1 run......

Downloading: 100%

464/464 [00:00<00:00, 13.2kB/s]

Downloading: 100%

270M/270M [00:09<00:00, 30.2MB/s]

training 1 epoch...

12.385462284088135 seconds used......

1 training finished! average train loss: 0.8298927761041202

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 0 95]

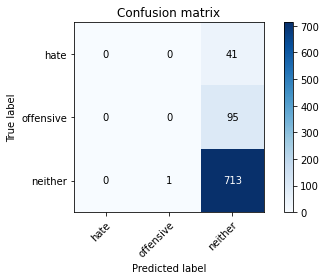
[ 0 1 713]]

Normalized confusion matrix

[[0. 0. 1. ]

[0. 0. 1. ]

[0. 0.00140056 0.99859944]]





/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 0.00 0.00 0.00 95

neither 0.84 1.00 0.91 714

accuracy 0.84 850

macro avg 0.28 0.33 0.30 850

weighted avg 0.71 0.84 0.77 850

average eval\_loss: 0.6030441007988397, accuracy: 0.8388235294117647

training 2 epoch...

11.175599813461304 seconds used......

2 training finished! average train loss: 0.6937036362978128

evaluating...

Confusion matrix, without normalization

[[ 0 12 29]

[ 0 52 43]

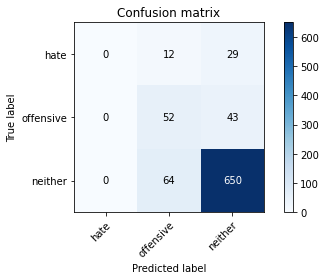
[ 0 64 650]]

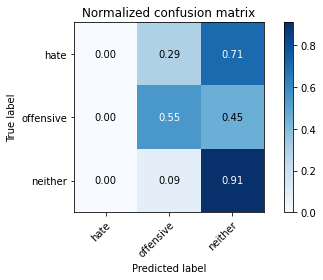
Normalized confusion matrix

[[0. 0.29268293 0.70731707]

[0. 0.54736842 0.45263158]

[0. 0.08963585 0.91036415]]





/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 0.41 0.55 0.47 95

neither 0.90 0.91 0.91 714

accuracy 0.83 850

macro avg 0.44 0.49 0.46 850

weighted avg 0.80 0.83 0.81 850

average eval\_loss: 0.5783651578564977, accuracy: 0.8258823529411765

training 3 epoch...

11.29759955406189 seconds used......

3 training finished! average train loss: 0.5295059185761671

evaluating...

Confusion matrix, without normalization

[[ 0 11 30]

[ 0 44 51]

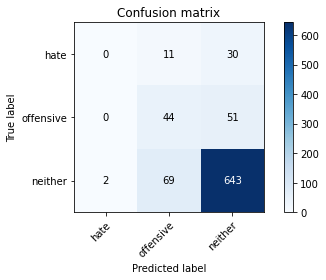
[ 2 69 643]]

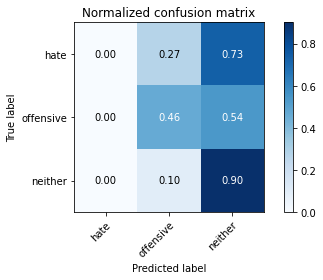
Normalized confusion matrix

[[0. 0.26829268 0.73170732]

[0. 0.46315789 0.53684211]

[0.00280112 0.09663866 0.90056022]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.35 0.46 0.40 95

neither 0.89 0.90 0.89 714

accuracy 0.81 850

macro avg 0.41 0.45 0.43 850

weighted avg 0.79 0.81 0.80 850

average eval\_loss: 0.509940107331373, accuracy: 0.808235294117647

training 4 epoch...

11.19869351387024 seconds used......

4 training finished! average train loss: 0.35229277461767194

evaluating...

Confusion matrix, without normalization

[[ 10 7 24]

[ 14 27 54]

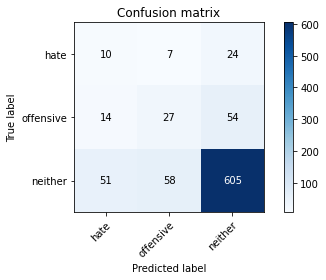
[ 51 58 605]]

Normalized confusion matrix

[[0.24390244 0.17073171 0.58536585]

[0.14736842 0.28421053 0.56842105]

[0.07142857 0.08123249 0.84733894]]





precision recall f1-score support

hate 0.13 0.24 0.17 41

offensive 0.29 0.28 0.29 95

neither 0.89 0.85 0.87 714

accuracy 0.76 850

macro avg 0.44 0.46 0.44 850

weighted avg 0.78 0.76 0.77 850

average eval\_loss: 0.6119872063918169, accuracy: 0.7552941176470588

training 5 epoch...

11.272398710250854 seconds used......

5 training finished! average train loss: 0.17965870084097751

evaluating...

Confusion matrix, without normalization

[[ 7 12 22]

[ 2 43 50]

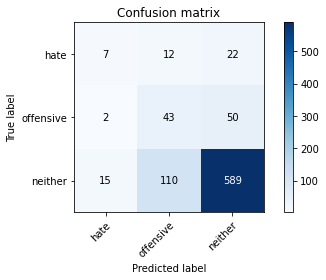
[ 15 110 589]]

Normalized confusion matrix

[[0.17073171 0.29268293 0.53658537]

[0.02105263 0.45263158 0.52631579]

[0.0210084 0.15406162 0.82492997]]





precision recall f1-score support

hate 0.29 0.17 0.22 41

offensive 0.26 0.45 0.33 95

neither 0.89 0.82 0.86 714

accuracy 0.75 850

macro avg 0.48 0.48 0.47 850

weighted avg 0.79 0.75 0.77 850

average eval\_loss: 0.7233313225307177, accuracy: 0.7517647058823529

training 6 epoch...

11.18929409980774 seconds used......

6 training finished! average train loss: 0.09118872012656469

evaluating...

Confusion matrix, without normalization

[[ 3 2 36]

[ 3 10 82]

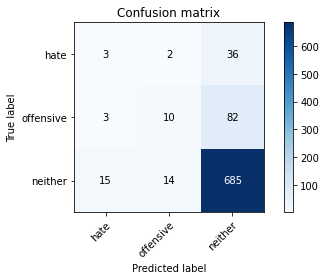
[ 15 14 685]]

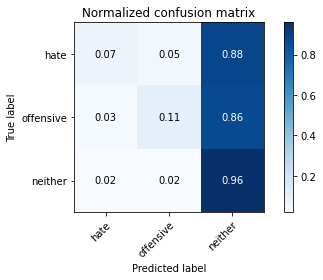
Normalized confusion matrix

[[0.07317073 0.04878049 0.87804878]

[0.03157895 0.10526316 0.86315789]

[0.0210084 0.01960784 0.95938375]]





precision recall f1-score support

hate 0.14 0.07 0.10 41

offensive 0.38 0.11 0.17 95

neither 0.85 0.96 0.90 714

accuracy 0.82 850

macro avg 0.46 0.38 0.39 850

weighted avg 0.77 0.82 0.78 850

average eval\_loss: 0.832267134312302, accuracy: 0.8211764705882353

training 7 epoch...

11.10390567779541 seconds used......

7 training finished! average train loss: 0.0477041587806665

evaluating...

Confusion matrix, without normalization

[[ 7 13 21]

[ 5 51 39]

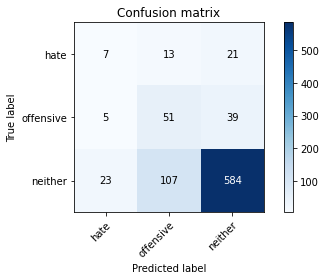
[ 23 107 584]]

Normalized confusion matrix

[[0.17073171 0.31707317 0.51219512]

[0.05263158 0.53684211 0.41052632]

[0.03221289 0.14985994 0.81792717]]





precision recall f1-score support

hate 0.20 0.17 0.18 41

offensive 0.30 0.54 0.38 95

neither 0.91 0.82 0.86 714

accuracy 0.76 850

macro avg 0.47 0.51 0.48 850

weighted avg 0.80 0.76 0.77 850

average eval\_loss: 0.9041825236798654, accuracy: 0.7552941176470588

training 8 epoch...

11.2339186668396 seconds used......

8 training finished! average train loss: 0.021033073202348672

evaluating...

Confusion matrix, without normalization

[[ 8 10 23]

[ 12 49 34]

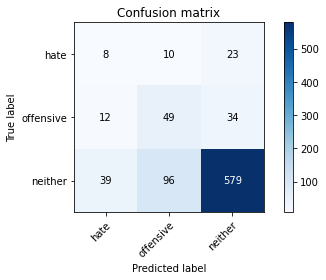
[ 39 96 579]]

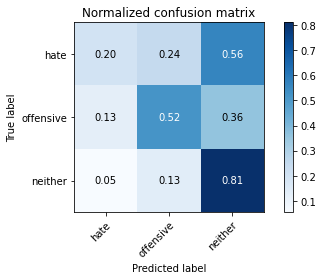
Normalized confusion matrix

[[0.19512195 0.24390244 0.56097561]

[0.12631579 0.51578947 0.35789474]

[0.05462185 0.13445378 0.81092437]]





precision recall f1-score support

hate 0.14 0.20 0.16 41

offensive 0.32 0.52 0.39 95

neither 0.91 0.81 0.86 714

accuracy 0.75 850

macro avg 0.45 0.51 0.47 850

weighted avg 0.81 0.75 0.77 850

average eval\_loss: 0.9237196424943503, accuracy: 0.7482352941176471

training 9 epoch...

11.144018650054932 seconds used......

9 training finished! average train loss: 0.008198556071147323

evaluating...

Confusion matrix, without normalization

[[ 7 8 26]

[ 8 51 36]

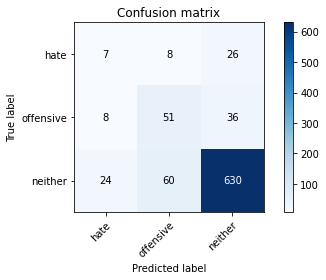
[ 24 60 630]]

Normalized confusion matrix

[[0.17073171 0.19512195 0.63414634]

[0.08421053 0.53684211 0.37894737]

[0.03361345 0.08403361 0.88235294]]





precision recall f1-score support

hate 0.18 0.17 0.18 41

offensive 0.43 0.54 0.48 95

neither 0.91 0.88 0.90 714

accuracy 0.81 850

macro avg 0.51 0.53 0.52 850

weighted avg 0.82 0.81 0.81 850

average eval\_loss: 0.8169448595280152, accuracy: 0.8094117647058824

training 10 epoch...

11.2383131980896 seconds used......

10 training finished! average train loss: 0.004629646352707194

evaluating...

Confusion matrix, without normalization

[[ 6 6 29]

[ 6 53 36]

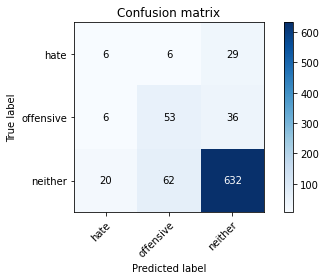
[ 20 62 632]]

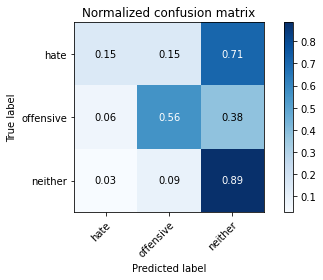
Normalized confusion matrix

[[0.14634146 0.14634146 0.70731707]

[0.06315789 0.55789474 0.37894737]

[0.0280112 0.08683473 0.88515406]]





precision recall f1-score support

hate 0.19 0.15 0.16 41

offensive 0.44 0.56 0.49 95

neither 0.91 0.89 0.90 714

accuracy 0.81 850

macro avg 0.51 0.53 0.52 850

weighted avg 0.82 0.81 0.82 850

average eval\_loss: 0.8657066473672383, accuracy: 0.8129411764705883

2 run......

training 1 epoch...

11.239588499069214 seconds used......

1 training finished! average train loss: 0.8709739808852855

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 0 95]

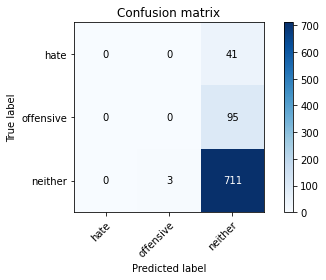
[ 0 3 711]]

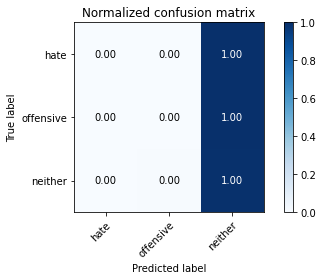
Normalized confusion matrix

[[0. 0. 1. ]

[0. 0. 1. ]

[0. 0.00420168 0.99579832]]





/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 0.00 0.00 0.00 95

neither 0.84 1.00 0.91 714

accuracy 0.84 850

macro avg 0.28 0.33 0.30 850

weighted avg 0.71 0.84 0.77 850

average eval\_loss: 0.5779704675078392, accuracy: 0.8364705882352941

training 2 epoch...

11.127774477005005 seconds used......

2 training finished! average train loss: 0.7116049706935883

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 0 95]

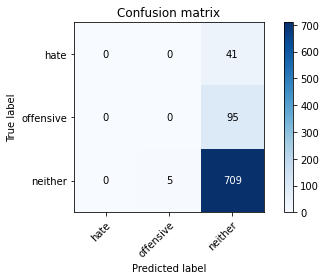
[ 0 5 709]]

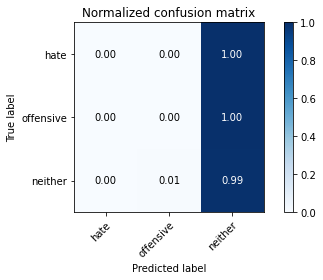
Normalized confusion matrix

[[0. 0. 1. ]

[0. 0. 1. ]

[0. 0.0070028 0.9929972]]





/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 0.00 0.00 0.00 95

neither 0.84 0.99 0.91 714

accuracy 0.83 850

macro avg 0.28 0.33 0.30 850

weighted avg 0.70 0.83 0.76 850

average eval\_loss: 0.5142251541621463, accuracy: 0.8341176470588235

training 3 epoch...

11.134482622146606 seconds used......

3 training finished! average train loss: 0.5783386471179816

evaluating...

Confusion matrix, without normalization

[[ 11 0 30]

[ 13 9 73]

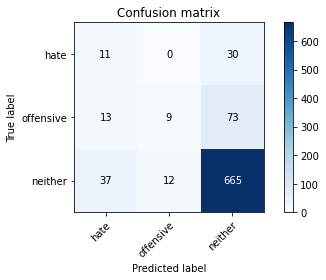
[ 37 12 665]]

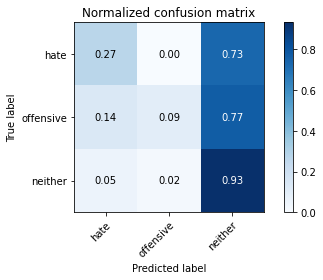
Normalized confusion matrix

[[0.26829268 0. 0.73170732]

[0.13684211 0.09473684 0.76842105]

[0.05182073 0.01680672 0.93137255]]





precision recall f1-score support

hate 0.18 0.27 0.22 41

offensive 0.43 0.09 0.16 95

neither 0.87 0.93 0.90 714

accuracy 0.81 850

macro avg 0.49 0.43 0.42 850

weighted avg 0.78 0.81 0.78 850

average eval\_loss: 0.5526784643704115, accuracy: 0.8058823529411765

training 4 epoch...

11.3996102809906 seconds used......

4 training finished! average train loss: 0.42173701708133404

evaluating...

Confusion matrix, without normalization

[[ 2 17 22]

[ 3 61 31]

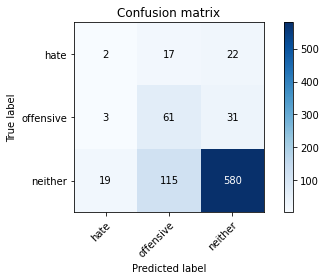
[ 19 115 580]]

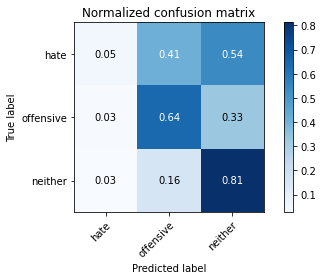
Normalized confusion matrix

[[0.04878049 0.41463415 0.53658537]

[0.03157895 0.64210526 0.32631579]

[0.02661064 0.16106443 0.81232493]]





precision recall f1-score support

hate 0.08 0.05 0.06 41

offensive 0.32 0.64 0.42 95

neither 0.92 0.81 0.86 714

accuracy 0.76 850

macro avg 0.44 0.50 0.45 850

weighted avg 0.81 0.76 0.77 850

average eval\_loss: 0.6016033617129852, accuracy: 0.7564705882352941

training 5 epoch...

11.142040491104126 seconds used......

5 training finished! average train loss: 0.22459192361969213

evaluating...

Confusion matrix, without normalization

[[ 8 0 33]

[ 6 7 82]

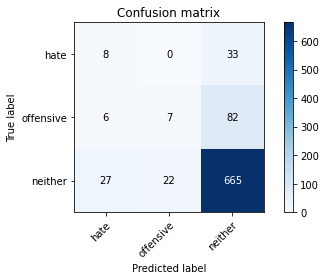
[ 27 22 665]]

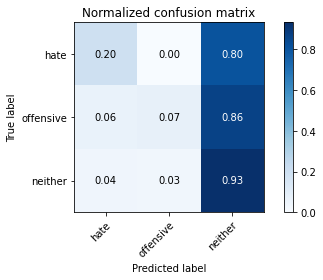
Normalized confusion matrix

[[0.19512195 0. 0.80487805]

[0.06315789 0.07368421 0.86315789]

[0.03781513 0.03081232 0.93137255]]





precision recall f1-score support

hate 0.20 0.20 0.20 41

offensive 0.24 0.07 0.11 95

neither 0.85 0.93 0.89 714

accuracy 0.80 850

macro avg 0.43 0.40 0.40 850

weighted avg 0.75 0.80 0.77 850

average eval\_loss: 0.7275464511205724, accuracy: 0.8

training 6 epoch...

11.295188188552856 seconds used......

6 training finished! average train loss: 0.09065572864447649

evaluating...

Confusion matrix, without normalization

[[ 10 0 31]

[ 12 19 64]

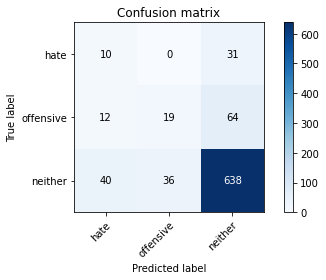
[ 40 36 638]]

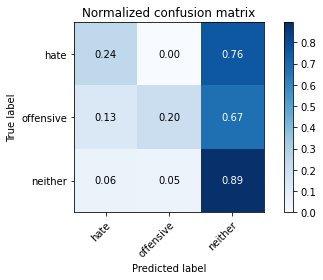
Normalized confusion matrix

[[0.24390244 0. 0.75609756]

[0.12631579 0.2 0.67368421]

[0.05602241 0.05042017 0.89355742]]





precision recall f1-score support

hate 0.16 0.24 0.19 41

offensive 0.35 0.20 0.25 95

neither 0.87 0.89 0.88 714

accuracy 0.78 850

macro avg 0.46 0.45 0.44 850

weighted avg 0.78 0.78 0.78 850

average eval\_loss: 0.8005452963989228, accuracy: 0.7847058823529411

training 7 epoch...

11.1183500289917 seconds used......

7 training finished! average train loss: 0.043701877645575087

evaluating...

Confusion matrix, without normalization

[[ 9 4 28]

[ 14 23 58]

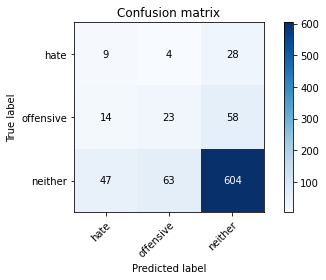
[ 47 63 604]]

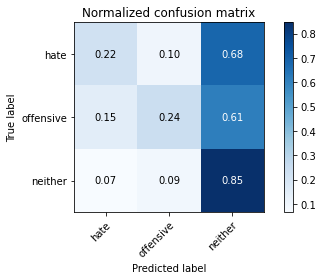
Normalized confusion matrix

[[0.2195122 0.09756098 0.68292683]

[0.14736842 0.24210526 0.61052632]

[0.06582633 0.08823529 0.84593838]]





precision recall f1-score support

hate 0.13 0.22 0.16 41

offensive 0.26 0.24 0.25 95

neither 0.88 0.85 0.86 714

accuracy 0.75 850

macro avg 0.42 0.44 0.42 850

weighted avg 0.77 0.75 0.76 850

average eval\_loss: 0.931531777302193, accuracy: 0.7482352941176471

training 8 epoch...

11.231523752212524 seconds used......

8 training finished! average train loss: 0.018545435823929996

evaluating...

Confusion matrix, without normalization

[[ 12 2 27]

[ 18 22 55]

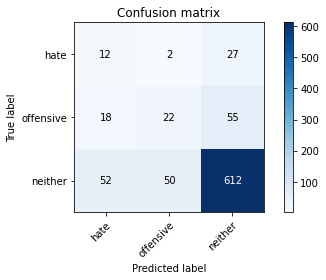
[ 52 50 612]]

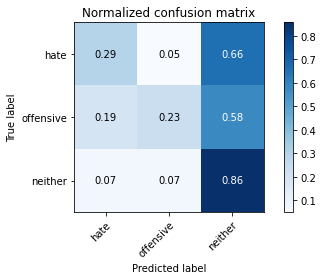
Normalized confusion matrix

[[0.29268293 0.04878049 0.65853659]

[0.18947368 0.23157895 0.57894737]

[0.07282913 0.07002801 0.85714286]]





precision recall f1-score support

hate 0.15 0.29 0.20 41

offensive 0.30 0.23 0.26 95

neither 0.88 0.86 0.87 714

accuracy 0.76 850

macro avg 0.44 0.46 0.44 850

weighted avg 0.78 0.76 0.77 850

average eval\_loss: 1.035063390159806, accuracy: 0.76

training 9 epoch...

11.102649450302124 seconds used......

9 training finished! average train loss: 0.006219260129504479

evaluating...

Confusion matrix, without normalization

[[ 4 5 32]

[ 6 28 61]

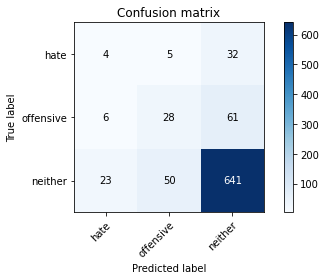
[ 23 50 641]]

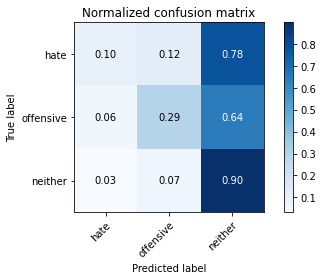
Normalized confusion matrix

[[0.09756098 0.12195122 0.7804878 ]

[0.06315789 0.29473684 0.64210526]

[0.03221289 0.07002801 0.8977591 ]]





precision recall f1-score support

hate 0.12 0.10 0.11 41

offensive 0.34 0.29 0.31 95

neither 0.87 0.90 0.89 714

accuracy 0.79 850

macro avg 0.44 0.43 0.44 850

weighted avg 0.78 0.79 0.78 850

average eval\_loss: 1.0130371461408538, accuracy: 0.7917647058823529

training 10 epoch...

11.266003608703613 seconds used......

10 training finished! average train loss: 0.003794373515456055

evaluating...

Confusion matrix, without normalization

[[ 7 4 30]

[ 8 30 57]

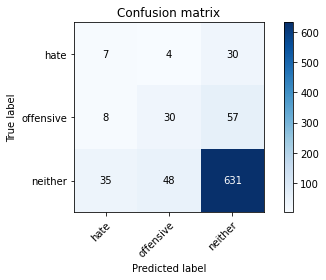
[ 35 48 631]]

Normalized confusion matrix

[[0.17073171 0.09756098 0.73170732]

[0.08421053 0.31578947 0.6 ]

[0.04901961 0.06722689 0.8837535 ]]





precision recall f1-score support

hate 0.14 0.17 0.15 41

offensive 0.37 0.32 0.34 95

neither 0.88 0.88 0.88 714

accuracy 0.79 850

macro avg 0.46 0.46 0.46 850

weighted avg 0.79 0.79 0.79 850

average eval\_loss: 1.0102944436914905, accuracy: 0.7858823529411765

3 run......

training 1 epoch...

11.105181217193604 seconds used......

1 training finished! average train loss: 0.8479992022881141

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 0 95]

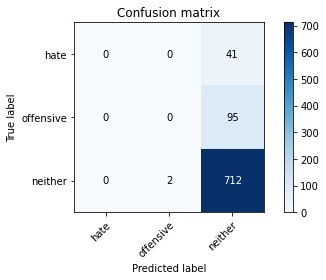
[ 0 2 712]]

Normalized confusion matrix

[[0. 0. 1. ]

[0. 0. 1. ]

[0. 0.00280112 0.99719888]]





/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 0.00 0.00 0.00 95

neither 0.84 1.00 0.91 714

accuracy 0.84 850

macro avg 0.28 0.33 0.30 850

weighted avg 0.71 0.84 0.77 850

average eval\_loss: 0.624360152281994, accuracy: 0.8376470588235294

training 2 epoch...

11.172459363937378 seconds used......

2 training finished! average train loss: 0.7071155328017015

evaluating...

Confusion matrix, without normalization

[[ 0 11 30]

[ 0 35 60]

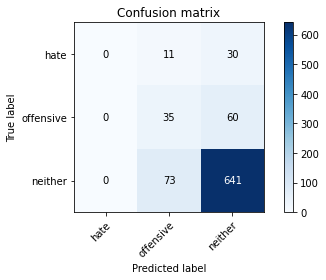
[ 0 73 641]]

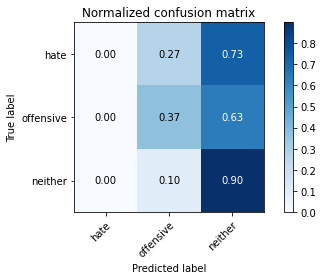
Normalized confusion matrix

[[0. 0.26829268 0.73170732]

[0. 0.36842105 0.63157895]

[0. 0.1022409 0.8977591 ]]





/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 0.29 0.37 0.33 95

neither 0.88 0.90 0.89 714

accuracy 0.80 850

macro avg 0.39 0.42 0.40 850

weighted avg 0.77 0.80 0.78 850

average eval\_loss: 0.5904795763970807, accuracy: 0.7952941176470588

training 3 epoch...

11.118898630142212 seconds used......

3 training finished! average train loss: 0.5521641621222863

evaluating...

Confusion matrix, without normalization

[[ 0 10 31]

[ 0 56 39]

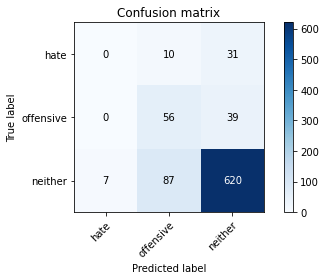
[ 7 87 620]]

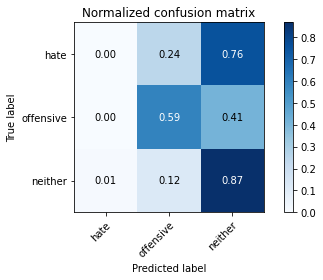
Normalized confusion matrix

[[0. 0.24390244 0.75609756]

[0. 0.58947368 0.41052632]

[0.00980392 0.12184874 0.86834734]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.37 0.59 0.45 95

neither 0.90 0.87 0.88 714

accuracy 0.80 850

macro avg 0.42 0.49 0.44 850

weighted avg 0.80 0.80 0.79 850

average eval\_loss: 0.5393109400431777, accuracy: 0.7952941176470588

training 4 epoch...

11.246728420257568 seconds used......

4 training finished! average train loss: 0.37507629291369365

evaluating...

Confusion matrix, without normalization

[[ 6 12 23]

[ 4 46 45]

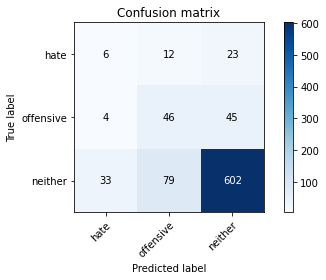
[ 33 79 602]]

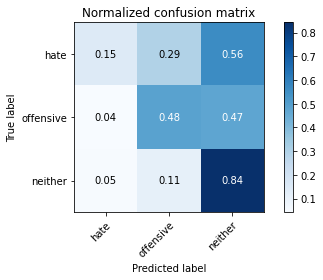
Normalized confusion matrix

[[0.14634146 0.29268293 0.56097561]

[0.04210526 0.48421053 0.47368421]

[0.04621849 0.11064426 0.84313725]]





precision recall f1-score support

hate 0.14 0.15 0.14 41

offensive 0.34 0.48 0.40 95

neither 0.90 0.84 0.87 714

accuracy 0.77 850

macro avg 0.46 0.49 0.47 850

weighted avg 0.80 0.77 0.78 850

average eval\_loss: 0.5899361027621253, accuracy: 0.7694117647058824

training 5 epoch...

11.103983879089355 seconds used......

5 training finished! average train loss: 0.1768188454497319

evaluating...

Confusion matrix, without normalization

[[ 7 3 31]

[ 7 17 71]

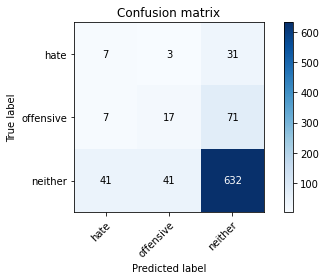
[ 41 41 632]]

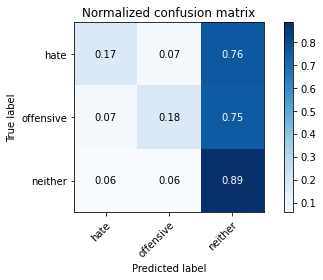
Normalized confusion matrix

[[0.17073171 0.07317073 0.75609756]

[0.07368421 0.17894737 0.74736842]

[0.05742297 0.05742297 0.88515406]]





precision recall f1-score support

hate 0.13 0.17 0.15 41

offensive 0.28 0.18 0.22 95

neither 0.86 0.89 0.87 714

accuracy 0.77 850

macro avg 0.42 0.41 0.41 850

weighted avg 0.76 0.77 0.76 850

average eval\_loss: 0.7442065439685139, accuracy: 0.7717647058823529

training 6 epoch...

11.201356410980225 seconds used......

6 training finished! average train loss: 0.07970260258477468

evaluating...

Confusion matrix, without normalization

[[ 7 6 28]

[ 5 25 65]

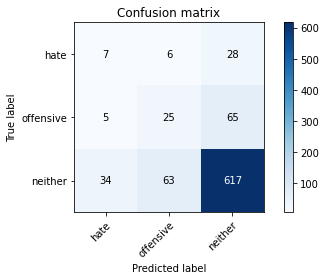
[ 34 63 617]]

Normalized confusion matrix

[[0.17073171 0.14634146 0.68292683]

[0.05263158 0.26315789 0.68421053]

[0.04761905 0.08823529 0.86414566]]





precision recall f1-score support

hate 0.15 0.17 0.16 41

offensive 0.27 0.26 0.26 95

neither 0.87 0.86 0.87 714

accuracy 0.76 850

macro avg 0.43 0.43 0.43 850

weighted avg 0.77 0.76 0.77 850

average eval\_loss: 0.825422523424116, accuracy: 0.7635294117647059

training 7 epoch...

11.122924566268921 seconds used......

7 training finished! average train loss: 0.03124855844877087

evaluating...

Confusion matrix, without normalization

[[ 8 14 19]

[ 6 43 46]

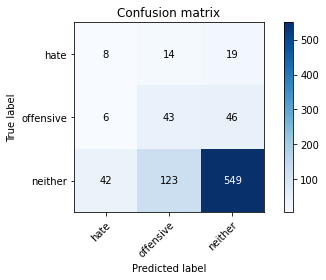
[ 42 123 549]]

Normalized confusion matrix

[[0.19512195 0.34146341 0.46341463]

[0.06315789 0.45263158 0.48421053]

[0.05882353 0.17226891 0.76890756]]





precision recall f1-score support

hate 0.14 0.20 0.16 41

offensive 0.24 0.45 0.31 95

neither 0.89 0.77 0.83 714

accuracy 0.71 850

macro avg 0.43 0.47 0.43 850

weighted avg 0.78 0.71 0.74 850

average eval\_loss: 1.1387675073225225, accuracy: 0.7058823529411765

training 8 epoch...

11.183669567108154 seconds used......

8 training finished! average train loss: 0.012403432021920497

evaluating...

Confusion matrix, without normalization

[[ 6 9 26]

[ 7 37 51]

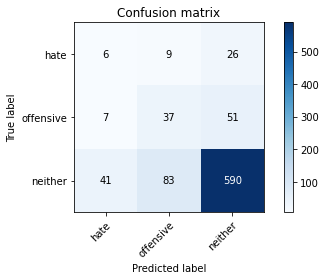
[ 41 83 590]]

Normalized confusion matrix

[[0.14634146 0.2195122 0.63414634]

[0.07368421 0.38947368 0.53684211]

[0.05742297 0.1162465 0.82633053]]





precision recall f1-score support

hate 0.11 0.15 0.13 41

offensive 0.29 0.39 0.33 95

neither 0.88 0.83 0.85 714

accuracy 0.74 850

macro avg 0.43 0.45 0.44 850

weighted avg 0.78 0.74 0.76 850

average eval\_loss: 1.0370490565587398, accuracy: 0.7447058823529412

training 9 epoch...

11.08559799194336 seconds used......

9 training finished! average train loss: 0.01246561638903446

evaluating...

Confusion matrix, without normalization

[[ 10 4 27]

[ 7 24 64]

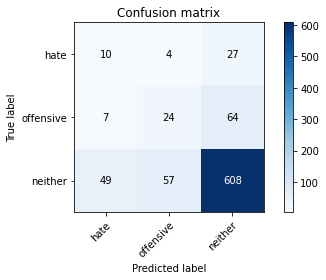
[ 49 57 608]]

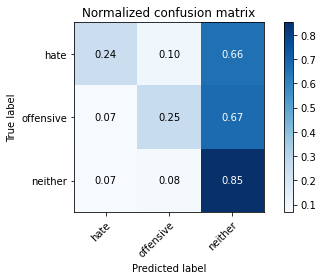
Normalized confusion matrix

[[0.24390244 0.09756098 0.65853659]

[0.07368421 0.25263158 0.67368421]

[0.06862745 0.07983193 0.85154062]]





precision recall f1-score support

hate 0.15 0.24 0.19 41

offensive 0.28 0.25 0.27 95

neither 0.87 0.85 0.86 714

accuracy 0.76 850

macro avg 0.43 0.45 0.44 850

weighted avg 0.77 0.76 0.76 850

average eval\_loss: 1.0893253047685247, accuracy: 0.7552941176470588

training 10 epoch...

11.09962773323059 seconds used......

10 training finished! average train loss: 0.006210031763363916

evaluating...

Confusion matrix, without normalization

[[ 8 3 30]

[ 4 14 77]

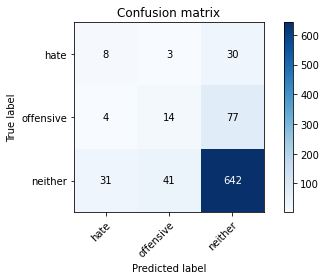
[ 31 41 642]]

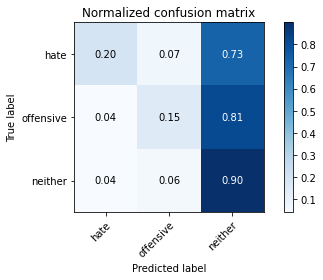
Normalized confusion matrix

[[0.19512195 0.07317073 0.73170732]

[0.04210526 0.14736842 0.81052632]

[0.04341737 0.05742297 0.89915966]]





precision recall f1-score support

hate 0.19 0.20 0.19 41

offensive 0.24 0.15 0.18 95

neither 0.86 0.90 0.88 714

accuracy 0.78 850

macro avg 0.43 0.41 0.42 850

weighted avg 0.76 0.78 0.77 850

average eval\_loss: 1.1481606676777945, accuracy: 0.7811764705882352

4 run......

training 1 epoch...

11.126285552978516 seconds used......

1 training finished! average train loss: 0.8413036759083088

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 0 95]

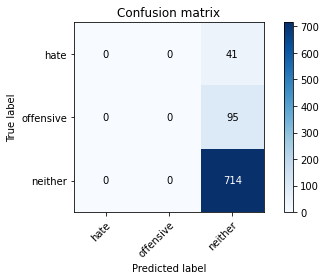
[ 0 0 714]]

Normalized confusion matrix

[[0. 0. 1.]

[0. 0. 1.]

[0. 0. 1.]]





/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 0.00 0.00 0.00 95

neither 0.84 1.00 0.91 714

accuracy 0.84 850

macro avg 0.28 0.33 0.30 850

weighted avg 0.71 0.84 0.77 850

average eval\_loss: 0.5762387614610583, accuracy: 0.84

training 2 epoch...

11.120228052139282 seconds used......

2 training finished! average train loss: 0.7156917406962469

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 0 95]

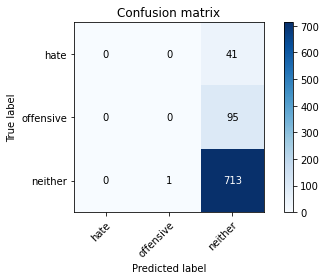
[ 0 1 713]]

Normalized confusion matrix

[[0. 0. 1. ]

[0. 0. 1. ]

[0. 0.00140056 0.99859944]]





/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 0.00 0.00 0.00 95

neither 0.84 1.00 0.91 714

accuracy 0.84 850

macro avg 0.28 0.33 0.30 850

weighted avg 0.71 0.84 0.77 850

average eval\_loss: 0.4977269888270733, accuracy: 0.8388235294117647

training 3 epoch...

11.105473279953003 seconds used......

3 training finished! average train loss: 0.5860593777436477

evaluating...

Confusion matrix, without normalization

[[ 0 11 30]

[ 0 52 43]

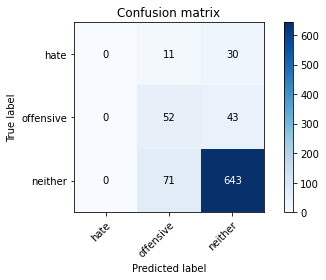
[ 0 71 643]]

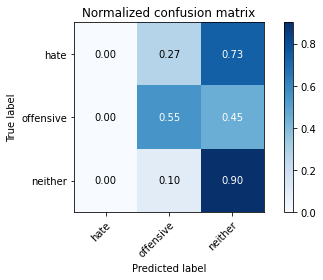
Normalized confusion matrix

[[0. 0.26829268 0.73170732]

[0. 0.54736842 0.45263158]

[0. 0.09943978 0.90056022]]





/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 0.39 0.55 0.45 95

neither 0.90 0.90 0.90 714

accuracy 0.82 850

macro avg 0.43 0.48 0.45 850

weighted avg 0.80 0.82 0.81 850

average eval\_loss: 0.5008498544609824, accuracy: 0.8176470588235294

training 4 epoch...

11.134618282318115 seconds used......

4 training finished! average train loss: 0.42403316429028143

evaluating...

Confusion matrix, without normalization

[[ 8 9 24]

[ 9 36 50]

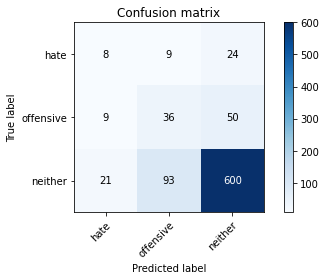
[ 21 93 600]]

Normalized confusion matrix

[[0.19512195 0.2195122 0.58536585]

[0.09473684 0.37894737 0.52631579]

[0.02941176 0.1302521 0.84033613]]





precision recall f1-score support

hate 0.21 0.20 0.20 41

offensive 0.26 0.38 0.31 95

neither 0.89 0.84 0.86 714

accuracy 0.76 850

macro avg 0.45 0.47 0.46 850

weighted avg 0.79 0.76 0.77 850

average eval\_loss: 0.5863606928912706, accuracy: 0.7576470588235295

training 5 epoch...

11.150301694869995 seconds used......

5 training finished! average train loss: 0.24980781903633706

evaluating...

Confusion matrix, without normalization

[[ 3 8 30]

[ 2 24 69]

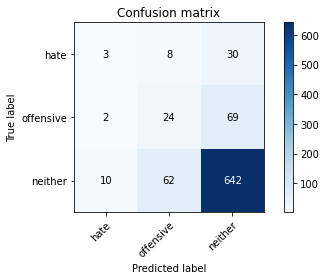
[ 10 62 642]]

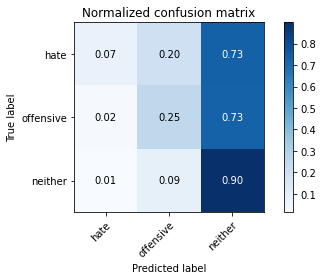
Normalized confusion matrix

[[0.07317073 0.19512195 0.73170732]

[0.02105263 0.25263158 0.72631579]

[0.0140056 0.08683473 0.89915966]]





precision recall f1-score support

hate 0.20 0.07 0.11 41

offensive 0.26 0.25 0.25 95

neither 0.87 0.90 0.88 714

accuracy 0.79 850

macro avg 0.44 0.41 0.41 850

weighted avg 0.77 0.79 0.77 850

average eval\_loss: 0.6398633934481632, accuracy: 0.7870588235294118

training 6 epoch...

11.099016666412354 seconds used......

6 training finished! average train loss: 0.09744161778631119

evaluating...

Confusion matrix, without normalization

[[ 3 9 29]

[ 2 37 56]

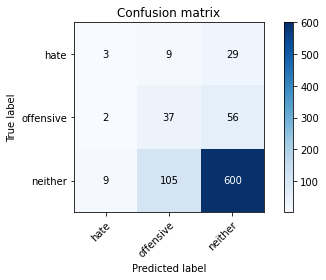
[ 9 105 600]]

Normalized confusion matrix

[[0.07317073 0.2195122 0.70731707]

[0.02105263 0.38947368 0.58947368]

[0.01260504 0.14705882 0.84033613]]





precision recall f1-score support

hate 0.21 0.07 0.11 41

offensive 0.25 0.39 0.30 95

neither 0.88 0.84 0.86 714

accuracy 0.75 850

macro avg 0.45 0.43 0.42 850

weighted avg 0.77 0.75 0.76 850

average eval\_loss: 0.7851703219483931, accuracy: 0.7529411764705882

training 7 epoch...

11.253954887390137 seconds used......

7 training finished! average train loss: 0.042164380977360104

evaluating...

Confusion matrix, without normalization

[[ 9 5 27]

[ 13 23 59]

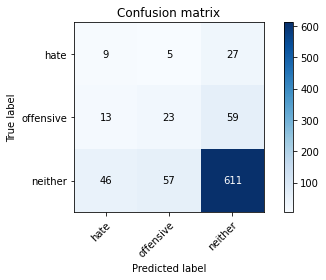
[ 46 57 611]]

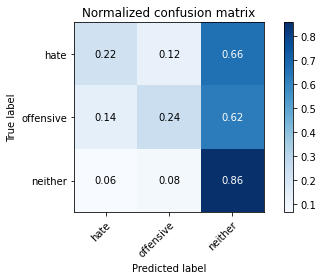
Normalized confusion matrix

[[0.2195122 0.12195122 0.65853659]

[0.13684211 0.24210526 0.62105263]

[0.06442577 0.07983193 0.8557423 ]]





precision recall f1-score support

hate 0.13 0.22 0.17 41

offensive 0.27 0.24 0.26 95

neither 0.88 0.86 0.87 714

accuracy 0.76 850

macro avg 0.43 0.44 0.43 850

weighted avg 0.77 0.76 0.76 850

average eval\_loss: 0.8665279492053615, accuracy: 0.7564705882352941

training 8 epoch...

11.08799123764038 seconds used......

8 training finished! average train loss: 0.013165582895565492

evaluating...

Confusion matrix, without normalization

[[ 8 6 27]

[ 9 23 63]

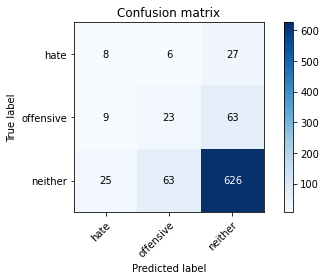
[ 25 63 626]]

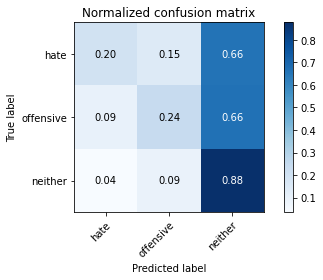
Normalized confusion matrix

[[0.19512195 0.14634146 0.65853659]

[0.09473684 0.24210526 0.66315789]

[0.03501401 0.08823529 0.8767507 ]]





precision recall f1-score support

hate 0.19 0.20 0.19 41

offensive 0.25 0.24 0.25 95

neither 0.87 0.88 0.88 714

accuracy 0.77 850

macro avg 0.44 0.44 0.44 850

weighted avg 0.77 0.77 0.77 850

average eval\_loss: 0.9293797930423712, accuracy: 0.7729411764705882

training 9 epoch...

11.261016607284546 seconds used......

9 training finished! average train loss: 0.007831550103862986

evaluating...

Confusion matrix, without normalization

[[ 4 7 30]

[ 4 24 67]

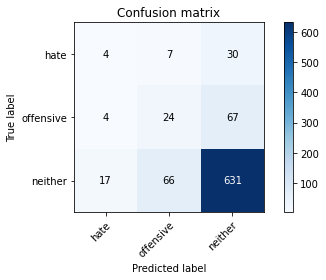
[ 17 66 631]]

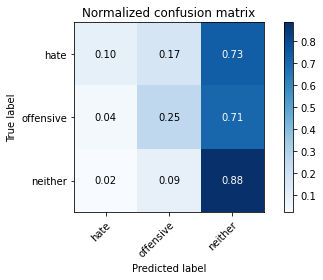
Normalized confusion matrix

[[0.09756098 0.17073171 0.73170732]

[0.04210526 0.25263158 0.70526316]

[0.02380952 0.09243697 0.8837535 ]]





precision recall f1-score support

hate 0.16 0.10 0.12 41

offensive 0.25 0.25 0.25 95

neither 0.87 0.88 0.88 714

accuracy 0.78 850

macro avg 0.42 0.41 0.42 850

weighted avg 0.76 0.78 0.77 850

average eval\_loss: 0.9754839244956519, accuracy: 0.7752941176470588

training 10 epoch...

11.101161479949951 seconds used......

10 training finished! average train loss: 0.0037763831277306263

evaluating...

Confusion matrix, without normalization

[[ 9 5 27]

[ 10 27 58]

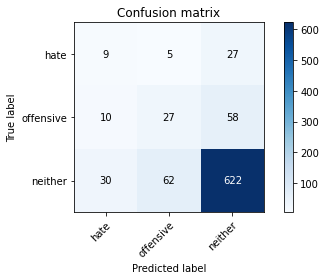
[ 30 62 622]]

Normalized confusion matrix

[[0.2195122 0.12195122 0.65853659]

[0.10526316 0.28421053 0.61052632]

[0.04201681 0.08683473 0.87114846]]





precision recall f1-score support

hate 0.18 0.22 0.20 41

offensive 0.29 0.28 0.29 95

neither 0.88 0.87 0.88 714

accuracy 0.77 850

macro avg 0.45 0.46 0.45 850

weighted avg 0.78 0.77 0.78 850

average eval\_loss: 1.0115038794018654, accuracy: 0.7741176470588236

5 run......

training 1 epoch...

11.221915245056152 seconds used......

1 training finished! average train loss: 0.8483704887903654

evaluating...

Confusion matrix, without normalization

[[ 0 5 36]

[ 0 1 94]

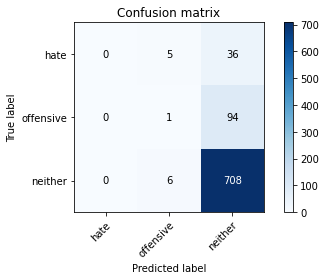
[ 0 6 708]]

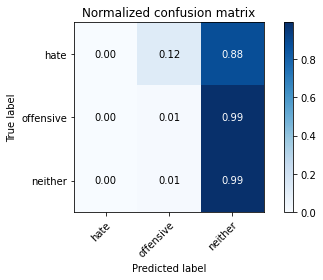
Normalized confusion matrix

[[0. 0.12195122 0.87804878]

[0. 0.01052632 0.98947368]

[0. 0.00840336 0.99159664]]





/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 0.01 0.02 95

neither 0.84 0.99 0.91 714

accuracy 0.83 850

macro avg 0.31 0.33 0.31 850

weighted avg 0.72 0.83 0.77 850

average eval\_loss: 0.6022140645357066, accuracy: 0.8341176470588235

training 2 epoch...

11.162645101547241 seconds used......

2 training finished! average train loss: 0.6961602761195256

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 0 95]

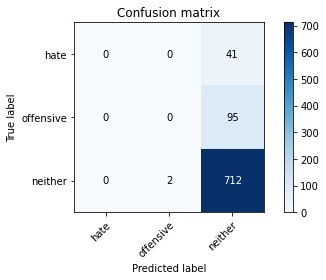
[ 0 2 712]]

Normalized confusion matrix

[[0. 0. 1. ]

[0. 0. 1. ]

[0. 0.00280112 0.99719888]]





/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 0.00 0.00 0.00 95

neither 0.84 1.00 0.91 714

accuracy 0.84 850

macro avg 0.28 0.33 0.30 850

weighted avg 0.71 0.84 0.77 850

average eval\_loss: 0.5094559925586678, accuracy: 0.8376470588235294

training 3 epoch...

11.215854406356812 seconds used......

3 training finished! average train loss: 0.5693644725359404

evaluating...

Confusion matrix, without normalization

[[ 0 13 28]

[ 2 27 66]

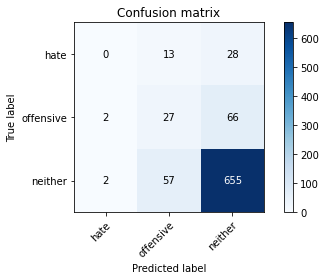
[ 2 57 655]]

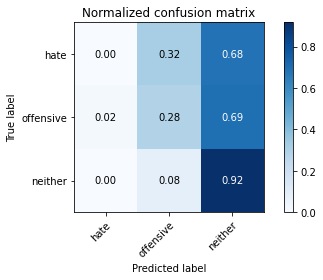
Normalized confusion matrix

[[0. 0.31707317 0.68292683]

[0.02105263 0.28421053 0.69473684]

[0.00280112 0.07983193 0.91736695]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.28 0.28 0.28 95

neither 0.87 0.92 0.90 714

accuracy 0.80 850

macro avg 0.38 0.40 0.39 850

weighted avg 0.77 0.80 0.78 850

average eval\_loss: 0.5382706604378168, accuracy: 0.8023529411764706

training 4 epoch...

11.120796918869019 seconds used......

4 training finished! average train loss: 0.3765780189862618

evaluating...

Confusion matrix, without normalization

[[ 4 7 30]

[ 4 28 63]

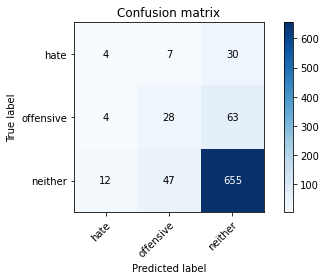
[ 12 47 655]]

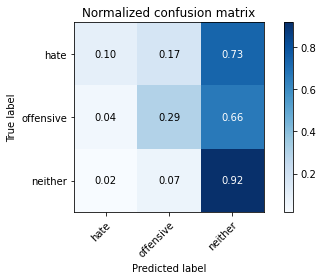
Normalized confusion matrix

[[0.09756098 0.17073171 0.73170732]

[0.04210526 0.29473684 0.66315789]

[0.01680672 0.06582633 0.91736695]]





precision recall f1-score support

hate 0.20 0.10 0.13 41

offensive 0.34 0.29 0.32 95

neither 0.88 0.92 0.90 714

accuracy 0.81 850

macro avg 0.47 0.44 0.45 850

weighted avg 0.78 0.81 0.79 850

average eval\_loss: 0.519029707842788, accuracy: 0.808235294117647

training 5 epoch...

11.176037073135376 seconds used......

5 training finished! average train loss: 0.19790949546373807

evaluating...

Confusion matrix, without normalization

[[ 9 2 30]

[ 6 21 68]

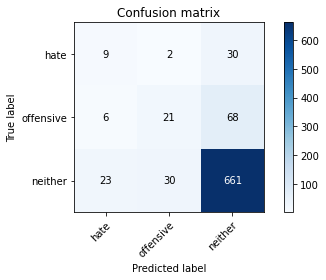
[ 23 30 661]]

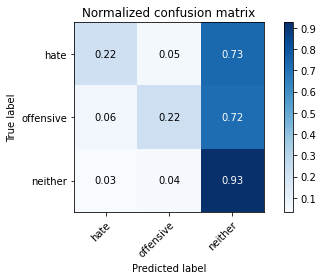
Normalized confusion matrix

[[0.2195122 0.04878049 0.73170732]

[0.06315789 0.22105263 0.71578947]

[0.03221289 0.04201681 0.92577031]]





precision recall f1-score support

hate 0.24 0.22 0.23 41

offensive 0.40 0.22 0.28 95

neither 0.87 0.93 0.90 714

accuracy 0.81 850

macro avg 0.50 0.46 0.47 850

weighted avg 0.79 0.81 0.80 850

average eval\_loss: 0.6296720018036404, accuracy: 0.8129411764705883

training 6 epoch...

11.103230714797974 seconds used......

6 training finished! average train loss: 0.08227018563506695

evaluating...

Confusion matrix, without normalization

[[ 11 0 30]

[ 13 9 73]

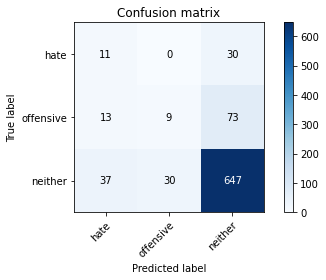
[ 37 30 647]]

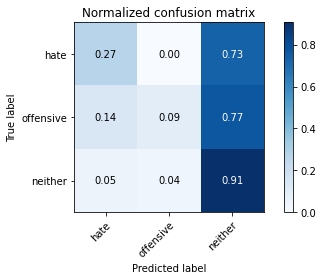
Normalized confusion matrix

[[0.26829268 0. 0.73170732]

[0.13684211 0.09473684 0.76842105]

[0.05182073 0.04201681 0.90616246]]





precision recall f1-score support

hate 0.18 0.27 0.22 41

offensive 0.23 0.09 0.13 95

neither 0.86 0.91 0.88 714

accuracy 0.78 850

macro avg 0.42 0.42 0.41 850

weighted avg 0.76 0.78 0.77 850

average eval\_loss: 0.8569301651888203, accuracy: 0.7847058823529411

training 7 epoch...

11.064327955245972 seconds used......

7 training finished! average train loss: 0.055553764498864226

evaluating...

Confusion matrix, without normalization

[[ 7 17 17]

[ 10 63 22]

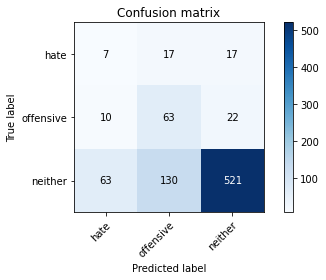
[ 63 130 521]]

Normalized confusion matrix

[[0.17073171 0.41463415 0.41463415]

[0.10526316 0.66315789 0.23157895]

[0.08823529 0.18207283 0.72969188]]





precision recall f1-score support

hate 0.09 0.17 0.12 41

offensive 0.30 0.66 0.41 95

neither 0.93 0.73 0.82 714

accuracy 0.70 850

macro avg 0.44 0.52 0.45 850

weighted avg 0.82 0.70 0.74 850

average eval\_loss: 1.120804090069087, accuracy: 0.6952941176470588

training 8 epoch...

11.092669248580933 seconds used......

8 training finished! average train loss: 0.023620272531675604

evaluating...

Confusion matrix, without normalization

[[ 7 7 27]

[ 8 31 56]

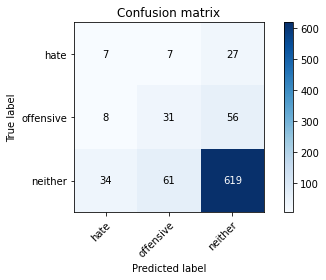
[ 34 61 619]]

Normalized confusion matrix

[[0.17073171 0.17073171 0.65853659]

[0.08421053 0.32631579 0.58947368]

[0.04761905 0.08543417 0.86694678]]





precision recall f1-score support

hate 0.14 0.17 0.16 41

offensive 0.31 0.33 0.32 95

neither 0.88 0.87 0.87 714

accuracy 0.77 850

macro avg 0.45 0.45 0.45 850

weighted avg 0.78 0.77 0.78 850

average eval\_loss: 0.9244850448097189, accuracy: 0.7729411764705882

training 9 epoch...

11.161595582962036 seconds used......

9 training finished! average train loss: 0.0085433493511608

evaluating...

Confusion matrix, without normalization

[[ 3 4 34]

[ 8 19 68]

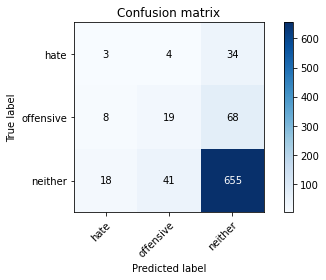
[ 18 41 655]]

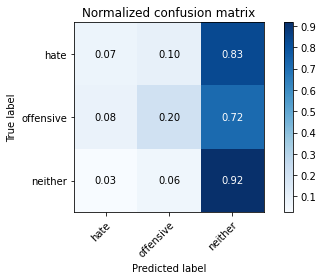
Normalized confusion matrix

[[0.07317073 0.09756098 0.82926829]

[0.08421053 0.2 0.71578947]

[0.02521008 0.05742297 0.91736695]]





precision recall f1-score support

hate 0.10 0.07 0.09 41

offensive 0.30 0.20 0.24 95

neither 0.87 0.92 0.89 714

accuracy 0.80 850

macro avg 0.42 0.40 0.41 850

weighted avg 0.76 0.80 0.78 850

average eval\_loss: 0.9732717604086691, accuracy: 0.7964705882352942

training 10 epoch...

11.122387886047363 seconds used......

10 training finished! average train loss: 0.003286699602344575

evaluating...

Confusion matrix, without normalization

[[ 7 3 31]

[ 10 22 63]

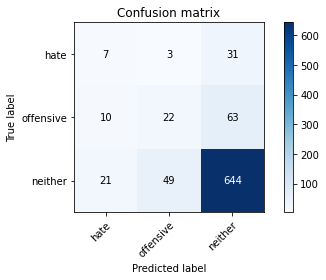
[ 21 49 644]]

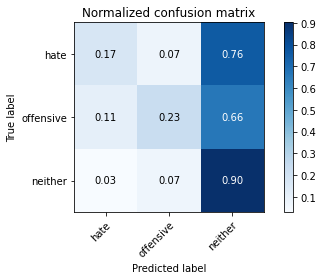
Normalized confusion matrix

[[0.17073171 0.07317073 0.75609756]

[0.10526316 0.23157895 0.66315789]

[0.02941176 0.06862745 0.90196078]]





precision recall f1-score support

hate 0.18 0.17 0.18 41

offensive 0.30 0.23 0.26 95

neither 0.87 0.90 0.89 714

accuracy 0.79 850

macro avg 0.45 0.43 0.44 850

weighted avg 0.78 0.79 0.78 850

average eval\_loss: 1.0055697072790746, accuracy: 0.7917647058823529

add Codeadd Markdown

**Joint training**

In this step, only the total loss which is defined as L = Lclf − α∗Ladv is used for backpropagation. Hence, α is also a import hyperparameter.

add Codeadd Markdown

[18]:

epochs **=** 10

alpha **=** 2

model\_name **=** 'adv\_hate\_as\_noffensive'

​

**for** r **in** range(1,6):

print(str(r)," run......")

classifier **=** Classifier(num\_label).to(DEVICE)

​

adversary **=** Adversary(2, 200).to(DEVICE)

​

loss\_function **=** nn.CrossEntropyLoss()

​

clf\_optimizer **=** AdamW(classifier.parameters(),lr **=** 2e-5, eps **=** 1e-8)

​

adv\_optimizer **=** torch.optim.AdamW(adversary.parameters(), lr**=**0.001)

joint\_training(classifier, adversary, epochs, clf\_optimizer, adv\_optimizer, train\_dataloader, test\_dataloader, alpha, model\_name)

1 run......

training 1 epoch...

12.640432596206665 seconds used......

1 training finished! average train loss: 2.140265173912048

total clf loss: 49.00475215911865 total adv loss: 29.004253059625626

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 1 94]

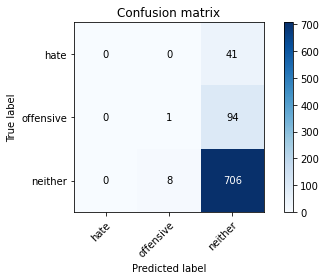
[ 0 8 706]]

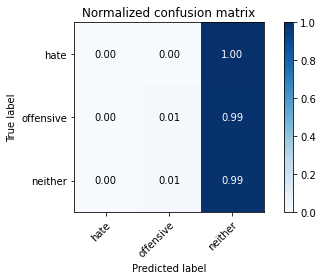
Normalized confusion matrix

[[0. 0. 1. ]

[0. 0.01052632 0.98947368]

[0. 0.01120448 0.98879552]]





/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 0.11 0.01 0.02 95

neither 0.84 0.99 0.91 714

accuracy 0.83 850

macro avg 0.32 0.33 0.31 850

weighted avg 0.72 0.83 0.76 850

average eval\_loss: 0.6681038184567939, accuracy: 0.831764705882353

training 2 epoch...

12.747937440872192 seconds used......

2 training finished! average train loss: 1.7504703736305236

total clf loss: 39.645162642002106 total adv loss: 23.93917815387249

evaluating...

Confusion matrix, without normalization

[[ 0 14 27]

[ 0 34 61]

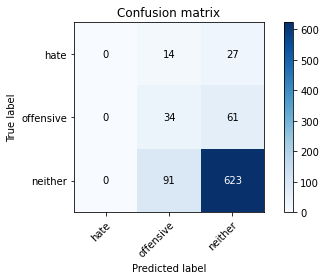
[ 0 91 623]]

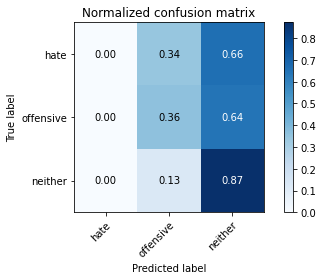
Normalized confusion matrix

[[0. 0.34146341 0.65853659]

[0. 0.35789474 0.64210526]

[0. 0.12745098 0.87254902]]





/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 0.24 0.36 0.29 95

neither 0.88 0.87 0.87 714

accuracy 0.77 850

macro avg 0.37 0.41 0.39 850

weighted avg 0.76 0.77 0.77 850

average eval\_loss: 0.6542204176617223, accuracy: 0.7729411764705882

training 3 epoch...

12.623417615890503 seconds used......

3 training finished! average train loss: 1.3258139562606812

total clf loss: 30.706442207098007 total adv loss: 17.792127832770348

evaluating...

Confusion matrix, without normalization

[[ 0 23 18]

[ 0 67 28]

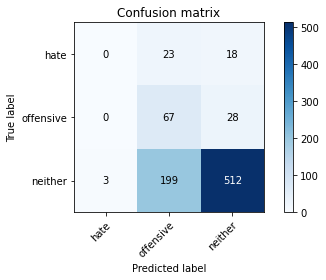
[ 3 199 512]]

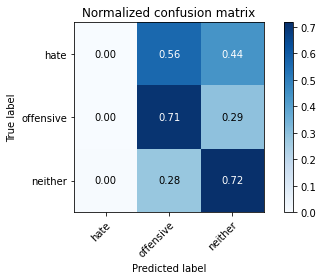
Normalized confusion matrix

[[0. 0.56097561 0.43902439]

[0. 0.70526316 0.29473684]

[0.00420168 0.27871148 0.71708683]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.23 0.71 0.35 95

neither 0.92 0.72 0.81 714

accuracy 0.68 850

macro avg 0.38 0.47 0.38 850

weighted avg 0.80 0.68 0.72 850

average eval\_loss: 0.795125859421353, accuracy: 0.6811764705882353

training 4 epoch...

12.783732175827026 seconds used......

4 training finished! average train loss: 0.8863988637924194

total clf loss: 22.84584730863571 total adv loss: 10.737048007547855

evaluating...

Confusion matrix, without normalization

[[ 0 23 18]

[ 0 71 24]

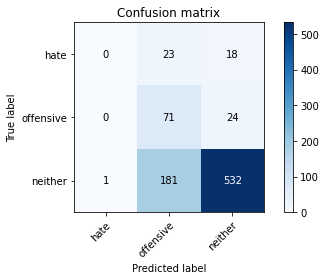
[ 1 181 532]]

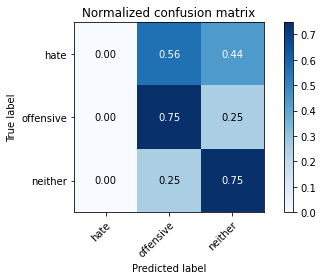
Normalized confusion matrix

[[0. 0.56097561 0.43902439]

[0. 0.74736842 0.25263158]

[0.00140056 0.2535014 0.74509804]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.26 0.75 0.38 95

neither 0.93 0.75 0.83 714

accuracy 0.71 850

macro avg 0.40 0.50 0.40 850

weighted avg 0.81 0.71 0.74 850

average eval\_loss: 0.7489077629911345, accuracy: 0.7094117647058824

training 5 epoch...

12.616151094436646 seconds used......

5 training finished! average train loss: 0.4398459738492966

total clf loss: 15.468628488481045 total adv loss: 3.261835142795462

evaluating...

Confusion matrix, without normalization

[[ 6 18 17]

[ 8 68 19]

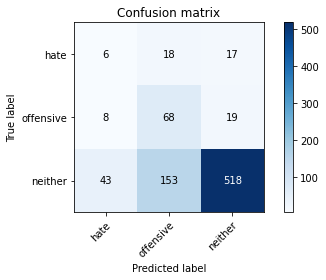
[ 43 153 518]]

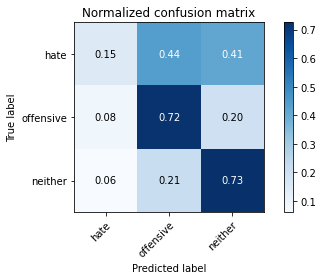
Normalized confusion matrix

[[0.14634146 0.43902439 0.41463415]

[0.08421053 0.71578947 0.2 ]

[0.06022409 0.21428571 0.7254902 ]]





precision recall f1-score support

hate 0.11 0.15 0.12 41

offensive 0.28 0.72 0.41 95

neither 0.94 0.73 0.82 714

accuracy 0.70 850

macro avg 0.44 0.53 0.45 850

weighted avg 0.82 0.70 0.74 850

average eval\_loss: 0.8643379310189292, accuracy: 0.6964705882352941

training 6 epoch...

12.758370637893677 seconds used......

6 training finished! average train loss: 0.2734730992466211

total clf loss: 11.939028155058622 total adv loss: 0.8673134327982552

evaluating...

Confusion matrix, without normalization

[[ 6 14 21]

[ 7 62 26]

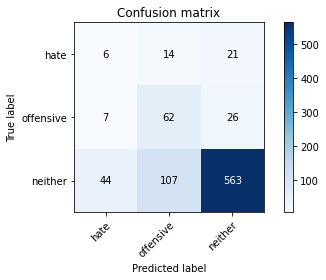
[ 44 107 563]]

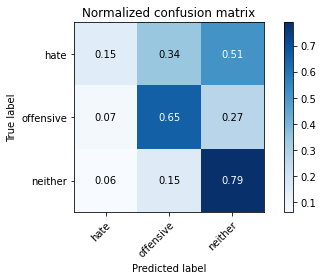
Normalized confusion matrix

[[0.14634146 0.34146341 0.51219512]

[0.07368421 0.65263158 0.27368421]

[0.06162465 0.14985994 0.78851541]]





precision recall f1-score support

hate 0.11 0.15 0.12 41

offensive 0.34 0.65 0.45 95

neither 0.92 0.79 0.85 714

accuracy 0.74 850

macro avg 0.46 0.53 0.47 850

weighted avg 0.82 0.74 0.77 850

average eval\_loss: 0.761579254122321, accuracy: 0.7423529411764705

training 7 epoch...

12.646974802017212 seconds used......

7 training finished! average train loss: 0.21140586502850056

total clf loss: 9.14667671173811 total adv loss: 0.7118082855631656

evaluating...

Confusion matrix, without normalization

[[ 7 15 19]

[ 8 59 28]

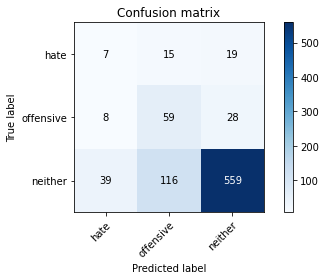
[ 39 116 559]]

Normalized confusion matrix

[[0.17073171 0.36585366 0.46341463]

[0.08421053 0.62105263 0.29473684]

[0.05462185 0.16246499 0.78291317]]





precision recall f1-score support

hate 0.13 0.17 0.15 41

offensive 0.31 0.62 0.41 95

neither 0.92 0.78 0.85 714

accuracy 0.74 850

macro avg 0.45 0.52 0.47 850

weighted avg 0.82 0.74 0.76 850

average eval\_loss: 0.8507759157394947, accuracy: 0.7352941176470589

training 8 epoch...

12.788854598999023 seconds used......

8 training finished! average train loss: 0.1694225939363241

total clf loss: 7.714941669255495 total adv loss: 0.37809403474238934

evaluating...

Confusion matrix, without normalization

[[ 11 14 16]

[ 20 55 20]

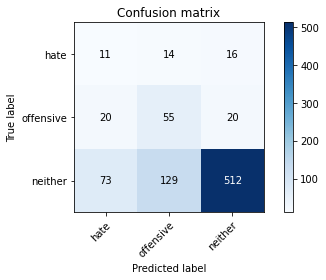
[ 73 129 512]]

Normalized confusion matrix

[[0.26829268 0.34146341 0.3902439 ]

[0.21052632 0.57894737 0.21052632]

[0.1022409 0.18067227 0.71708683]]





precision recall f1-score support

hate 0.11 0.27 0.15 41

offensive 0.28 0.58 0.38 95

neither 0.93 0.72 0.81 714

accuracy 0.68 850

macro avg 0.44 0.52 0.45 850

weighted avg 0.82 0.68 0.73 850

average eval\_loss: 1.0059515387686186, accuracy: 0.68

training 9 epoch...

12.667274713516235 seconds used......

9 training finished! average train loss: 0.12414983130991458

total clf loss: 5.745239516720176 total adv loss: 0.2311260437272722

evaluating...

Confusion matrix, without normalization

[[ 10 18 13]

[ 10 67 18]

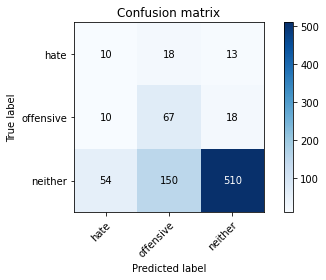
[ 54 150 510]]

Normalized confusion matrix

[[0.24390244 0.43902439 0.31707317]

[0.10526316 0.70526316 0.18947368]

[0.07563025 0.21008403 0.71428571]]





precision recall f1-score support

hate 0.14 0.24 0.17 41

offensive 0.29 0.71 0.41 95

neither 0.94 0.71 0.81 714

accuracy 0.69 850

macro avg 0.45 0.55 0.46 850

weighted avg 0.83 0.69 0.74 850

average eval\_loss: 1.138072766147034, accuracy: 0.6905882352941176

training 10 epoch...

12.668768167495728 seconds used......

10 training finished! average train loss: 0.13604472594335676

total clf loss: 5.618199126794934 total adv loss: 0.5920185524446424

evaluating...

Confusion matrix, without normalization

[[ 8 12 21]

[ 13 50 32]

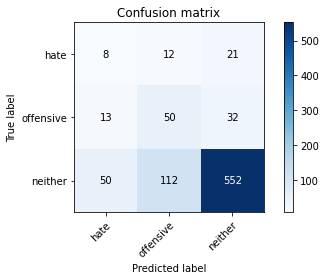
[ 50 112 552]]

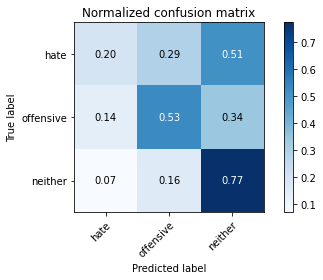
Normalized confusion matrix

[[0.19512195 0.29268293 0.51219512]

[0.13684211 0.52631579 0.33684211]

[0.07002801 0.15686275 0.77310924]]





precision recall f1-score support

hate 0.11 0.20 0.14 41

offensive 0.29 0.53 0.37 95

neither 0.91 0.77 0.84 714

accuracy 0.72 850

macro avg 0.44 0.50 0.45 850

weighted avg 0.80 0.72 0.75 850

average eval\_loss: 0.9893807482494171, accuracy: 0.7176470588235294

2 run......

training 1 epoch...

12.715333938598633 seconds used......

1 training finished! average train loss: 2.176018545627594

total clf loss: 49.68963676691055 total adv loss: 29.555644989013672

evaluating...

Confusion matrix, without normalization

[[ 0 16 25]

[ 0 8 87]

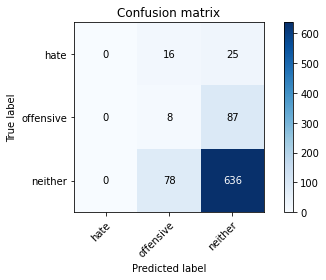
[ 0 78 636]]

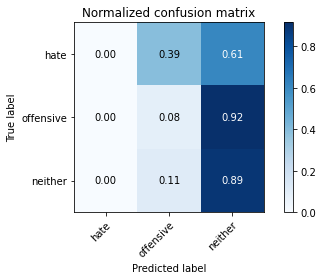
Normalized confusion matrix

[[0. 0.3902439 0.6097561 ]

[0. 0.08421053 0.91578947]

[0. 0.1092437 0.8907563 ]]





/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 0.08 0.08 95

neither 0.85 0.89 0.87 714

accuracy 0.76 850

macro avg 0.31 0.32 0.32 850

weighted avg 0.72 0.76 0.74 850

average eval\_loss: 0.7360936285451402, accuracy: 0.7576470588235295

training 2 epoch...

12.67758584022522 seconds used......

2 training finished! average train loss: 1.7905834436416626

total clf loss: 40.094738721847534 total adv loss: 24.71721686422825

evaluating...

Confusion matrix, without normalization

[[ 0 21 20]

[ 0 56 39]

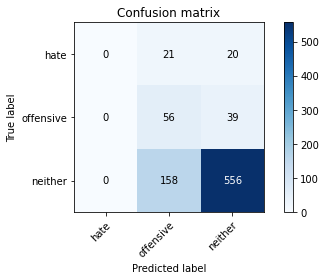
[ 0 158 556]]

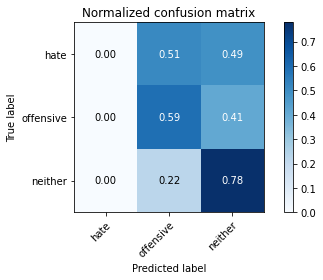
Normalized confusion matrix

[[0. 0.51219512 0.48780488]

[0. 0.58947368 0.41052632]

[0. 0.22128852 0.77871148]]





/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 0.24 0.59 0.34 95

neither 0.90 0.78 0.84 714

accuracy 0.72 850

macro avg 0.38 0.46 0.39 850

weighted avg 0.79 0.72 0.74 850

average eval\_loss: 0.7159008674843367, accuracy: 0.72

training 3 epoch...

12.62491226196289 seconds used......

3 training finished! average train loss: 1.316790807247162

total clf loss: 30.94277185201645 total adv loss: 17.448384284973145

evaluating...

Confusion matrix, without normalization

[[ 0 17 24]

[ 0 55 40]

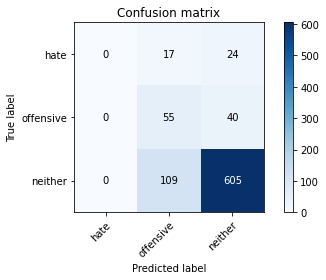
[ 0 109 605]]

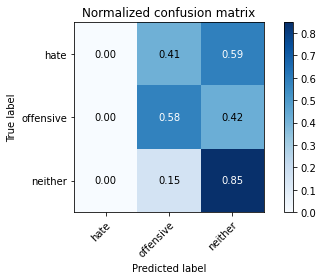
Normalized confusion matrix

[[0. 0.41463415 0.58536585]

[0. 0.57894737 0.42105263]

[0. 0.15266106 0.84733894]]





/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 0.30 0.58 0.40 95

neither 0.90 0.85 0.87 714

accuracy 0.78 850

macro avg 0.40 0.48 0.42 850

weighted avg 0.79 0.78 0.78 850

average eval\_loss: 0.6020023546853038, accuracy: 0.7764705882352941

training 4 epoch...

12.684543132781982 seconds used......

4 training finished! average train loss: 0.7936732977628708

total clf loss: 21.765369400382042 total adv loss: 8.959147840738297

evaluating...

Confusion matrix, without normalization

[[ 3 14 24]

[ 2 53 40]

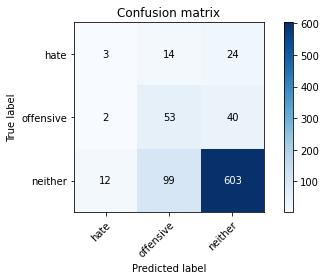
[ 12 99 603]]

Normalized confusion matrix

[[0.07317073 0.34146341 0.58536585]

[0.02105263 0.55789474 0.42105263]

[0.01680672 0.13865546 0.84453782]]





precision recall f1-score support

hate 0.18 0.07 0.10 41

offensive 0.32 0.56 0.41 95

neither 0.90 0.84 0.87 714

accuracy 0.78 850

macro avg 0.47 0.49 0.46 850

weighted avg 0.80 0.78 0.78 850

average eval\_loss: 0.6227547291579635, accuracy: 0.7752941176470588

training 5 epoch...

12.63560152053833 seconds used......

5 training finished! average train loss: 0.4521890737116337

total clf loss: 14.958425655961037 total adv loss: 3.825513922609389

evaluating...

Confusion matrix, without normalization

[[ 5 18 18]

[ 6 68 21]

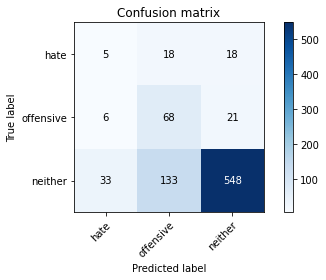
[ 33 133 548]]

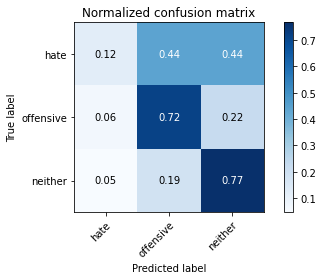
Normalized confusion matrix

[[0.12195122 0.43902439 0.43902439]

[0.06315789 0.71578947 0.22105263]

[0.04621849 0.18627451 0.767507 ]]





precision recall f1-score support

hate 0.11 0.12 0.12 41

offensive 0.31 0.72 0.43 95

neither 0.93 0.77 0.84 714

accuracy 0.73 850

macro avg 0.45 0.54 0.46 850

weighted avg 0.82 0.73 0.76 850

average eval\_loss: 0.7451909116014491, accuracy: 0.7305882352941176

training 6 epoch...

12.690715312957764 seconds used......

6 training finished! average train loss: 0.2920872385799885

total clf loss: 11.122239090502262 total adv loss: 1.7410614229156636

evaluating...

Confusion matrix, without normalization

[[ 6 19 16]

[ 7 66 22]

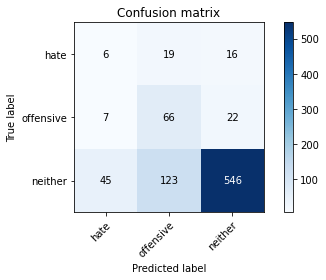
[ 45 123 546]]

Normalized confusion matrix

[[0.14634146 0.46341463 0.3902439 ]

[0.07368421 0.69473684 0.23157895]

[0.06302521 0.17226891 0.76470588]]





precision recall f1-score support

hate 0.10 0.15 0.12 41

offensive 0.32 0.69 0.44 95

neither 0.93 0.76 0.84 714

accuracy 0.73 850

macro avg 0.45 0.54 0.47 850

weighted avg 0.83 0.73 0.76 850

average eval\_loss: 0.8181774701576593, accuracy: 0.7270588235294118

training 7 epoch...

12.65464472770691 seconds used......

7 training finished! average train loss: 0.16664328664541245

total clf loss: 7.916050355881453 total adv loss: 0.20805696059687762

evaluating...

Confusion matrix, without normalization

[[ 18 6 17]

[ 24 37 34]

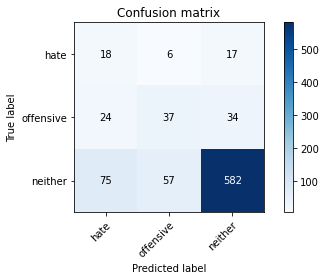
[ 75 57 582]]

Normalized confusion matrix

[[0.43902439 0.14634146 0.41463415]

[0.25263158 0.38947368 0.35789474]

[0.10504202 0.07983193 0.81512605]]





precision recall f1-score support

hate 0.15 0.44 0.23 41

offensive 0.37 0.39 0.38 95

neither 0.92 0.82 0.86 714

accuracy 0.75 850

macro avg 0.48 0.55 0.49 850

weighted avg 0.82 0.75 0.78 850

average eval\_loss: 0.780233160441005, accuracy: 0.7494117647058823

training 8 epoch...

12.732200384140015 seconds used......

8 training finished! average train loss: 0.19509808730334044

total clf loss: 6.826582340523601 total adv loss: 1.464160968484066

evaluating...

Confusion matrix, without normalization

[[ 10 3 28]

[ 8 37 50]

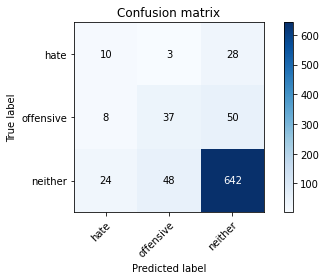
[ 24 48 642]]

Normalized confusion matrix

[[0.24390244 0.07317073 0.68292683]

[0.08421053 0.38947368 0.52631579]

[0.03361345 0.06722689 0.89915966]]





precision recall f1-score support

hate 0.24 0.24 0.24 41

offensive 0.42 0.39 0.40 95

neither 0.89 0.90 0.90 714

accuracy 0.81 850

macro avg 0.52 0.51 0.51 850

weighted avg 0.81 0.81 0.81 850

average eval\_loss: 0.6999699220678575, accuracy: 0.8105882352941176

training 9 epoch...

12.666252374649048 seconds used......

9 training finished! average train loss: 0.2057387324422598

total clf loss: 6.992934027686715 total adv loss: 1.6470012851641513

evaluating...

Confusion matrix, without normalization

[[ 17 14 10]

[ 13 58 24]

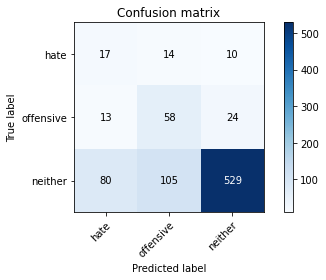
[ 80 105 529]]

Normalized confusion matrix

[[0.41463415 0.34146341 0.24390244]

[0.13684211 0.61052632 0.25263158]

[0.11204482 0.14705882 0.74089636]]





precision recall f1-score support

hate 0.15 0.41 0.23 41

offensive 0.33 0.61 0.43 95

neither 0.94 0.74 0.83 714

accuracy 0.71 850

macro avg 0.47 0.59 0.49 850

weighted avg 0.83 0.71 0.75 850

average eval\_loss: 0.9959006169161131, accuracy: 0.7105882352941176

training 10 epoch...

12.674255609512329 seconds used......

10 training finished! average train loss: 0.09778008501976729

total clf loss: 4.166393468156457 total adv loss: 0.36130539999794564

evaluating...

Confusion matrix, without normalization

[[ 23 3 15]

[ 29 17 49]

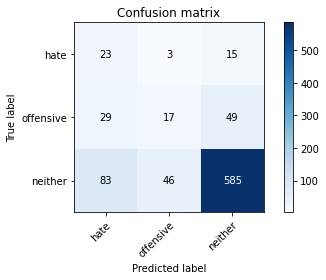
[ 83 46 585]]

Normalized confusion matrix

[[0.56097561 0.07317073 0.36585366]

[0.30526316 0.17894737 0.51578947]

[0.1162465 0.06442577 0.81932773]]





precision recall f1-score support

hate 0.17 0.56 0.26 41

offensive 0.26 0.18 0.21 95

neither 0.90 0.82 0.86 714

accuracy 0.74 850

macro avg 0.44 0.52 0.44 850

weighted avg 0.79 0.74 0.76 850

average eval\_loss: 0.9229776466178686, accuracy: 0.7352941176470589

3 run......

training 1 epoch...

12.644453525543213 seconds used......

1 training finished! average train loss: 2.1563890528678895

total clf loss: 49.07576537132263 total adv loss: 29.371843695640564

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 1 94]

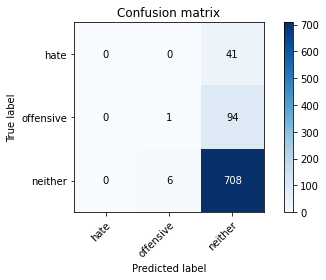
[ 0 6 708]]

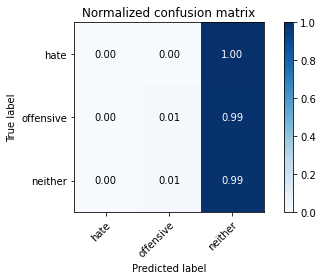
Normalized confusion matrix

[[0. 0. 1. ]

[0. 0.01052632 0.98947368]

[0. 0.00840336 0.99159664]]





/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 0.14 0.01 0.02 95

neither 0.84 0.99 0.91 714

accuracy 0.83 850

macro avg 0.33 0.33 0.31 850

weighted avg 0.72 0.83 0.77 850

average eval\_loss: 0.6449354703343192, accuracy: 0.8341176470588235

training 2 epoch...

12.663822412490845 seconds used......

2 training finished! average train loss: 1.6923502123355865

total clf loss: 38.2829754948616 total adv loss: 23.16726726293564

evaluating...

Confusion matrix, without normalization

[[ 0 17 24]

[ 0 56 39]

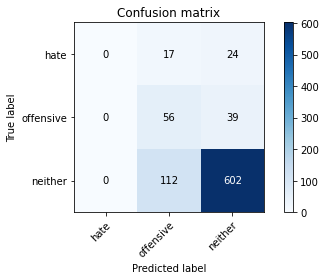
[ 0 112 602]]

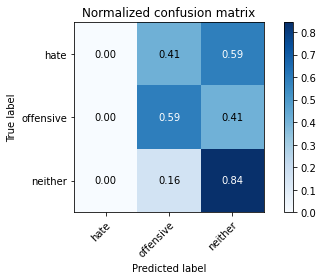
Normalized confusion matrix

[[0. 0.41463415 0.58536585]

[0. 0.58947368 0.41052632]

[0. 0.15686275 0.84313725]]





/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 0.30 0.59 0.40 95

neither 0.91 0.84 0.87 714

accuracy 0.77 850

macro avg 0.40 0.48 0.42 850

weighted avg 0.79 0.77 0.78 850

average eval\_loss: 0.6283843858644019, accuracy: 0.7741176470588236

training 3 epoch...

12.66604208946228 seconds used......

3 training finished! average train loss: 1.2684740042686462

total clf loss: 30.126335069537163 total adv loss: 16.648682579398155

evaluating...

Confusion matrix, without normalization

[[ 0 15 26]

[ 0 54 41]

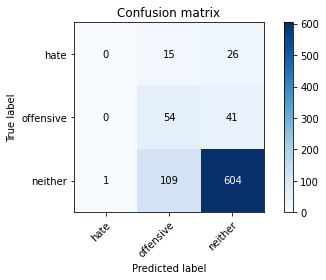
[ 1 109 604]]

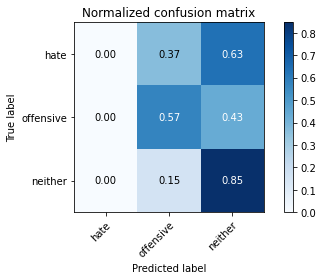
Normalized confusion matrix

[[0. 0.36585366 0.63414634]

[0. 0.56842105 0.43157895]

[0.00140056 0.15266106 0.84593838]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.30 0.57 0.40 95

neither 0.90 0.85 0.87 714

accuracy 0.77 850

macro avg 0.40 0.47 0.42 850

weighted avg 0.79 0.77 0.78 850

average eval\_loss: 0.6128189623529134, accuracy: 0.7741176470588236

training 4 epoch...

12.714701890945435 seconds used......

4 training finished! average train loss: 0.8130794113874436

total clf loss: 21.728677600622177 total adv loss: 9.462646331638098

evaluating...

Confusion matrix, without normalization

[[ 1 18 22]

[ 1 65 29]

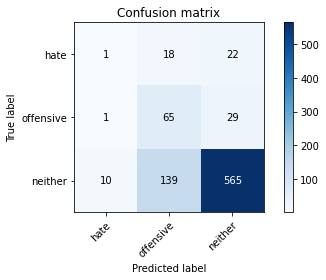
[ 10 139 565]]

Normalized confusion matrix

[[0.02439024 0.43902439 0.53658537]

[0.01052632 0.68421053 0.30526316]

[0.0140056 0.19467787 0.79131653]]





precision recall f1-score support

hate 0.08 0.02 0.04 41

offensive 0.29 0.68 0.41 95

neither 0.92 0.79 0.85 714

accuracy 0.74 850

macro avg 0.43 0.50 0.43 850

weighted avg 0.81 0.74 0.76 850

average eval\_loss: 0.6792915640182273, accuracy: 0.7423529411764705

training 5 epoch...

12.652348756790161 seconds used......

5 training finished! average train loss: 0.4376115620136261

total clf loss: 15.0972261428833 total adv loss: 3.391675940947607

evaluating...

Confusion matrix, without normalization

[[ 3 16 22]

[ 1 72 22]

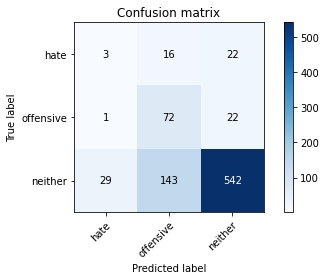
[ 29 143 542]]

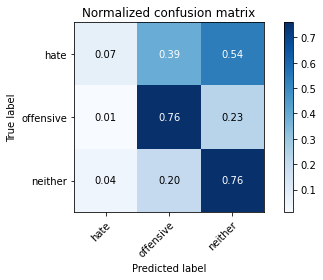
Normalized confusion matrix

[[0.07317073 0.3902439 0.53658537]

[0.01052632 0.75789474 0.23157895]

[0.04061625 0.20028011 0.75910364]]





precision recall f1-score support

hate 0.09 0.07 0.08 41

offensive 0.31 0.76 0.44 95

neither 0.92 0.76 0.83 714

accuracy 0.73 850

macro avg 0.44 0.53 0.45 850

weighted avg 0.82 0.73 0.75 850

average eval\_loss: 0.792572307292112, accuracy: 0.7258823529411764

training 6 epoch...

12.683165550231934 seconds used......

6 training finished! average train loss: 0.28894675940275194

total clf loss: 11.79271199554205 total adv loss: 1.3273129201261327

evaluating...

Confusion matrix, without normalization

[[ 9 7 25]

[ 12 43 40]

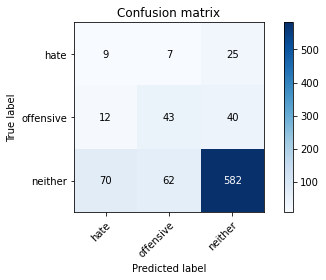
[ 70 62 582]]

Normalized confusion matrix

[[0.2195122 0.17073171 0.6097561 ]

[0.12631579 0.45263158 0.42105263]

[0.09803922 0.08683473 0.81512605]]





precision recall f1-score support

hate 0.10 0.22 0.14 41

offensive 0.38 0.45 0.42 95

neither 0.90 0.82 0.86 714

accuracy 0.75 850

macro avg 0.46 0.50 0.47 850

weighted avg 0.80 0.75 0.77 850

average eval\_loss: 0.7356743567384952, accuracy: 0.7458823529411764

training 7 epoch...

12.623869180679321 seconds used......

7 training finished! average train loss: 0.19478605039417743

total clf loss: 8.989633332937956 total adv loss: 0.3748346061511256

evaluating...

Confusion matrix, without normalization

[[ 5 10 26]

[ 8 52 35]

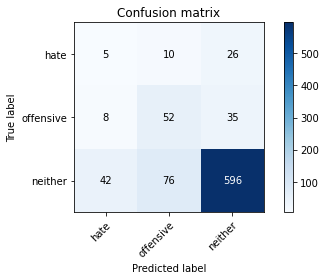
[ 42 76 596]]

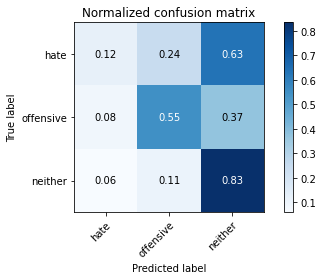
Normalized confusion matrix

[[0.12195122 0.24390244 0.63414634]

[0.08421053 0.54736842 0.36842105]

[0.05882353 0.10644258 0.83473389]]





precision recall f1-score support

hate 0.09 0.12 0.10 41

offensive 0.38 0.55 0.45 95

neither 0.91 0.83 0.87 714

accuracy 0.77 850

macro avg 0.46 0.50 0.47 850

weighted avg 0.81 0.77 0.79 850

average eval\_loss: 0.7182513871121892, accuracy: 0.768235294117647

training 8 epoch...

12.675272464752197 seconds used......

8 training finished! average train loss: 0.15070260770618915

total clf loss: 6.202897347509861 total adv loss: 0.6661165228942991

evaluating...

Confusion matrix, without normalization

[[ 14 6 21]

[ 27 39 29]

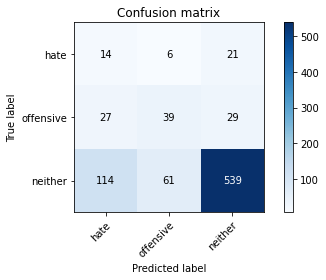
[114 61 539]]

Normalized confusion matrix

[[0.34146341 0.14634146 0.51219512]

[0.28421053 0.41052632 0.30526316]

[0.15966387 0.08543417 0.75490196]]





precision recall f1-score support

hate 0.09 0.34 0.14 41

offensive 0.37 0.41 0.39 95

neither 0.92 0.75 0.83 714

accuracy 0.70 850

macro avg 0.46 0.50 0.45 850

weighted avg 0.81 0.70 0.75 850

average eval\_loss: 0.9544462658205005, accuracy: 0.6964705882352941

training 9 epoch...

12.649204015731812 seconds used......

9 training finished! average train loss: 0.09507606722414494

total clf loss: 4.635088782757521 total adv loss: 0.05935728907570592

evaluating...

Confusion matrix, without normalization

[[ 16 5 20]

[ 26 39 30]

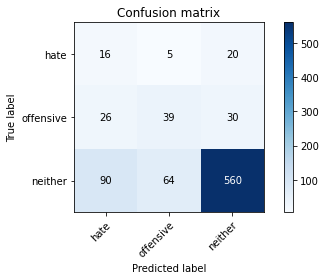
[ 90 64 560]]

Normalized confusion matrix

[[0.3902439 0.12195122 0.48780488]

[0.27368421 0.41052632 0.31578947]

[0.12605042 0.08963585 0.78431373]]





precision recall f1-score support

hate 0.12 0.39 0.18 41

offensive 0.36 0.41 0.38 95

neither 0.92 0.78 0.85 714

accuracy 0.72 850

macro avg 0.47 0.53 0.47 850

weighted avg 0.82 0.72 0.76 850

average eval\_loss: 0.9877639613006004, accuracy: 0.7235294117647059

training 10 epoch...

12.703433752059937 seconds used......

10 training finished! average train loss: 0.06428617609664798

total clf loss: 3.1363957775756717 total adv loss: 0.038956522413172934

evaluating...

Confusion matrix, without normalization

[[ 18 4 19]

[ 27 36 32]

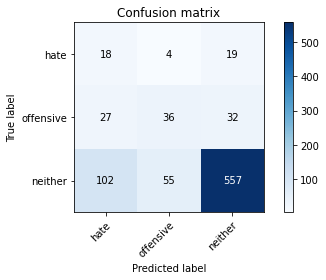
[102 55 557]]

Normalized confusion matrix

[[0.43902439 0.09756098 0.46341463]

[0.28421053 0.37894737 0.33684211]

[0.14285714 0.07703081 0.78011204]]





precision recall f1-score support

hate 0.12 0.44 0.19 41

offensive 0.38 0.38 0.38 95

neither 0.92 0.78 0.84 714

accuracy 0.72 850

macro avg 0.47 0.53 0.47 850

weighted avg 0.82 0.72 0.76 850

average eval\_loss: 1.0655289376804302, accuracy: 0.7188235294117648

4 run......

training 1 epoch...

12.659214735031128 seconds used......

1 training finished! average train loss: 2.091639792919159

total clf loss: 47.04907512664795 total adv loss: 28.766457229852676

evaluating...

Confusion matrix, without normalization

[[ 0 0 41]

[ 0 1 94]

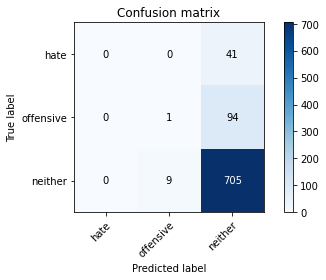
[ 0 9 705]]

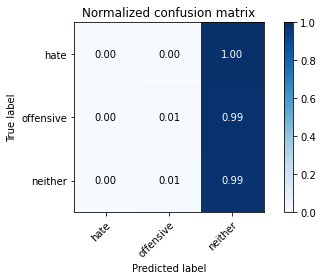
Normalized confusion matrix

[[0. 0. 1. ]

[0. 0.01052632 0.98947368]

[0. 0.01260504 0.98739496]]





/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 0.01 0.02 95

neither 0.84 0.99 0.91 714

accuracy 0.83 850

macro avg 0.31 0.33 0.31 850

weighted avg 0.72 0.83 0.76 850

average eval\_loss: 0.6482944090005963, accuracy: 0.8305882352941176

training 2 epoch...

12.636448860168457 seconds used......

2 training finished! average train loss: 1.690715742111206

total clf loss: 38.65292713046074 total adv loss: 22.941429883241653

evaluating...

Confusion matrix, without normalization

[[ 0 14 27]

[ 0 48 47]

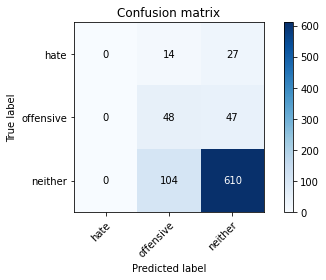
[ 0 104 610]]

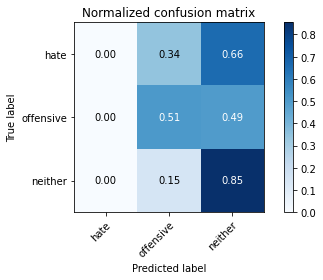
Normalized confusion matrix

[[0. 0.34146341 0.65853659]

[0. 0.50526316 0.49473684]

[0. 0.14565826 0.85434174]]





/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 0.29 0.51 0.37 95

neither 0.89 0.85 0.87 714

accuracy 0.77 850

macro avg 0.39 0.45 0.41 850

weighted avg 0.78 0.77 0.77 850

average eval\_loss: 0.6259668342076069, accuracy: 0.7741176470588236

training 3 epoch...

12.63971495628357 seconds used......

3 training finished! average train loss: 1.2786623418331147

total clf loss: 29.90329721570015 total adv loss: 17.01491016894579

evaluating...

Confusion matrix, without normalization

[[ 1 13 27]

[ 0 50 45]

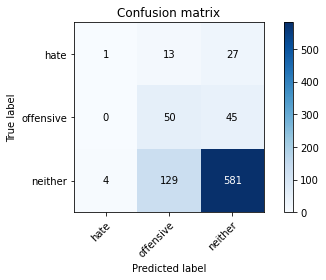
[ 4 129 581]]

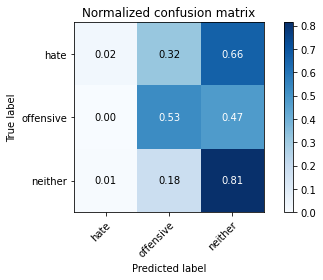
Normalized confusion matrix

[[0.02439024 0.31707317 0.65853659]

[0. 0.52631579 0.47368421]

[0.00560224 0.18067227 0.81372549]]





precision recall f1-score support

hate 0.20 0.02 0.04 41

offensive 0.26 0.53 0.35 95

neither 0.89 0.81 0.85 714

accuracy 0.74 850

macro avg 0.45 0.45 0.41 850

weighted avg 0.79 0.74 0.76 850

average eval\_loss: 0.6492529102536135, accuracy: 0.7435294117647059

training 4 epoch...

12.635771751403809 seconds used......

4 training finished! average train loss: 0.7795588409900666

total clf loss: 20.834475830197334 total adv loss: 9.071733167394996

evaluating...

Confusion matrix, without normalization

[[ 2 11 28]

[ 4 47 44]

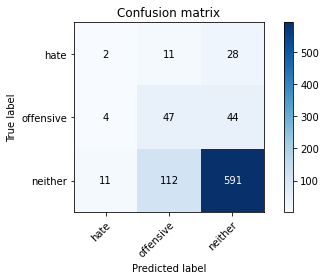
[ 11 112 591]]

Normalized confusion matrix

[[0.04878049 0.26829268 0.68292683]

[0.04210526 0.49473684 0.46315789]

[0.01540616 0.15686275 0.82773109]]





precision recall f1-score support

hate 0.12 0.05 0.07 41

offensive 0.28 0.49 0.35 95

neither 0.89 0.83 0.86 714

accuracy 0.75 850

macro avg 0.43 0.46 0.43 850

weighted avg 0.79 0.75 0.76 850

average eval\_loss: 0.6531987232697565, accuracy: 0.7529411764705882

training 5 epoch...

12.613650798797607 seconds used......

5 training finished! average train loss: 0.36333985984325406

total clf loss: 14.234423160552979 total adv loss: 1.9662849302403629

evaluating...

Confusion matrix, without normalization

[[ 5 20 16]

[ 5 68 22]

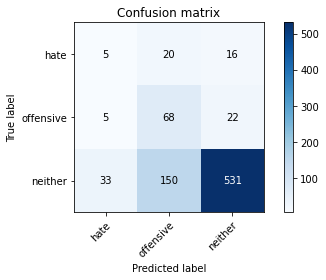
[ 33 150 531]]

Normalized confusion matrix

[[0.12195122 0.48780488 0.3902439 ]

[0.05263158 0.71578947 0.23157895]

[0.04621849 0.21008403 0.74369748]]





precision recall f1-score support

hate 0.12 0.12 0.12 41

offensive 0.29 0.72 0.41 95

neither 0.93 0.74 0.83 714

accuracy 0.71 850

macro avg 0.45 0.53 0.45 850

weighted avg 0.82 0.71 0.75 850

average eval\_loss: 0.8493467840517677, accuracy: 0.7105882352941176

training 6 epoch...

12.626386880874634 seconds used......

6 training finished! average train loss: 0.2954507891833782

total clf loss: 10.99551546573639 total adv loss: 1.8885120271588676

evaluating...

Confusion matrix, without normalization

[[ 15 7 19]

[ 11 63 21]

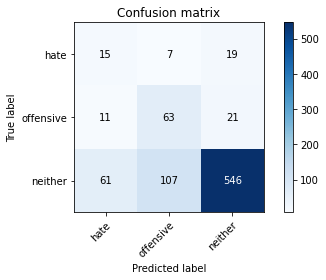
[ 61 107 546]]

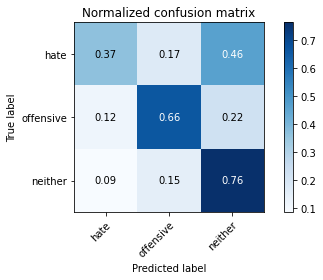
Normalized confusion matrix

[[0.36585366 0.17073171 0.46341463]

[0.11578947 0.66315789 0.22105263]

[0.08543417 0.14985994 0.76470588]]





precision recall f1-score support

hate 0.17 0.37 0.23 41

offensive 0.36 0.66 0.46 95

neither 0.93 0.76 0.84 714

accuracy 0.73 850

macro avg 0.49 0.60 0.51 850

weighted avg 0.83 0.73 0.77 850

average eval\_loss: 0.7866123103385054, accuracy: 0.7341176470588235

training 7 epoch...

12.618418216705322 seconds used......

7 training finished! average train loss: 0.18482181914150714

total clf loss: 8.341172985732555 total adv loss: 0.4499589697952615

evaluating...

Confusion matrix, without normalization

[[ 9 6 26]

[ 9 58 28]

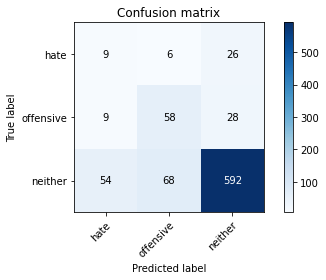
[ 54 68 592]]

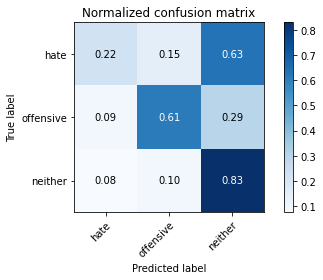
Normalized confusion matrix

[[0.2195122 0.14634146 0.63414634]

[0.09473684 0.61052632 0.29473684]

[0.07563025 0.0952381 0.82913165]]





precision recall f1-score support

hate 0.12 0.22 0.16 41

offensive 0.44 0.61 0.51 95

neither 0.92 0.83 0.87 714

accuracy 0.78 850

macro avg 0.49 0.55 0.51 850

weighted avg 0.82 0.78 0.80 850

average eval\_loss: 0.6825909651815891, accuracy: 0.7752941176470588

training 8 epoch...

12.632163763046265 seconds used......

8 training finished! average train loss: 0.1440836749970913

total clf loss: 6.129329759627581 total adv loss: 0.5374269826788804

evaluating...

Confusion matrix, without normalization

[[ 12 8 21]

[ 5 63 27]

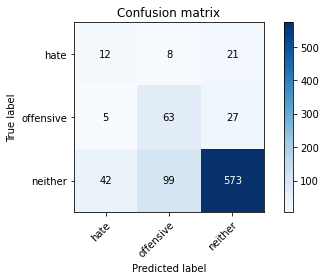
[ 42 99 573]]

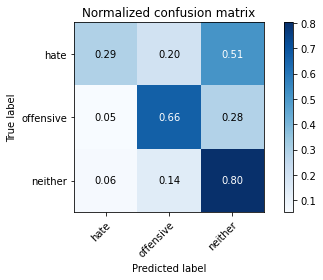
Normalized confusion matrix

[[0.29268293 0.19512195 0.51219512]

[0.05263158 0.66315789 0.28421053]

[0.05882353 0.13865546 0.80252101]]





precision recall f1-score support

hate 0.20 0.29 0.24 41

offensive 0.37 0.66 0.48 95

neither 0.92 0.80 0.86 714

accuracy 0.76 850

macro avg 0.50 0.59 0.52 850

weighted avg 0.83 0.76 0.79 850

average eval\_loss: 0.8164496076683145, accuracy: 0.7623529411764706

training 9 epoch...

12.630355834960938 seconds used......

9 training finished! average train loss: 0.11818711571395397

total clf loss: 5.452725775539875 total adv loss: 0.22831501234031748

evaluating...

Confusion matrix, without normalization

[[ 26 6 9]

[ 24 56 15]

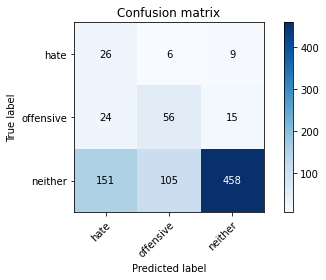
[151 105 458]]

Normalized confusion matrix

[[0.63414634 0.14634146 0.2195122 ]

[0.25263158 0.58947368 0.15789474]

[0.21148459 0.14705882 0.64145658]]





precision recall f1-score support

hate 0.13 0.63 0.21 41

offensive 0.34 0.59 0.43 95

neither 0.95 0.64 0.77 714

accuracy 0.64 850

macro avg 0.47 0.62 0.47 850

weighted avg 0.84 0.64 0.70 850

average eval\_loss: 1.181224420450108, accuracy: 0.6352941176470588

training 10 epoch...

12.649110794067383 seconds used......

10 training finished! average train loss: 0.07821621427312493

total clf loss: 3.8357432493939996 total adv loss: 0.037533736907107595

evaluating...

Confusion matrix, without normalization

[[ 19 6 16]

[ 14 51 30]

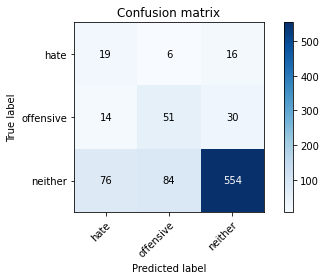
[ 76 84 554]]

Normalized confusion matrix

[[0.46341463 0.14634146 0.3902439 ]

[0.14736842 0.53684211 0.31578947]

[0.10644258 0.11764706 0.77591036]]





precision recall f1-score support

hate 0.17 0.46 0.25 41

offensive 0.36 0.54 0.43 95

neither 0.92 0.78 0.84 714

accuracy 0.73 850

macro avg 0.49 0.59 0.51 850

weighted avg 0.82 0.73 0.77 850

average eval\_loss: 0.9693857192462527, accuracy: 0.7341176470588235

5 run......

training 1 epoch...

12.695341110229492 seconds used......

1 training finished! average train loss: 2.1146429228782653

total clf loss: 47.119808077812195 total adv loss: 29.30616894364357

evaluating...

Confusion matrix, without normalization

[[ 0 4 37]

[ 0 3 92]

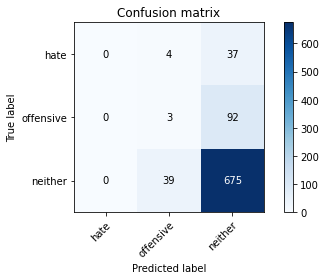
[ 0 39 675]]

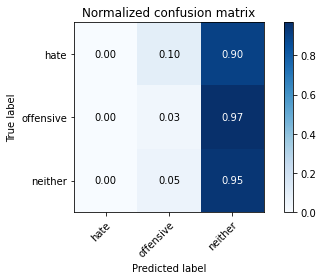
Normalized confusion matrix

[[0. 0.09756098 0.90243902]

[0. 0.03157895 0.96842105]

[0. 0.05462185 0.94537815]]





/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 0.07 0.03 0.04 95

neither 0.84 0.95 0.89 714

accuracy 0.80 850

macro avg 0.30 0.33 0.31 850

weighted avg 0.71 0.80 0.75 850

average eval\_loss: 0.6562414886646493, accuracy: 0.7976470588235294

training 2 epoch...

12.624331951141357 seconds used......

2 training finished! average train loss: 1.720886106491089

total clf loss: 39.72279345989227 total adv loss: 23.160755902528763

evaluating...

Confusion matrix, without normalization

[[ 0 15 26]

[ 0 37 58]

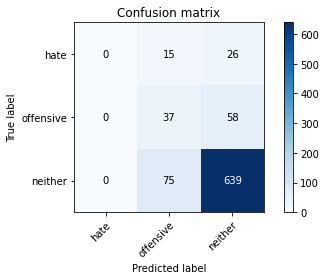
[ 0 75 639]]

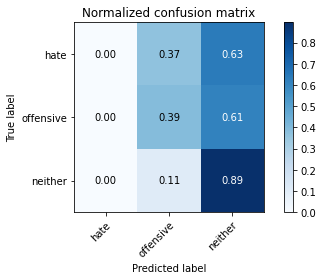
Normalized confusion matrix

[[0. 0.36585366 0.63414634]

[0. 0.38947368 0.61052632]

[0. 0.10504202 0.89495798]]





/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 0.29 0.39 0.33 95

neither 0.88 0.89 0.89 714

accuracy 0.80 850

macro avg 0.39 0.43 0.41 850

weighted avg 0.77 0.80 0.78 850

average eval\_loss: 0.5972155181821003, accuracy: 0.7952941176470588

training 3 epoch...

12.675314903259277 seconds used......

3 training finished! average train loss: 1.297461233139038

total clf loss: 30.705428570508957 total adv loss: 17.083816699683666

evaluating...

Confusion matrix, without normalization

[[ 0 23 18]

[ 0 71 24]

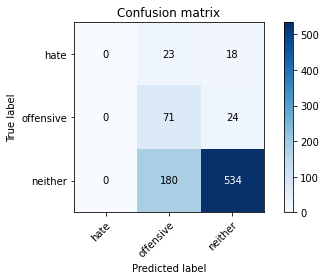
[ 0 180 534]]

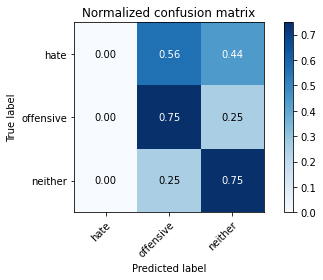
Normalized confusion matrix

[[0. 0.56097561 0.43902439]

[0. 0.74736842 0.25263158]

[0. 0.25210084 0.74789916]]





/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 0.26 0.75 0.38 95

neither 0.93 0.75 0.83 714

accuracy 0.71 850

macro avg 0.40 0.50 0.40 850

weighted avg 0.81 0.71 0.74 850

average eval\_loss: 0.7060049812980863, accuracy: 0.711764705882353

training 4 epoch...

12.628438711166382 seconds used......

4 training finished! average train loss: 0.7445435699820518

total clf loss: 21.066772013902664 total adv loss: 8.080203129444271

evaluating...

Confusion matrix, without normalization

[[ 0 15 26]

[ 0 62 33]

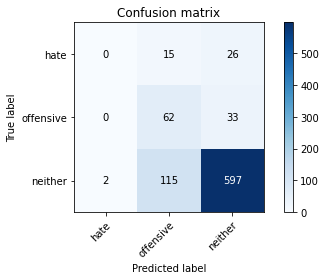
[ 2 115 597]]

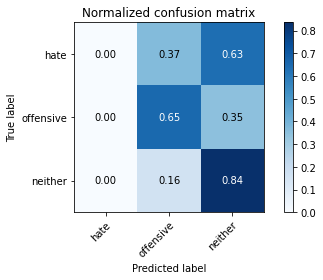
Normalized confusion matrix

[[0. 0.36585366 0.63414634]

[0. 0.65263158 0.34736842]

[0.00280112 0.16106443 0.83613445]]





precision recall f1-score support

hate 0.00 0.00 0.00 41

offensive 0.32 0.65 0.43 95

neither 0.91 0.84 0.87 714

accuracy 0.78 850

macro avg 0.41 0.50 0.43 850

weighted avg 0.80 0.78 0.78 850

average eval\_loss: 0.6151260290901328, accuracy: 0.7752941176470588

training 5 epoch...

12.65387773513794 seconds used......

5 training finished! average train loss: 0.38587033316493036

total clf loss: 14.310022935271263 total adv loss: 2.491746873711236

evaluating...

Confusion matrix, without normalization

[[ 5 13 23]

[ 7 66 22]

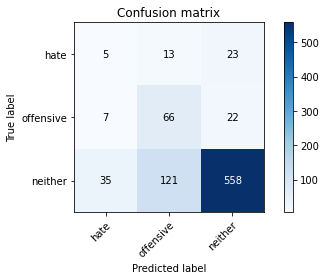
[ 35 121 558]]

Normalized confusion matrix

[[0.12195122 0.31707317 0.56097561]

[0.07368421 0.69473684 0.23157895]

[0.04901961 0.16946779 0.78151261]]





precision recall f1-score support

hate 0.11 0.12 0.11 41

offensive 0.33 0.69 0.45 95

neither 0.93 0.78 0.85 714

accuracy 0.74 850

macro avg 0.45 0.53 0.47 850

weighted avg 0.82 0.74 0.77 850

average eval\_loss: 0.76024865844222, accuracy: 0.74

training 6 epoch...

12.6311936378479 seconds used......

6 training finished! average train loss: 0.2557105890661478

total clf loss: 10.600427828729153 total adv loss: 1.0925507620268036

evaluating...

Confusion matrix, without normalization

[[ 6 5 30]

[ 10 34 51]

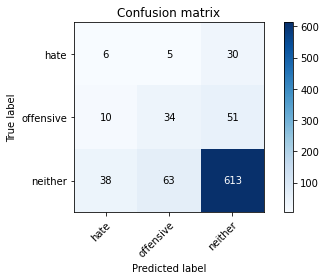
[ 38 63 613]]

Normalized confusion matrix

[[0.14634146 0.12195122 0.73170732]

[0.10526316 0.35789474 0.53684211]

[0.05322129 0.08823529 0.85854342]]





precision recall f1-score support

hate 0.11 0.15 0.13 41

offensive 0.33 0.36 0.35 95

neither 0.88 0.86 0.87 714

accuracy 0.77 850

macro avg 0.44 0.45 0.45 850

weighted avg 0.78 0.77 0.78 850

average eval\_loss: 0.6674906918129255, accuracy: 0.768235294117647

training 7 epoch...

12.634494543075562 seconds used......

7 training finished! average train loss: 0.21981083005666732

total clf loss: 8.799218215048313 total adv loss: 1.095661619285238

evaluating...

Confusion matrix, without normalization

[[ 8 10 23]

[ 14 48 33]

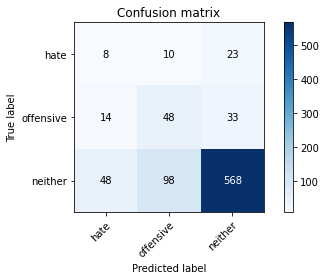
[ 48 98 568]]

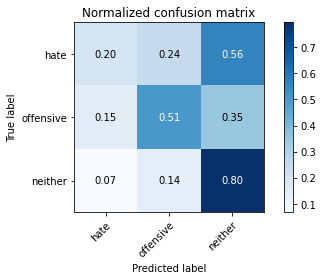
Normalized confusion matrix

[[0.19512195 0.24390244 0.56097561]

[0.14736842 0.50526316 0.34736842]

[0.06722689 0.1372549 0.79551821]]





precision recall f1-score support

hate 0.11 0.20 0.14 41

offensive 0.31 0.51 0.38 95

neither 0.91 0.80 0.85 714

accuracy 0.73 850

macro avg 0.44 0.50 0.46 850

weighted avg 0.80 0.73 0.76 850

average eval\_loss: 0.8189725805680419, accuracy: 0.7341176470588235

training 8 epoch...

12.65627908706665 seconds used......

8 training finished! average train loss: 0.1344757827371359

total clf loss: 6.4824041575193405 total adv loss: 0.12069249499108992

evaluating...

Confusion matrix, without normalization

[[ 12 9 20]

[ 21 41 33]

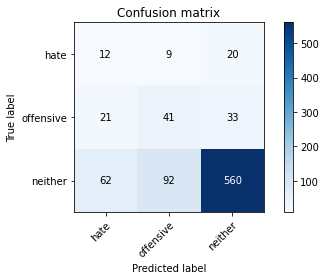
[ 62 92 560]]

Normalized confusion matrix

[[0.29268293 0.2195122 0.48780488]

[0.22105263 0.43157895 0.34736842]

[0.08683473 0.12885154 0.78431373]]





precision recall f1-score support

hate 0.13 0.29 0.18 41

offensive 0.29 0.43 0.35 95

neither 0.91 0.78 0.84 714

accuracy 0.72 850

macro avg 0.44 0.50 0.46 850

weighted avg 0.81 0.72 0.76 850

average eval\_loss: 0.9257628016693648, accuracy: 0.7211764705882353

training 9 epoch...

12.63722038269043 seconds used......

9 training finished! average train loss: 0.11274523310363292

total clf loss: 4.951865391805768 total adv loss: 0.3426981074953801

evaluating...

Confusion matrix, without normalization

[[ 13 11 17]

[ 32 45 18]

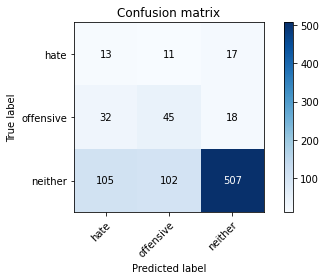
[105 102 507]]

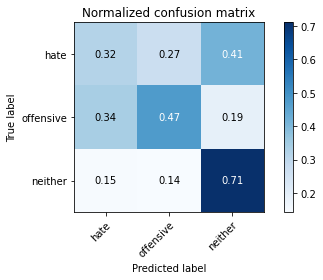
Normalized confusion matrix

[[0.31707317 0.26829268 0.41463415]

[0.33684211 0.47368421 0.18947368]

[0.14705882 0.14285714 0.71008403]]





precision recall f1-score support

hate 0.09 0.32 0.14 41

offensive 0.28 0.47 0.36 95

neither 0.94 0.71 0.81 714

accuracy 0.66 850

macro avg 0.44 0.50 0.43 850

weighted avg 0.82 0.66 0.72 850

average eval\_loss: 1.0789939081079738, accuracy: 0.6647058823529411

training 10 epoch...

12.638134479522705 seconds used......

10 training finished! average train loss: 0.11529332023113965

total clf loss: 4.3700979966670275 total adv loss: 0.6972840081689355

evaluating...

Confusion matrix, without normalization

[[ 14 13 14]

[ 13 63 19]

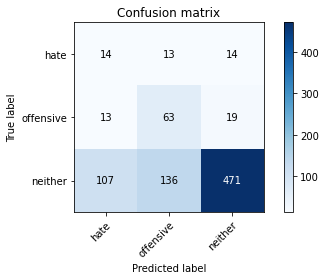
[107 136 471]]

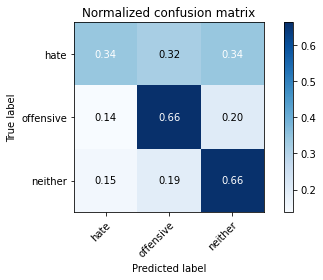
Normalized confusion matrix

[[0.34146341 0.31707317 0.34146341]

[0.13684211 0.66315789 0.2 ]

[0.14985994 0.19047619 0.65966387]]





precision recall f1-score support

hate 0.10 0.34 0.16 41

offensive 0.30 0.66 0.41 95

neither 0.93 0.66 0.77 714

accuracy 0.64 850

macro avg 0.45 0.55 0.45 850

weighted avg 0.82 0.64 0.70 850

average eval\_loss: 1.270661440807893, accuracy: 0.6447058823529411