Sarcasm detection in plain text using deep learning model with Tensorflow

Sanjay Haresh Khatwani (sxk6714), Savitha Jayasankar (skj9180), Saurabh Parekh (sbp4709)

Dependencies:

- Anaconda 4.3.1*
- Python 3.5.x
- TextBlob 0.12.0
- Tensorflow 1.0.1**
- Scikit-learn 0.18.1
- Scipy 0.18.1
- Numpy 1.12.1
- Nltk 3.2.2

There are 4 files in the project:

1. create_feature_sets.py
2. train_and_test.py
3. exp_replace.py
4. Use_NN.py

There are two dataset files in the project:

1. negproc.npy
2. posproc.npy

Feature-sets are stored in featuresets.npy

The model is stored inside folder /model/

Run ***create_feature_sets.py*** to extract features from the two dataset files and get ***featuresets.npy*** file.

Run ***train_and_test.py*** file after the ***create_feature_sets.py*** to use the ***featuresets.npy*** just created and train the neural network. After ***train_and_test.py*** is finished, the model will be saved inside ***/model/*** and can be accessed from there.

exp_replace.py is used by ***create_feature_sets.py*** to preprocess the data.

Use_NN.py can be used after we have model saved inside ***/model/*** to use the neural network to make predictions. The input sentence needs to be supplied as a method argument to *'use_neural_network()'* at the end of the file.

Visualization:

To get visualization in Tensorboard, do the following steps:

- After running ***train_and_test.py***, the logs are collected in ***/tmp/logs/***. Tensorflow uses these logs to generate the visualization.
- Go to terminal, make sure the location is same as the project location. Run the following command there:

  tensorboard --logdir=/tmp/logs

- As part of the output, a URL is provided. The visualization could be accessed by navigating to that URL.

*Install Anaconda: https://docs.continuum.io/anaconda/install

**Install Tensorflow: https://www.tensorflow.org/install/