Analyze tweets to detect emotion of people on desired topic

Project Mentor-
Dr. Girdhari Singh
Associate Professor
MNIT Jaipur

Submitted By -
Parth Patel
2013UCP1652
Nishu Jain
2013UCP1089
Chandraprakash Avasthi
2013UCP1709

Outline

- Problem Statement
- Motivation
- Objectives
- Flowchart of algorithm designed
- Emotion Detection Steps
- Examples of Output
- Results obtained
- Applications
- Conclusions and Future Scope
- References

Problem Statement

To design an algorithm for detecting presence of different emotions (sad, surprise, fear, disgust, sad and happy) and deploying it to analyze emotions of people towards different trending (or desired) topics by extracting tweets from twitter and putting them together in a horizontal bar graph.

Motivation

- Blogging sites like Twitter attracts people to share their opinion about various topics.
- For future predictions and statistics involved , we must have a summary of opinion of public.
- Public sentiments shows the satisfaction, acceptance, reluctance, hatred, etc towards the object/topic concerned.
- Emotion detection is the study of these opinions of people and using them into better use for future statistics/predictions.

Objectives

- To design an algorithm which detects presence of different emotion in a sentence.
- To develop an application which summarizes the emotions present in different tweets and display the cumulative result showing percentage of people happy, sad, angry, etc about the topic concerned.
- To print the output in a horizontal bar graph for better understanding of the results.
- To design an algorithm without any training data and using rule based approach

Flowcharts of Algorithm designed

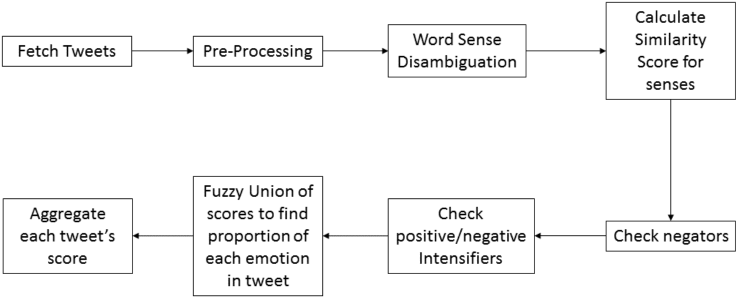

Figure: Block Diagram

Flowcharts of Algorithm designed

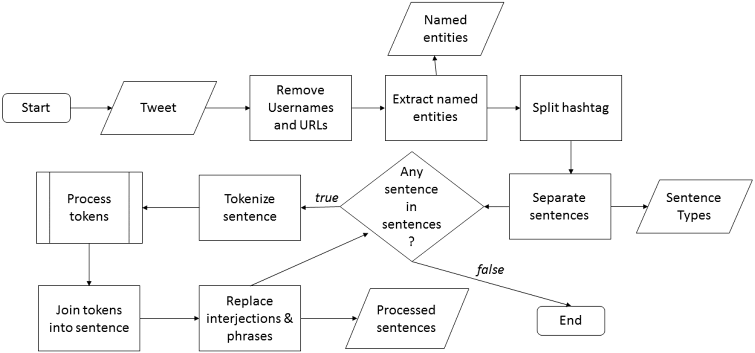

Figure: Flow chart for pre-processing tweets

Flowcharts of Algorithm designed

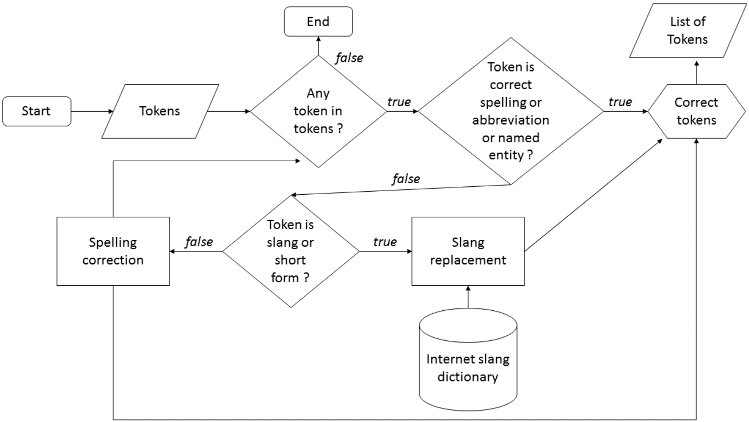

Figure: Flow chart for processing tokens

Flowcharts of Algorithm designed

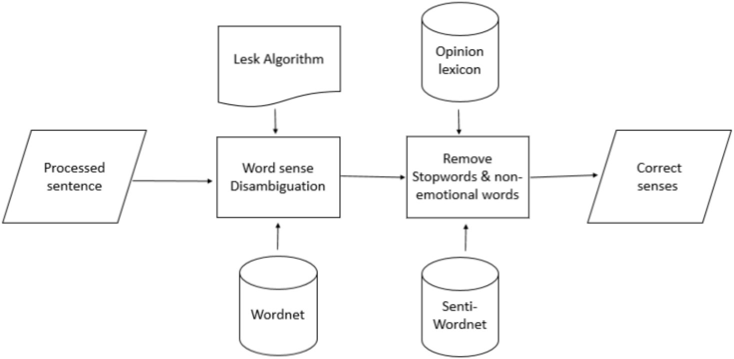

Figure: Flow chart for extracting correct senses

Flowcharts of Algorithm designed

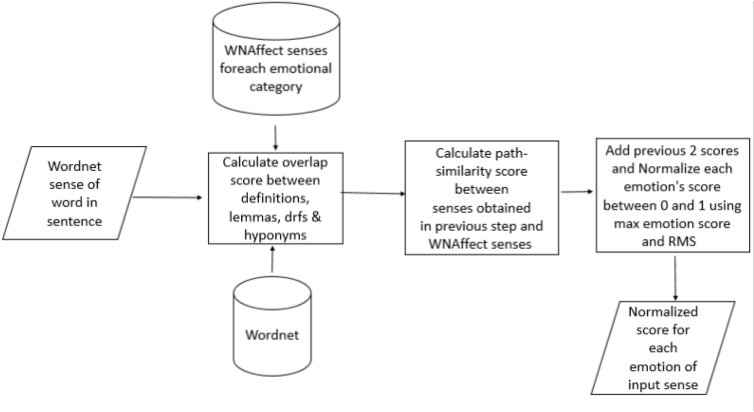

Figure: Flow chart for Calculating similarity scores

Modification rules for intensifiers and negators

$$\mu(pos\_intensifier(word)) = \sqrt{Score(word)} \forall Score(word) \geq 0.5 \quad (1)$$

$$\mu(pos\_intensifier(word)) = Score(word)^2 \forall Score(word) < 0.5 \quad (2)$$

For negative intensifiers scores are modified according to equations (3) and (4).

$$\mu(neg\_intensifier(word)) = \sqrt{Score(word)} \forall Score(word) < 0.5 \quad (3)$$

$$\mu(neg\_intensifier(word)) = Score(word)^2 \forall Score(word) \geq 0.5 \quad (4)$$

In case of combinators where negators and intensifier both are present, scores are modified according to equation (5).

$$\mu(combinator(word)) = \sqrt{Score(word) * \mu(intensifier(word))} \quad (5)$$

Difference between emotions with intensifiers and negators

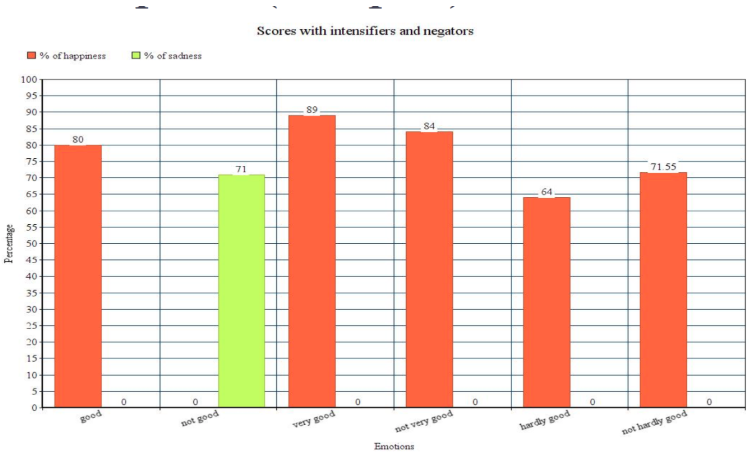

Figure: Scores in presence of intensifiers and negators

Fuzzy Union of similarity scores

- The fuzzy union of all the similarity scores for all keywords is taken to obtain the score for the entire sentence corresponding to each of the six emotions.
- Same method is applied to score of each sentence to find similarity score of tweet corresponding to each emotion.

Aggregate tweets' score

- If the score of any emotion is greater than 0.5, we consider that emotion is present in tweet.
- If none of the emotions' score is greater than 0.5, then that tweet is considered neutral.
- We maintain a counter for each emotion initially set to 0. For each emotional score greater than 0.5 for tweet we increment the counter of that particular emotion which means that emotion is present in tweet.
- At the end after aggregating scores for all tweets, we divide the counter for each emotion with total number of emotions in all tweets which gives us the proportion of each emotion in collected tweets.

1) BoycottSnapchat

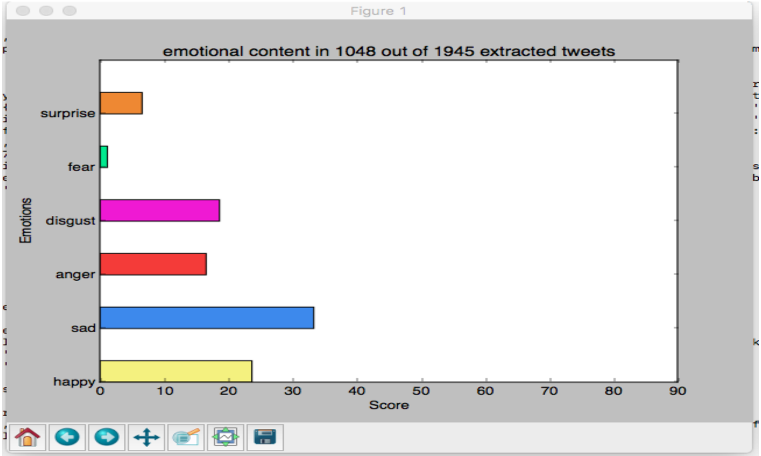

Figure: Graph showing emotions on tweets having BoycottSnapchat

Examples of Output

2) GujaratWelcomesPMModi

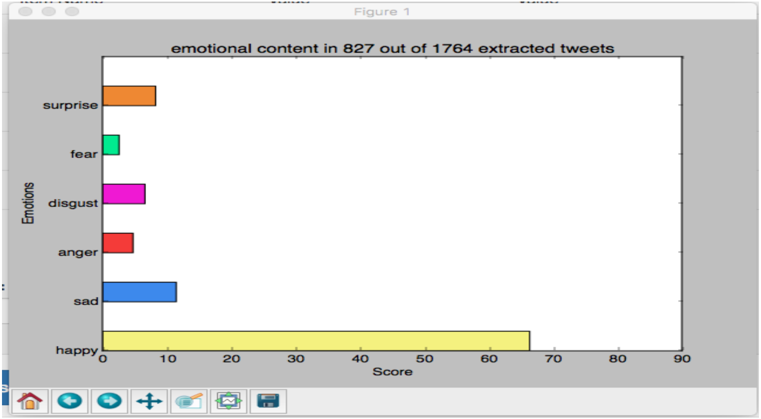

Figure: Graph showing emotions on tweets having Gujarat Welcomes PM Modi

3) Justice Leila Seth passed away

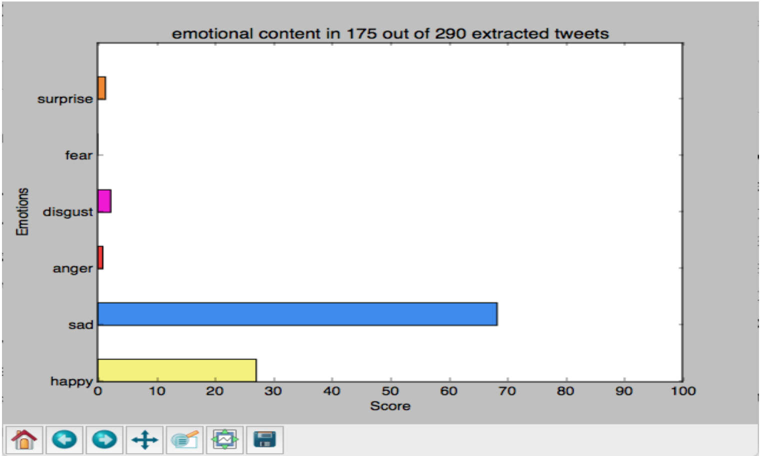

Figure: Graph showing emotions on tweets having Leila Seth

4) Baahubali2

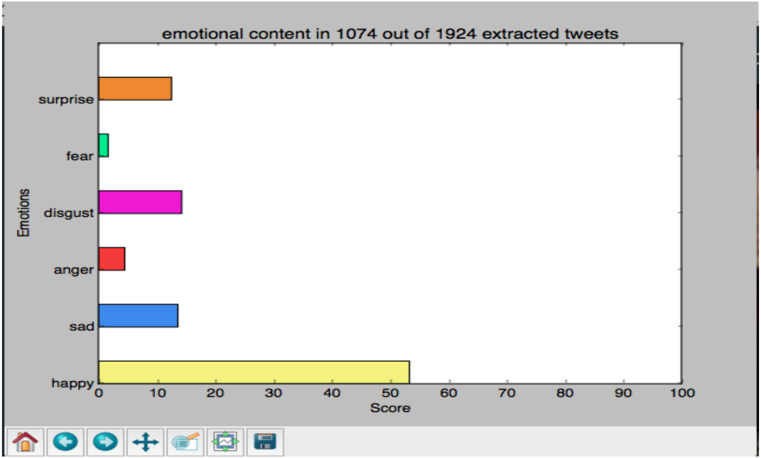

Figure: Graph showing emotions on tweets having Baahubali2

Results obtained

We tested the accuracy and performance of our algorithm by taking Aman's Dataset as standard input. The results are shown below :

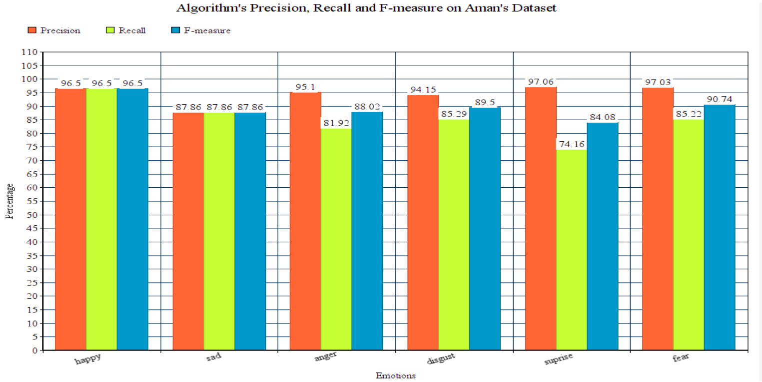

Figure: Precision, Recall and F-measure of the algorithm designed on Aman's Dataset

Results obtained

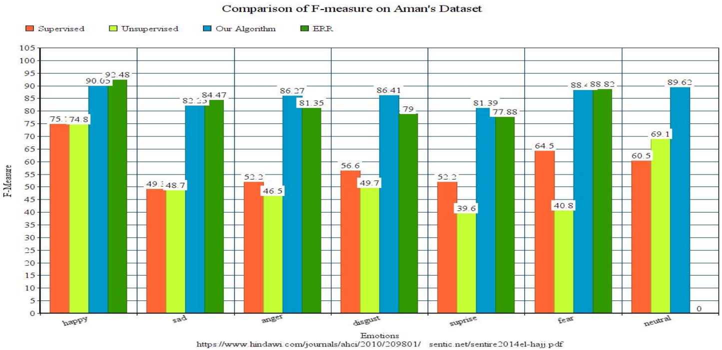

Figure: Comparison of F-measure in the results with other previously published works

Applications

- To prioritize emails of angry customers first in the customer care department of any big company or firm which uses online feedback system.
- Provide better counselling to patients and students based on the amount of fear and sad content .
- Sentiments of public towards different contestants of an election by the popularity of the person and number of tweets for him .
- To find Happiness Index : The UK government measures people's wellbeing; their statistics can be found on (https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing) . Other countries and cities such as Seattle, Dubai, and South Korea, have similar measures.

Applications

- Enhancing quality and brand image of your own restaurants , shops , etc by analysing reviews / feedbacks of customers and giving them discounts accordingly.
- Understanding the consumer. Improving perception of a customer with the ultimate goal to increase brand reputation and sales.
- To warn employees before sending an email by detecting the anger and disgust content so that he may not get fired by seniors because of rudeness in the email.

Conclusions and Future Scope

- Our proposed model neither depends on finite number of keywords nor on trained dataset for ML. We calculate similarity scores of each word with each of the emotion based on their wordnet definition, hyponyms, lemma names, derivationally related forms and hierarchy. But our model can not find embedded emotions in sentence without any emotional word. For example in a sentence *'My grandma passed away.'*, there is no emotional word but there is sadness expressed in sentence by phrase *'passed away'*. Our model will give this a neutral sentence.

Conclusions and Future Scope

- A PMI-based (Pointwise Mutual Information) model can be used here. It needs a manually emotion-annotated large dataset with such sentences. This model stores patterns showed in and corresponding emotion in database. For a live sentence, we would extract these patterns from it and check in our database in which of the class does this pattern belong to detect emotion expressed by pattern. We require manually annotated dataset for this model, which we did not have. Amazon mechanical Turk can manually annotate given dataset with six emotion but it is not currently available in India so we could not implement this model. But this model can be implemented with such dataset which would further increase our model's accuracy.

References

- M. Hasan, E. Rundensteiner, and E. Agu, "Emotex: Detecting emotions in twitter messages," 2014.
- R. V. M and M. Kumar, "Emotion and sarcasm identification of posts from facebook data using a hybrid approach," 2017.
- C.-Y. Yam. Emotion detection and recognition from text using deep learning. [Online]. Available: https://www.microsoft.com/reallifecode/2015/11/29/emotion-detection-and-recognition-from-text-using-deep-learning/
- S. Shaheen, W. El-Hajj, H. Hajj, and S. Elbassuoni, "Emotion recognition from text based on automatically generated rules," 2014