RUNNING BASELINE MODEL

We are preprocessing the data, basically tokenizing the raw medical reports files into sentences and further tokenizing the sentences into words. So, ultimately the raw medical report files are converted into list of list of words with their corresponding NER type tags (BIO format). There are 29 tags/labels for the task which we mentioned in Milestone 2.

Step 1. How to preprocess the raw text files -

Training files -
$python preprocessing_data.py --dataDir ./train/gold/ --train

After this pre-processing step, a pickle file - train_word_ner_startidx_dict.pickle - is created which has the data in the NER type tag format (BIO) which is used for building the model.

Dev files -
$python preprocessing_data.py --dataDir ./dev/gold/ --dev

After this pre-processing step, a pickle file - dev_word_ner_startidx_dict.pickle - is created which has the data in the NER type tag format (BIO) which is used for building the model.

Test files -
$python preprocessing_data.py --dataDir ./test/gold/ --test

After pre-processing step, a pickle file - test_word_ner_startidx_dict.pickle - is created which has the data in the NER type tag format (BIO) which is used for building the model.

Step 2. How to run baseline model (baseline.py)

baseline.py assumes that the data should be in the current working directory in the same folder hierarchy as is submitted with this Milestone 3.

$python baseline.py

How to evaluate -

Evaluation on Training Data -

$ python evaluate.py brat ner ./train/gold/ ./train/system

Report (SYSTEM: system):

```
------------------------------------------------------------
SubTrack 1 [NER]              Measure      Micro
------------------------------------------------------------
Total (401 docs)             Precision     0.9792
                   Recall        0.9408
                   F1            0.9597
------------------------------------------------------------
```

Evaluation on Dev Data -

$ python evaluate.py brat ner ./dev/gold/ ./dev/system

Report (SYSTEM: system):

```
------------------------------------------------------------
SubTrack 1 [NER]              Measure      Micro
------------------------------------------------------------
Total (193 docs)             Precision     0.8868
                   Recall        0.8199
                   F1            0.852
------------------------------------------------------------
```

Evaluation on Test Data -

$ python evaluate.py brat ner ./test/gold/ ./test/system

Report (SYSTEM: system):

```
------------------------------------------------------------
SubTrack 1 [NER]              Measure      Micro
------------------------------------------------------------
Total (156 docs)             Precision     0.8918
                   Recall        0.8337
                   F1            0.8618
------------------------------------------------------------
```