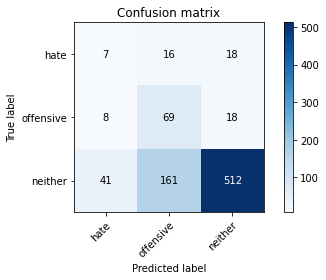
[ 41 161 512]]

Normalized confusion matrix

[[0.17073171 0.3902439 0.43902439]

[0.08421053 0.72631579 0.18947368]

[0.05742297 0.2254902 0.71708683]]





precision recall f1-score support

hate 0.12 0.17 0.14 41

offensive 0.28 0.73 0.40 95

neither 0.93 0.72 0.81 714

accuracy 0.69 850

macro avg 0.45 0.54 0.45 850

weighted avg 0.82 0.69 0.73 850

average eval\_loss: 0.7496654221484828, accuracy: 0.691764705882353

training 7 epoch...

8.939224481582642 seconds used......

7 training finished! average train loss: 0.14560101429621378

total clf loss: 5.359092555940151 total adv loss: 2.34912101412192

evaluating...

Confusion matrix, without normalization

[[ 6 24 11]

[ 7 73 15]

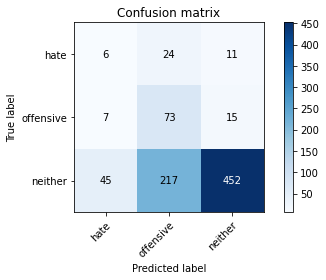
[ 45 217 452]]

Normalized confusion matrix

[[0.14634146 0.58536585 0.26829268]

[0.07368421 0.76842105 0.15789474]

[0.06302521 0.30392157 0.63305322]]





precision recall f1-score support

hate 0.10 0.15 0.12 41

offensive 0.23 0.77 0.36 95

neither 0.95 0.63 0.76 714

accuracy 0.62 850

macro avg 0.43 0.52 0.41 850

weighted avg 0.83 0.62 0.68 850

average eval\_loss: 1.0685414921405703, accuracy: 0.6247058823529412

training 8 epoch...

8.898464679718018 seconds used......

8 training finished! average train loss: 0.05583898200550013

total clf loss: 2.0362054128199816 total adv loss: 0.5200415817671455

evaluating...

Confusion matrix, without normalization

[[ 9 22 10]

[ 9 67 19]

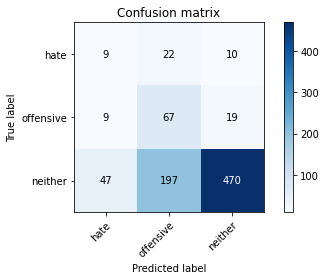
[ 47 197 470]]

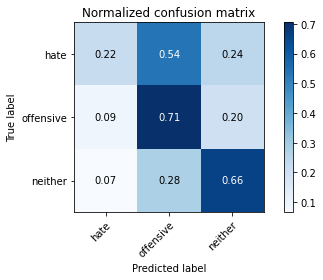
Normalized confusion matrix

[[0.2195122 0.53658537 0.24390244]

[0.09473684 0.70526316 0.2 ]

[0.06582633 0.27591036 0.65826331]]





precision recall f1-score support

hate 0.14 0.22 0.17 41

offensive 0.23 0.71 0.35 95

neither 0.94 0.66 0.77 714

accuracy 0.64 850

macro avg 0.44 0.53 0.43 850

weighted avg 0.82 0.64 0.70 850

average eval\_loss: 1.1737253775950087, accuracy: 0.6423529411764706

training 9 epoch...

8.967592716217041 seconds used......

9 training finished! average train loss: 0.024802848826059036

total clf loss: 0.9036054434254766 total adv loss: 0.2140575608909785

evaluating...

Confusion matrix, without normalization

[[ 17 16 8]

[ 21 63 11]

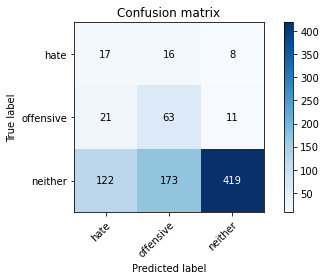
[122 173 419]]

Normalized confusion matrix

[[0.41463415 0.3902439 0.19512195]

[0.22105263 0.66315789 0.11578947]

[0.17086835 0.24229692 0.58683473]]





precision recall f1-score support

hate 0.11 0.41 0.17 41

offensive 0.25 0.66 0.36 95

neither 0.96 0.59 0.73 714

accuracy 0.59 850

macro avg 0.44 0.55 0.42 850

weighted avg 0.84 0.59 0.66 850

average eval\_loss: 1.3733385885350926, accuracy: 0.5870588235294117

training 10 epoch...

8.943658113479614 seconds used......

10 training finished! average train loss: 0.013065964914858341

total clf loss: 0.47269418369978666 total adv loss: 0.04638897864788305

evaluating...

Confusion matrix, without normalization

[[ 10 19 12]

[ 13 65 17]

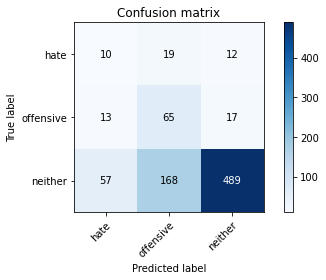
[ 57 168 489]]

Normalized confusion matrix

[[0.24390244 0.46341463 0.29268293]

[0.13684211 0.68421053 0.17894737]

[0.07983193 0.23529412 0.68487395]]





precision recall f1-score support

hate 0.12 0.24 0.17 41

offensive 0.26 0.68 0.37 95

neither 0.94 0.68 0.79 714

accuracy 0.66 850

macro avg 0.44 0.54 0.44 850

weighted avg 0.83 0.66 0.72 850

average eval\_loss: 1.19971375002764, accuracy: 0.6635294117647059

training 11 epoch...

8.942253112792969 seconds used......

11 training finished! average train loss: 0.007871828423554285

total clf loss: 0.28405826503876597 total adv loss: 0.013448847016661603

evaluating...

Confusion matrix, without normalization

[[ 11 19 11]

[ 12 66 17]

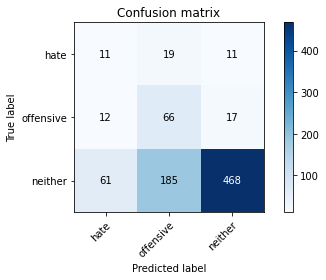
[ 61 185 468]]

Normalized confusion matrix

[[0.26829268 0.46341463 0.26829268]

[0.12631579 0.69473684 0.17894737]

[0.08543417 0.25910364 0.65546218]]





precision recall f1-score support

hate 0.13 0.27 0.18 41

offensive 0.24 0.69 0.36 95

neither 0.94 0.66 0.77 714

accuracy 0.64 850

macro avg 0.44 0.54 0.44 850

weighted avg 0.83 0.64 0.70 850

average eval\_loss: 1.3521610107272863, accuracy: 0.6411764705882353

training 12 epoch...

8.996700763702393 seconds used......

12 training finished! average train loss: 0.0058843122274589

total clf loss: 0.21205084543908015 total adv loss: 0.004312102338531076

evaluating...

Confusion matrix, without normalization

[[ 11 19 11]

[ 11 66 18]

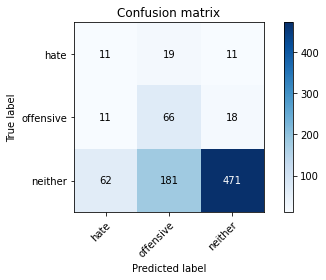
[ 62 181 471]]

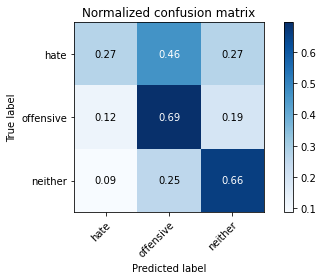
Normalized confusion matrix

[[0.26829268 0.46341463 0.26829268]

[0.11578947 0.69473684 0.18947368]

[0.08683473 0.2535014 0.65966387]]





precision recall f1-score support

hate 0.13 0.27 0.18 41

offensive 0.25 0.69 0.37 95

neither 0.94 0.66 0.78 714

accuracy 0.64 850

macro avg 0.44 0.54 0.44 850

weighted avg 0.83 0.64 0.70 850

average eval\_loss: 1.403278457044169, accuracy: 0.6447058823529411

training 13 epoch...

8.932039499282837 seconds used......

13 training finished! average train loss: 0.003879864434970336

total clf loss: 0.139771637506783 total adv loss: 0.0019303776603010192

evaluating...

Confusion matrix, without normalization

[[ 10 21 10]

[ 11 71 13]

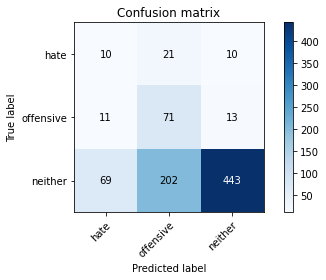
[ 69 202 443]]

Normalized confusion matrix

[[0.24390244 0.51219512 0.24390244]

[0.11578947 0.74736842 0.13684211]

[0.09663866 0.28291317 0.62044818]]





precision recall f1-score support

hate 0.11 0.24 0.15 41

offensive 0.24 0.75 0.37 95

neither 0.95 0.62 0.75 714

accuracy 0.62 850

macro avg 0.43 0.54 0.42 850

weighted avg 0.83 0.62 0.68 850

average eval\_loss: 1.620745146690413, accuracy: 0.6164705882352941

training 14 epoch...

8.883931159973145 seconds used......

14 training finished! average train loss: 0.0032728378977885265

total clf loss: 0.11785679939202964 total adv loss: 0.0006926966556193292

evaluating...

Confusion matrix, without normalization

[[ 10 20 11]

[ 11 68 16]

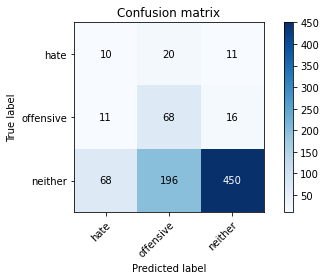
[ 68 196 450]]

Normalized confusion matrix

[[0.24390244 0.48780488 0.26829268]

[0.11578947 0.71578947 0.16842105]

[0.0952381 0.2745098 0.6302521 ]]





precision recall f1-score support

hate 0.11 0.24 0.15 41

offensive 0.24 0.72 0.36 95

neither 0.94 0.63 0.76 714

accuracy 0.62 850

macro avg 0.43 0.53 0.42 850

weighted avg 0.82 0.62 0.68 850

average eval\_loss: 1.5653372186474328, accuracy: 0.6211764705882353

training 15 epoch...

8.885833024978638 seconds used......

15 training finished! average train loss: 0.002682234639198416

total clf loss: 0.09658270870568231 total adv loss: 0.00044521919159024037

evaluating...

Confusion matrix, without normalization

[[ 10 19 12]

[ 10 64 21]

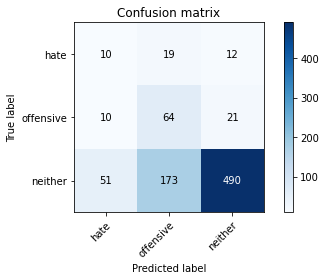
[ 51 173 490]]

Normalized confusion matrix

[[0.24390244 0.46341463 0.29268293]

[0.10526316 0.67368421 0.22105263]

[0.07142857 0.24229692 0.68627451]]





precision recall f1-score support

hate 0.14 0.24 0.18 41

offensive 0.25 0.67 0.36 95

neither 0.94 0.69 0.79 714

accuracy 0.66 850

macro avg 0.44 0.53 0.45 850

weighted avg 0.82 0.66 0.71 850

average eval\_loss: 1.419854096775918, accuracy: 0.6635294117647059

training 16 epoch...

8.908878087997437 seconds used......

16 training finished! average train loss: 0.002215661858725879

total clf loss: 0.07977501268032938 total adv loss: 0.00022372369551248994

evaluating...

Confusion matrix, without normalization

[[ 10 20 11]

[ 10 67 18]

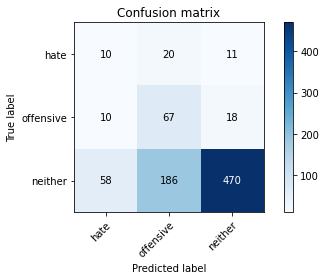
[ 58 186 470]]

Normalized confusion matrix

[[0.24390244 0.48780488 0.26829268]

[0.10526316 0.70526316 0.18947368]

[0.08123249 0.2605042 0.65826331]]





precision recall f1-score support

hate 0.13 0.24 0.17 41

offensive 0.25 0.71 0.36 95

neither 0.94 0.66 0.77 714

accuracy 0.64 850

macro avg 0.44 0.54 0.44 850

weighted avg 0.82 0.64 0.70 850

average eval\_loss: 1.5622841907491865, accuracy: 0.6435294117647059

training 17 epoch...

8.981194496154785 seconds used......

17 training finished! average train loss: 0.0019624044491340304

total clf loss: 0.070653457834851 total adv loss: 0.00013796931060028328

evaluating...

Confusion matrix, without normalization

[[ 10 20 11]

[ 10 68 17]

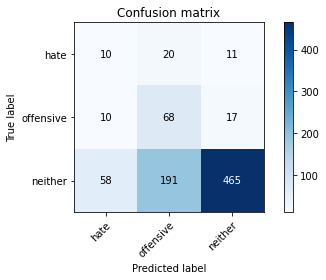
[ 58 191 465]]

Normalized confusion matrix

[[0.24390244 0.48780488 0.26829268]

[0.10526316 0.71578947 0.17894737]

[0.08123249 0.267507 0.6512605 ]]





precision recall f1-score support

hate 0.13 0.24 0.17 41

offensive 0.24 0.72 0.36 95

neither 0.94 0.65 0.77 714

accuracy 0.64 850

macro avg 0.44 0.54 0.43 850

weighted avg 0.83 0.64 0.70 850

average eval\_loss: 1.5987094055081523, accuracy: 0.6388235294117647

training 18 epoch...

8.939445972442627 seconds used......

18 training finished! average train loss: 0.0015496578918019724

total clf loss: 0.05579215323086828 total adv loss: 8.937232487937763e-05

evaluating...

Confusion matrix, without normalization

[[ 10 19 12]

[ 8 67 20]

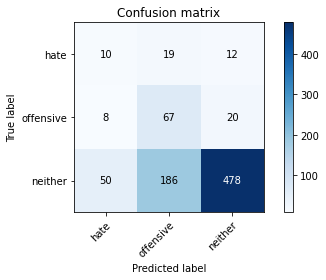
[ 50 186 478]]

Normalized confusion matrix

[[0.24390244 0.46341463 0.29268293]

[0.08421053 0.70526316 0.21052632]

[0.07002801 0.2605042 0.66946779]]





precision recall f1-score support

hate 0.15 0.24 0.18 41

offensive 0.25 0.71 0.37 95

neither 0.94 0.67 0.78 714

accuracy 0.65 850

macro avg 0.44 0.54 0.44 850

weighted avg 0.82 0.65 0.71 850

average eval\_loss: 1.5666220685423806, accuracy: 0.6529411764705882

training 19 epoch...

8.951361179351807 seconds used......

19 training finished! average train loss: 0.0015687625418649986

total clf loss: 0.05647961742943153 total adv loss: 8.332290167345491e-05

evaluating...

Confusion matrix, without normalization

[[ 10 19 12]

[ 10 62 23]

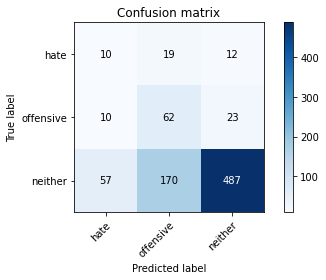
[ 57 170 487]]

Normalized confusion matrix

[[0.24390244 0.46341463 0.29268293]

[0.10526316 0.65263158 0.24210526]

[0.07983193 0.23809524 0.68207283]]





precision recall f1-score support

hate 0.13 0.24 0.17 41

offensive 0.25 0.65 0.36 95

neither 0.93 0.68 0.79 714

accuracy 0.66 850

macro avg 0.44 0.53 0.44 850

weighted avg 0.82 0.66 0.71 850

average eval\_loss: 1.5100519805388568, accuracy: 0.6576470588235294

training 20 epoch...

8.994513988494873 seconds used......

20 training finished! average train loss: 0.001315603111935262

total clf loss: 0.04736500087892637 total adv loss: 6.577664279916462e-05

evaluating...

Confusion matrix, without normalization

[[ 10 19 12]

[ 9 70 16]

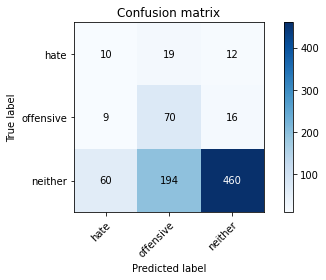
[ 60 194 460]]

Normalized confusion matrix

[[0.24390244 0.46341463 0.29268293]

[0.09473684 0.73684211 0.16842105]

[0.08403361 0.27170868 0.6442577 ]]





precision recall f1-score support

hate 0.13 0.24 0.17 41

offensive 0.25 0.74 0.37 95

neither 0.94 0.64 0.77 714

accuracy 0.64 850

macro avg 0.44 0.54 0.43 850

weighted avg 0.83 0.64 0.69 850

average eval\_loss: 1.6758221230729549, accuracy: 0.6352941176470588