Emotion Detection in Tweets

12/6/2019

Introduction

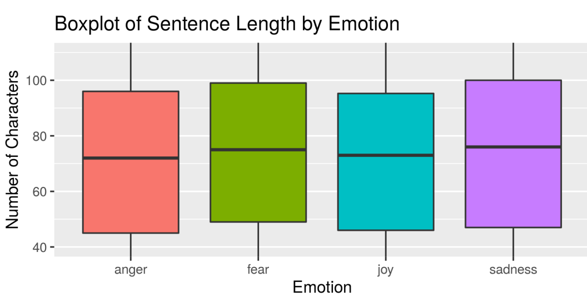
As amount of text data on social media platforms increases rapidly, we can use such data such as tweets to detect public's emotional reaction towards certain topics. In this project, we present an end-to-end model for emotion detection in tweets using R. From text data representation and classification to novel visualizations, we aim to build powerful model to detect emotion in social media text data and present statistical textual information behind each emotion category through visualization.

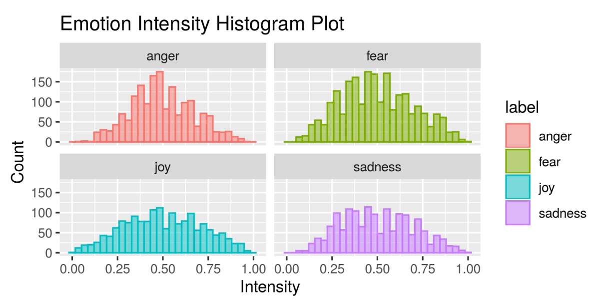
Dataset Overview

We used **Emotion Intensity Dataset** by Mohammad and Bravo-Marquez. It contains 7102 tweets each annotated with single emotion from anger, fear, joy or sadness and an emotion intensity score scaled to 0-1. We use 80% data for training and 20% for testing. Below are some examples from the data:

	tweets	label	intensity
908	Today was horrible and it was only half a day	fear	0.854
3614	@ZubairSabirPTI pls dont insult the word 'Molna'	anger	0.479
2903	Good morning lovely people. Not gonna lie I've woken up feeling pretty glum.	sadness	0.754
6000	We have been a better second half team this season. So there's that.	joy	0.200

From the following plots, we see that emotion intensity of tweets is normally distirbuted in all categories and sentence length distribution is similar across each category.





Preprocessing

We did following preprocessing on our data and store preprocessed data in "preprocessed_data/" folder in .csv format:

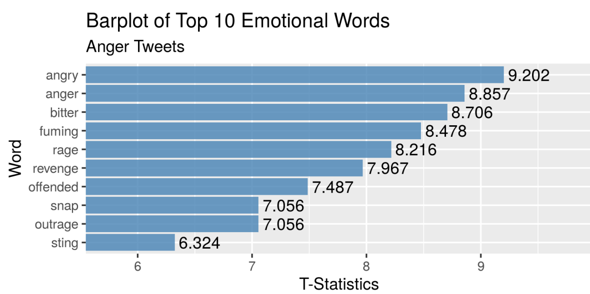
- Change all characters to lower case
- Remove punctuations, redundant white spaces and special characters (excluding "#" for mining hashtags)
- Create a new column "hashtags" which contains all hashtags in each tweet

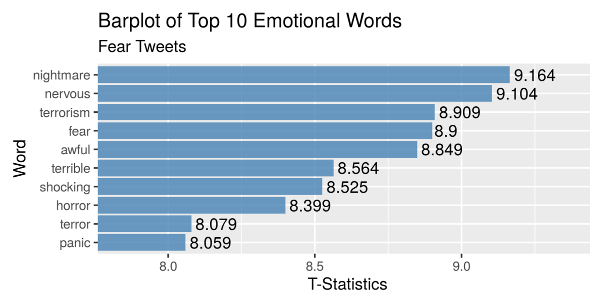
Mining Top Emotional Words & Hashtags

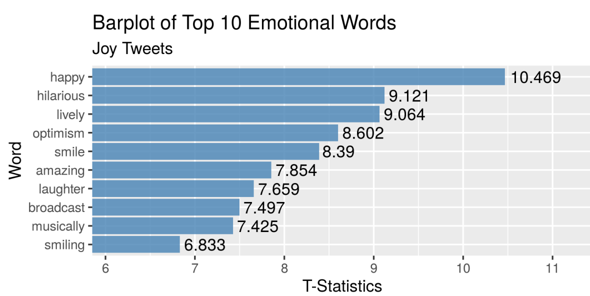
To mine most representitve emotional words in tweets for each emotion, we use two sample t-test of word distribution in tweets of a specific category vs word distribution in tweets not of that category. The larger the t-statistic of a word, the more representitve it is for that category. Using this method, we find top positive and negative words in each category and store them in "emotion_words/" and "emotion_hashtags/" folders. We then use barplot to visualize their ranking and t-statistics.

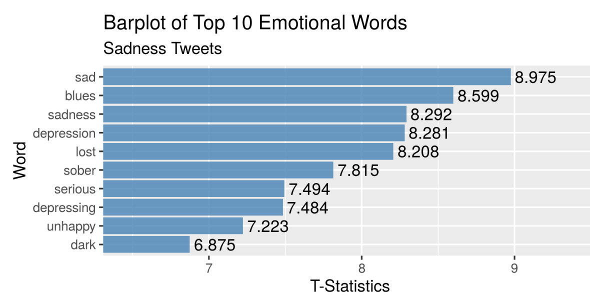
Visualization of Emotional Words

From the following barplots, we see that the top ranked words resemble their emotion category. With the exception of the category 'fear', the top ranked word for each category was a synonym to the category itself. For example, word 'angry' had the highest t-statistics with 9.202 for the category 'anger'. From the results, we can see that our mining method correctly identifies reasonable representative words from emotion tweets.



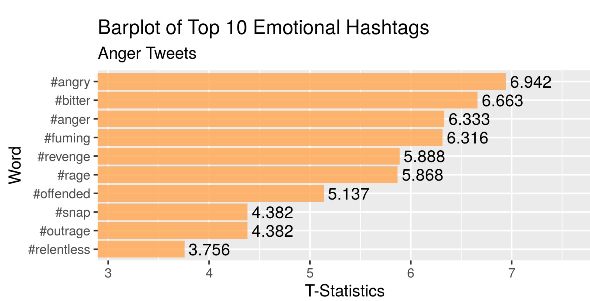


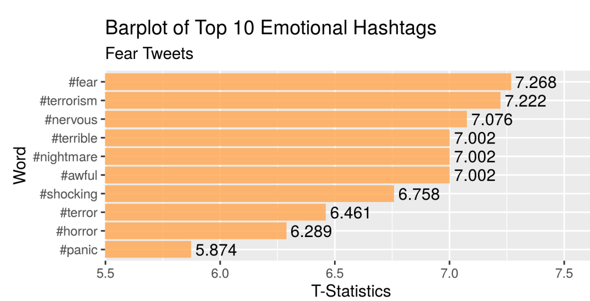


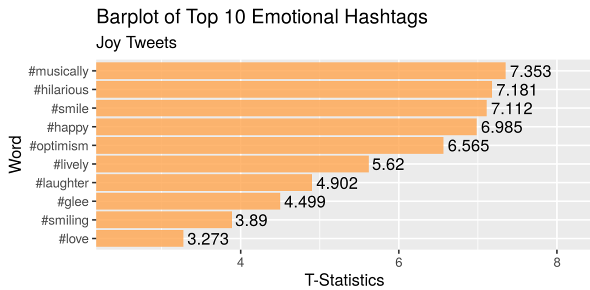


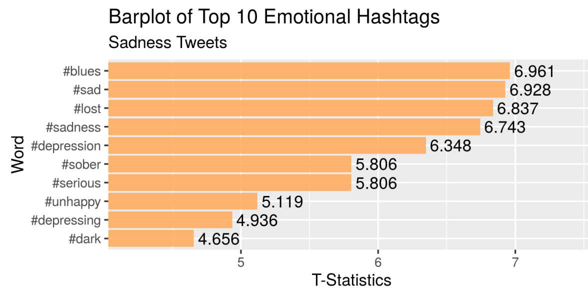
Visualization of Emotional Hashtags

From the following barplots, we see that the top ranked hastags reflect their emotion category. Similar to emotion words above, top ranked hashtags are often the category word itself. For example, '#fear' had the highest t-statistics with 7.268 for the category 'fear'. From the results, we can see that our mining method correctly identifies reasonable representative hashtags from emotion tweets.









Emotion Prediction Model

To build a model for emotion detection in tweets, we fisrt obtain vector representation of each tweet sentence with "fastrtext" package. The vectors are obtained from recent state-of-the-art text representation models with some simple fine-tuning. We then obtain vector-label pairs as our training data and use XGBoost to train our model. Our model achieves following results:

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1   2   3
##          0 160  91  37  58
##          1  67 265  60  59
##          2  33  60 173  46
##          3  54 102  57  99
##
## Overall Statistics
##
##                Accuracy : 0.4905
##                  95% CI : (0.4642, 0.5168)
##     No Information Rate : 0.3645
##     P-Value [Acc > NIR] : < 2e-16
##
##                   Kappa : 0.311
##
##  Mcnemar's Test P-Value : 0.01055
##
## Statistics by Class:
##
##                      Class: 0 Class: 1 Class: 2 Class: 3
## Sensitivity            0.5096   0.5116   0.5291  0.37786
## Specificity            0.8320   0.7940   0.8729  0.81622
## Pos Pred Value         0.4624   0.5876   0.5545  0.31731
## Neg Pred Value         0.8567   0.7392   0.8611  0.85302
## Precision              0.4624   0.5876   0.5545  0.31731
## Recall                 0.5096   0.5116   0.5291  0.37786
## F1                     0.4848   0.5470   0.5415  0.34495
## Prevalence             0.2210   0.3645   0.2301  0.18438
## Detection Rate         0.1126   0.1865   0.1217  0.06967
## Detection Prevalence   0.2435   0.3174   0.2196  0.21956
## Balanced Accuracy      0.6708   0.6528   0.7010  0.59704
```

Note: classes 0-3 are "anger", "fear", "joy" and "sadness" respectively.

The model achieves around 0.5 F1-score on first three classes and around 0.34 F1-score in “sadness” class. The results are acceptable as it is a 4-class classification model. Such model can be used to approximate emotion of unlabeled tweets.

Word Cloud Visualization

We create word cloud for preprocessed tweets in each category to get a big picture of vocabularies in emotional tweets. The size and color signify the frequency of the word in the specified emotion category. We have limited the maximum number of words for each cloud to 100.

Word Cloud for Emotion ‘Anger’

As shown in the word cloud below, we see that the most frequent words are rage, offend, anger, etc… Not all words in the cloud directly resemble the emotion ‘anger’, but majority of the words reasonably represent the emotion tweets for ‘anger’.

Figure 1: Word Cloud for Emotion ‘Anger’

Word Cloud for Emotion 'Fear'

As shown in the word cloud below, we see that the most frequent words are terror, fear, amp, etc... Not all words in the cloud directly resemble the emotion 'fear', but majority of the words reasonably represent the emotion tweets for 'fear'.

Figure 2: Word Cloud for Emotion 'Fear'

Word Cloud for Emotion 'Joy'

As shown in the word cloud below, we see that the most frequent words are happy, cheer, smile, etc... Not all words in the cloud directly resemble the emotion 'joy', but majority of the words reasonably represent the emotion tweets for 'joy'.

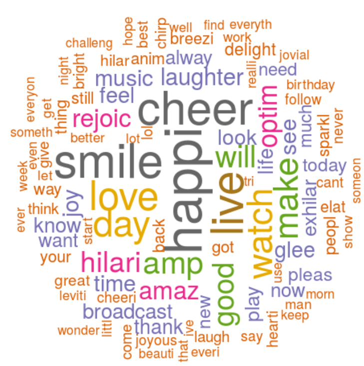


Figure 3: Word Cloud for Emotion 'Joy'

Word Cloud for Emotion ‘Sadness’

As shown in the word cloud below, we see that the most frequent words are sad, depress, lost, etc… Not all words in the cloud directly resemble the emotion ‘sadness’, but majority of the words reasonably represent the emotion tweets for ‘sadness’.

Figure 4: Word Cloud for Emotion ‘Sadness’