Future Work

Following is the list of functionalities/modules that will be added in the model:

1. Spell Correction module: This module is required to convert the shorthand notations that usually used in tweets into proper words so as to make them visible to POS\_tag function. This module has to be implemented before the usage of POS\_tagging.
2. Emoticon processor: Emoticons are heavily used in tweets but since they are non-ASCII characters so they are removed in the cleaning module. So they need to be identified before that. Right now we don’t have pre-defined values for these emoticons so we haven’t added such module yet.
3. Apriori Association: We haven’t used any association rule between adverb/verb-adjective pair (or triplet). This method will provide a method that shows how much the adverb/verb affects the following adjective. It will provide stronger lexical affinity between the words and hence better contextual understanding by the machine.
4. kNN Classifier: This classifier has to be used after the scoring module. When we encounter a tweet from the same hashtag we will classify using the kNN classification method into one of the pre-defined classes. The actual use of this kNN classifier will be in a separate application that will use the functionality of all the modules. We can enter a tweet and then compare it with other tweets of the same hash tags and then classify that tweet into one of the five emotion classes.