A PROJECT REPORT

ON

Analysing Tweets for Emotion Detection of people

For the partial fulfilment of the requirements of the degree of

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE & ENGINEERING

Submitted to

Malaviya National Institute of Technology, Jaipur

BY

Parth Patel (2013UCP1652)

Nishu Jain (2013UCP1089)

Chandraprakash Avasthi (2013UCP1709)

UNDER THE GUIDANCE OF

Dr. Girhdari Singh (Associate Professor)

Malaviya National Institute of Technology, Jaipur
Department of Computer Science & Engineering

CERTIFICATE

This is to certify that the project entitled

Analysing Tweets for Emotion Detection of people

submitted by

Parth Patel (2013UCP1652)
Nishu Jain (2013UCP1089)
Chandraprakash Avasthi (2013UCP1709)

is a record of bonafide work carried out by them , in the partial fulfilment of the requirement for the award of Degree of Bachelor of Technology in Computer Science Engineering at Malaviya National Institute of Technology, Jaipur.This work is done during academic year 2016-2017, under my guidance and supervision.

Date: / /

Dr. Girdhari Singh
(Project Guide)
Malaviya National Institute of Technology, Jaipur

ACKNOWLEDGEMENT

We are profoundly grateful to **Dr. Girdhari Singh (Associate Professor)** for his expert guidance and continuous encouragement throughout to see that this project reach its target since its commencement to its completion.

We extend our thanks to **Dr. Namita Mittal** and **Dr. Yogesh Kumar Meena** for their guidance and support in completing this project.

At last we must express our sincere heartfelt gratitude to all the staff members of Computer Science Engineering Department who helped us directly or indirectly during this course of work.

Parth Patel (2013UCP1652)
Nishu Jain (2013UCP1089)
Chandraprakash Avasthi (2013UCP1709)

DECLARATION

We, Parth Patel (2013UCP1652), Nishu Jain (2013UCP1089) and Chandraprakash Avasthi (2013UCP1709) declare that this project titled, **Analysing Tweets for Emotion Detection of people** and the work presented in it is an authentic record of our own work. This work is done wholly and mainly towards fulfilment of the dissertation report at MNIT, Jaipur where we have consulted the published work of others,this is always clearly attributed where we have quoted from the work of others, the source is always given. We have acknowledged all sources of help and no part of the report is plagiarized and report does not suffer from any act of plagiarism.

Parth Patel (2013UCP1652)

Nishu Jain (2013UCP1089)

Chandraprakash Avasthi (2013UCP1709)

Date: / /

ABSTRACT

People today are using social networking sites and microblogging platforms to express their opinions about something , someone, some event or some place. These platforms contain text abundant with people's reaction towards their surrounding. Detecting these emotions of people is an emerging topic in the field of Computer Science and have various applications. The personality of a person can be judged by his posts he usually uploads, the success of an event or product or law can be judged by the happiness in the posts of the people and this knowledge of detecting emotions of people of different caste, regions and race and religion can be useful to an analyser or a surveyor.

Here we have proposed a method of detecting emotions from tweets of twitter over a desired topic. We select Twitter messages as input data set, as they provide a very large, diverse and freely available ensemble of emotions. We also tried many alternative methods like SVM, PMI and Naive Bayes for classifying Twitter messages. Our method has an accuracy of over 85%, while demonstrating robustness across learning algorithms.

Contents

List of Figures

1 Introduction

1.1 Background

Emotion Detection is the process of identifying human emotions and attitude in the form of facial expressions, tone and pitch of speaker, messages typed by the user, etc. In this project, we have tried to detect emotions of humans by extracting their tweets from twitter. Twitter can be considered as a huge dataset that contains emotions and moods. For ex: "Going to grandma's house. Excited!!" showing happiness , "This movie was awful and such a waste of time and money" showing disgustness, "We will miss you grandma. RIP" showing sadness, etc.

Table below shows examples of different emotions present in tweets :

Tweet	Emotions
I miss you grandma.RIP	Sadness
This place is so comfortable and beautiful.	Happiness
You are bloody idiot.	Anger
What the fuck you think of yourself ? #Kejriwal	Disgust
I cant believe Arman Malik is in our college today.	Surprise
I'm gonna fail in these dreadful exams.	Fear

1.2 Motivation

Blogging sites like Twitter and social networking sites like Facebook attracts people to share their opinions 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.

Messages or posts of people on twitter are generally short limited upto only 140 characters which bound them to write the gist of their mind and conclude details all together. If all posts about a topic is analyzed thoroughly then a study can be maintained according to the results obtained in the form of emotions of people. This information may help any product company to enhance the quality of the product as per customer's need and demand. It may also help government to analyze the impact of any law by the outrage of people and helps take government proper preventive measures in case of large anger among citizens. This may also help government to survey about the happiness index of their nation. [7]

With the rapid emergence of Internet and e-business, e-tourism, e-commerce companies are very fond of positioning their brand image high and top in the market and are keen to exploit emotion analysis for carrying out their marketing schemes for evaluating public response and attitude towards their products and services. [8]

This application can be applied to any platform where text analysis for emotion detection is required. For instance, customer call centers for prioritizing mails of angry customers first for better service. All these vast applications make a pathway for the development of this project.

2 State of the Art

Owing to the importance of this topic, several techniques have been proposed for emotion detection from text, some of which a

2.1 Ho & Cao (2012) [1]

Their research was based on the approach that emotions are related to mental states of humans occurred due to some emotional events. They concluded that human mind move from initial state to final state on the occurrence of some emotional event. For this they used Hidden Markov Model where each sentence consisted of various sub ideas and each individual idea is treated as an event responsible for transition to a emotional state. By tracing the sequence of these transitional events of the sentence, the system detects the most appropriate emotion of the sentence.The best precision obtained by the system was 47% and this low accuracy was due to the ignorance of the semantic and syntactic analysis of the sentence which turned it into non context sensitive.

2.2 Yang et al (2012) [2]

They focussed on a hybrid approach for emotion classification by generating results of various techniques and integrating them using a voting based sytem. Techniques included lexicon keyword spotting, Conditional Random Field (CRF) emotion cue identification and machine learning based emotion classification using Support Vector Machine , Naive Bayes and Max Entropy classifiers. They achieved better results with F-score of 61% on a dataset of suicide notes however these datasets are not available.

2.3 Burget R. et al. (2011) [3]

Their research was based majorily on the pre processing of the input data which was Czech newspaper headlines and labelling it using a classifier. The pre processing was carried out at the word and sentence levels. The steps of pre processing includes applying POS tagging, then lemmatization and removal of stop words. Term Frequency - Inverse Document Frequency (Tf- Idf) was used to measure the relatedness between each term and each emotion class. This model achieved an average accuracy of 80% on 1000 Czech news headlines using Support Vector Machine with 10-fold cross validation. However this method was not deployed over English datasets and this was also context insensitive as it focusses only on presence of emotional words as their features of classification.

2.4 Ghazi et al. (2010) [4]

They proposed a method for hierarchical classification to classify the six Ekman emotions. Firstly, they checked whether a sentence holds an emotion or not, then they classified the emotion as either positive or negative and finally classified the emotion on a fine grained level. For each stage of classification the features for Support Vector Machine classifier were different and they achieved a better accuracy (+7%) over the flat classification where directly a sentence is classified into fine grained level. The major drawback of this model was its context insensitiveness.

2.5 Maryam Hasan (2014) [5]

To detect emotions in text messages such as tweets, they applied supervised learning methods to automatically classify short texts, according to a finer-grained category of the emotions. They labelled data from twitter using Hashtags for different emotion classes and then selected their features. Both these were given as inputs to Support Vector Machine classifier and K nearest neighbour classifier as training data. Test data is also given as input to the classifier and decided features are extracted and it is classified to appropriate emotional class by the classifier using training dataset. This model achieved better accuracy as they didn't conducted their analysis using 10 fold cross validation on the collected data itself.

2.6 Shadi Shaheen & Wassim El-Hajj (2014) [6]

They proposed a methodology where emotions were treated as concepts extracted from the sentences and can be expressed as nouns, adjectives, verbs and adverbs forms or as a combination of these various forms. Firstly, the dependency tree of a given sentence is constructed for syntactic and semantic analysis and then apply rules on the tree to truncate it such that the subtree that represents the sentence contains only the emotional part of the sentence. The reason for this truncation step is to make a lexicon of generally used emotion related phrases. This phrases are then called Emotion Recognition Rules (ERRs). ERRs of labelled training dataset are constructed and stored. For test data, ERRs are contructed using the same set of rules and the ERR obtained from the test data are then compared with the training dataset's ERRs for the classification of the test data into proper emotional class. They used K nearest neighbour (KNN) classifier for ERRs that have emotional words in their phrases but to cover cases containing contextual emotion they used PMI. Their method achieved better results however the method is not clearly described in their published work. The similarity between training data and test data is measured using ConceptNet which is deprecated now and hence the method is not feasible. The major drawback was this model was unable to classify neutral sentences and is only capable of classifying sentences which contains emotion.

2.7 Chew-Yean Yam (2015)[7]

They proposed the method of detection of emotions using deep learning technique. A multi-layered neural network with 3 hidden layers of 125, 25 and 5 neurons respectively, is used to tackle the task of learning to identify emotions from text using a bi-gram as the text feature representation. They had a training dataset of about 8 lakhs and obtained better results.

2.8 Raghavan V M (2017) [8]

They have proposed a rule based approach for emotion detection from text. Firstly , the test data is collected and pre processed to remove all unwanted slangs, symbols and scripts and obtain proper English sentences. Then rules are applied on these test data which checks the similarity of the words of the tokens with the synsets of each emotion classes and then return emotion of the most similar class. This algorithm has an advantage as it doesn't require any labelled training dataset which was really difficult to get. This method has major drawback as it checks the simlarity of only noun forms of the tokens present in the sentence or derivationally related noun forms but otherwise it ignore all emotion keywards which are not noun.

3 Proposed Methodology

We have proposed a score based model for detecting proportion of 6 of Ekman's basic emotions[9] in tweets collected from twitter on a desired topic. We give scores from 0 to 1 to each of the 6 emotions : happiness, sadness, anger, fear, disgust and surprise to each tweet and aggregate each tweet's score to final output. Figure 1 shows step by step procedure of the model.

Following sections describe each part in detail.

3.1 Fetching Tweets

In Twitter trending topics are tweeted with hashtag. A hashtag—written with a symbol—is used to index keywords or topics on Twitter. This function was created on Twitter, and allows people to easily follow

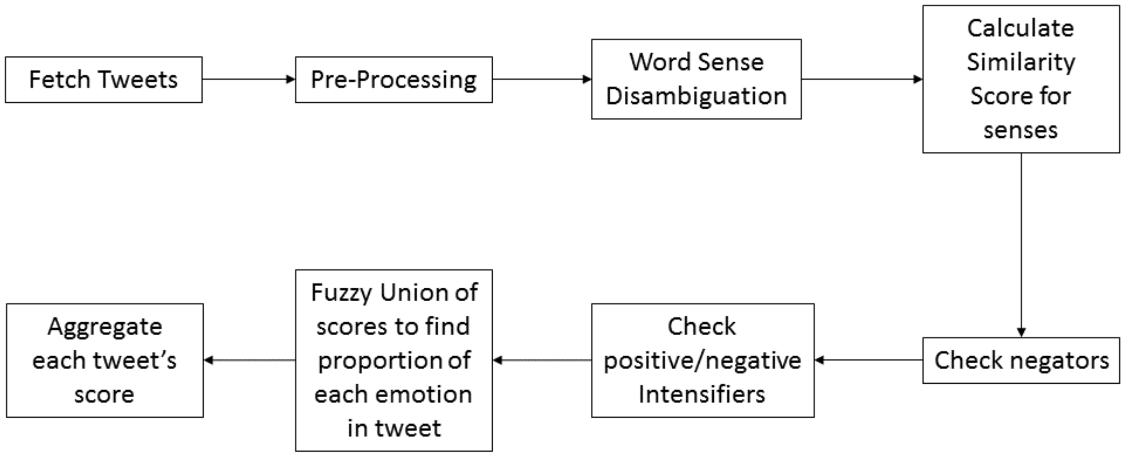

Figure 1: Block Diagram

topics they are interested in.[10] People use the hashtag symbol (#) before a relevant keyword or phrase in their Tweet to categorize those Tweets and help them show more easily in Twitter search. Now to gather tweets of a desired topic we search for the tweets with the hashtag of that topic (for example #bahubali to gather tweets for Bahubali movie). To fetch tweets from Twitter to our local computer, Twitter provides a search API[11]. Twitter4j[12] is a open source free of charge Java library that uses this API to fetch tweets in a Java program.

We have used Twitter4j to fetch tweets of desired topic/hashtag from twitter. Twitter4j has different filters for filtering tweets. We have only fetched English tweets and also excluded retweets.

3.2 Pre-Process Tweets

Flowchart for pre-processing tweets is shown in Figures 2 and 3.

Tweets fetched contain too much noise such as Usernames, URLs, some unidentified characters. These noise has to be removed to get clean tweets. Username is Twitter starts with @ symbol (for example Virat Kohli's username is @imVKohli). We removed usernames and URLs with regular expression. We extract named entities from tweet using Stanford Named Entity Recognizer[13] as named entities are names of place, person, animal, thing, date and organization which do not contain any emotion. Next step is to split hashtags. As a single hashtag can only have one word people generally write hashtag in camelcase *(ex. #WeWillWinTheMatch)* or words seperated by '_' *(ex. #we_will_win_the_match)*. In both of these examples we have to split hashtag to *'we will win the match'*. Then we separate sentences from tweet as each sentence may have different emotion. We also separate sentences from conjunctions such as *but, yet, except, still* and so on. For example in the tweet *"fielding was quite poor but batting was nice"*, the part before but express sad emotion and the part after but contain happiness. So to express individual part's emotion we separate them from such conjunctions. For each sentence we also store their type whether is interrogative sentence or exclamatory sentence or declarative sentence by checking their ending punctuation.

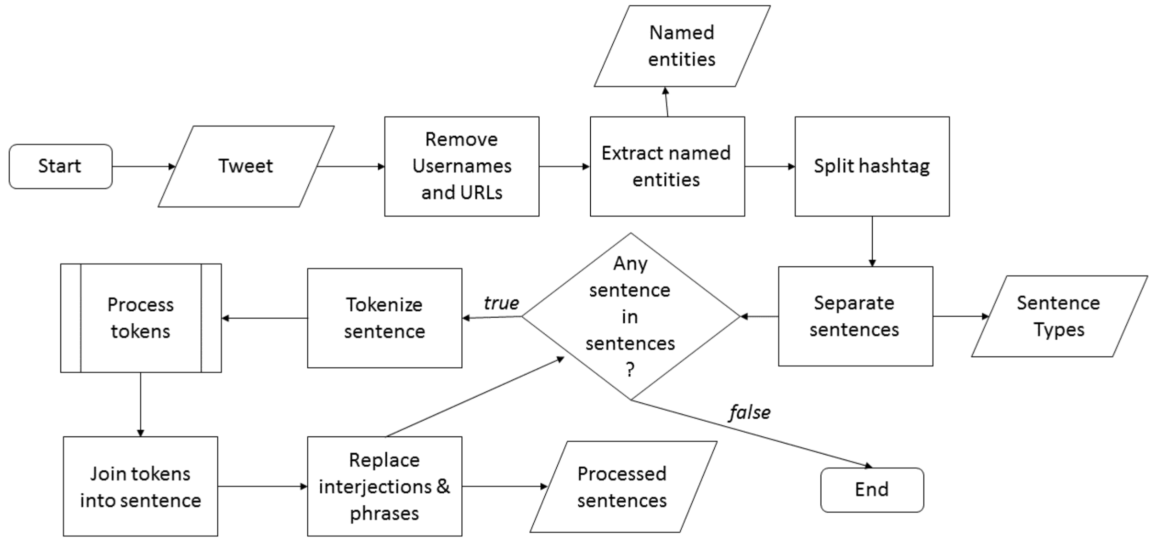

Figure 2: Flow chart for pre-processing tweets

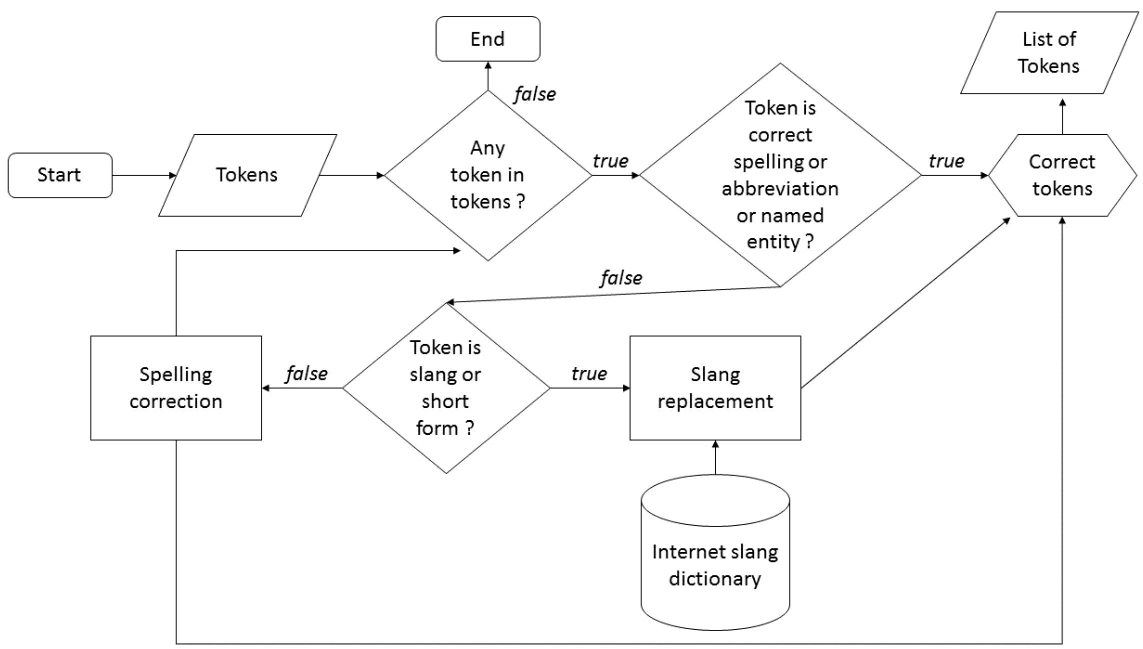

Figure 3: Flow chart for processing tokens

Now for each sentence obtained, we first tokenize the sentence using NLTK's[14] tweetTokenizer[15]. NLTK is a leading platform for building Python programs to work with human language data. It provides easy-to-use interfaces to over 50 corpora and lexical resources such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active discussion forum. Tweet tokenizer outputs tokens from a sentence. Now for each token in a sentence, if the token is abbreviation, pronoun, named entity or correct dictionary spelling we append it to list of correct tokens. Tokens with short form such as *he's, isn't, can't, we'd* are replaced with their full forms. People also use slangs in twitter such as *LOL for laughing out loud, RIP for rest in peace, gr8 for great, tq for thank you*. We have prepared **Internet slang dictionary** to replace these slangs with their correct forms. The correct forms are again appended to list of correct tokens. Also in twitter to put emphasis on certain words people write their spelling with multiple letters such as *'partyyyyyy', 'crrrrrrazyyyyy'*. We first remove the occurrence of more than 2 letters. Now if the spelling is incorrect we correct the spelling using python's enchant dictionary[16] and append to list of corrected tokens. We count the number of corrected spellings. If the number of corrected spellings is more than 50% of total tokens in sentence, we reject the sentence as in twitter many people write hindi words in English. (ex. *"aaj India match jeet jayega"*).

At this step we have tokens corresponding to each sentence. We join tokens by space to sentences. We replace interjections and some phrases such as *hooray, bravo, alas, oh my god, hats off, what the hell* from sentence with their meaning or emotion. Output of pre-processing is named entities and these sentences with their types.

3.3 Extracting correct senses of words

Wordnet® [17] is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. The resulting network of meaningfully related words and concepts can be navigated with the browser. WordNet is also freely and publicly available for download. WordNet's structure makes it a useful tool for computational linguistics and natural language processing.

Figure 4 shows flowchart of the process.

For each sentence obtained from pre-processing, we extract correct sense of each word of sentence from Wordnet using Lesk's word sense disambiguation algorithm [18]. Wordnet contain many senses of each word each having different meaning. For example word [**?**] contains 10 senses. Out of them sense 9 corresponds to financial institution bank and sense 1 corresponds to river bank. Different senses may be used in different sentences. Depending on context we have to find correct sense of word in a sentence. Lesk algorithm uses sense's

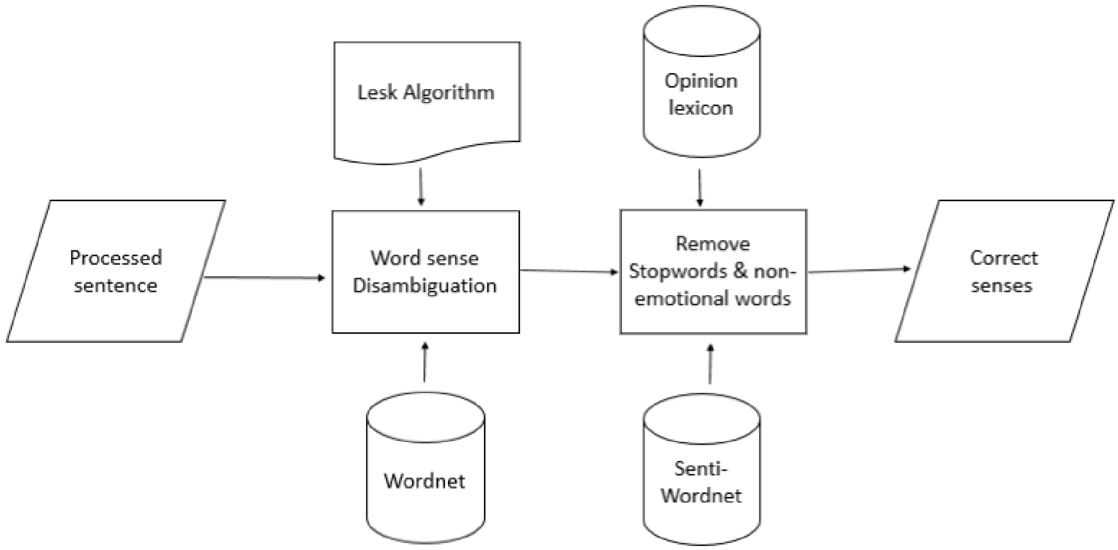

Figure 4: Flow chart for extracting correct senses

definition and lemma names and overlap it with sentence to find which sense overlaps maximum with the sentence. That is the correct sense. We remove stopwords such as pronouns, is, am, are and also non-emotional words. We check whether a word is emotional or not using Senti-Wordnet or opinion lexicon. SentiWordNet [19] is a lexical resource for opinion mining. SentiWordNet assigns to each synset of WordNet three sentiment scores: positivity, negativity, objectivity. Depending on these scores we can tell whether a word is emotional or not. Opinion Lexicon [20] contains list of positive and negative words. These list can be used to check whether a word is contains opinion or not. If no emotional word is contained in a sentence, then it is considered neutral.

3.4 Calculate similarity scores

For each sense obtained from previous step, we calculate their similarity score with happiness and surprise if sense is positive else with sadness, anger, disgust, fear and surprise. We have collected a list of senses under each of the six emotions from Wordnet-Affect [21]. We calculate similarity score of each emotional sense in sentence with collected senses under each emotion.

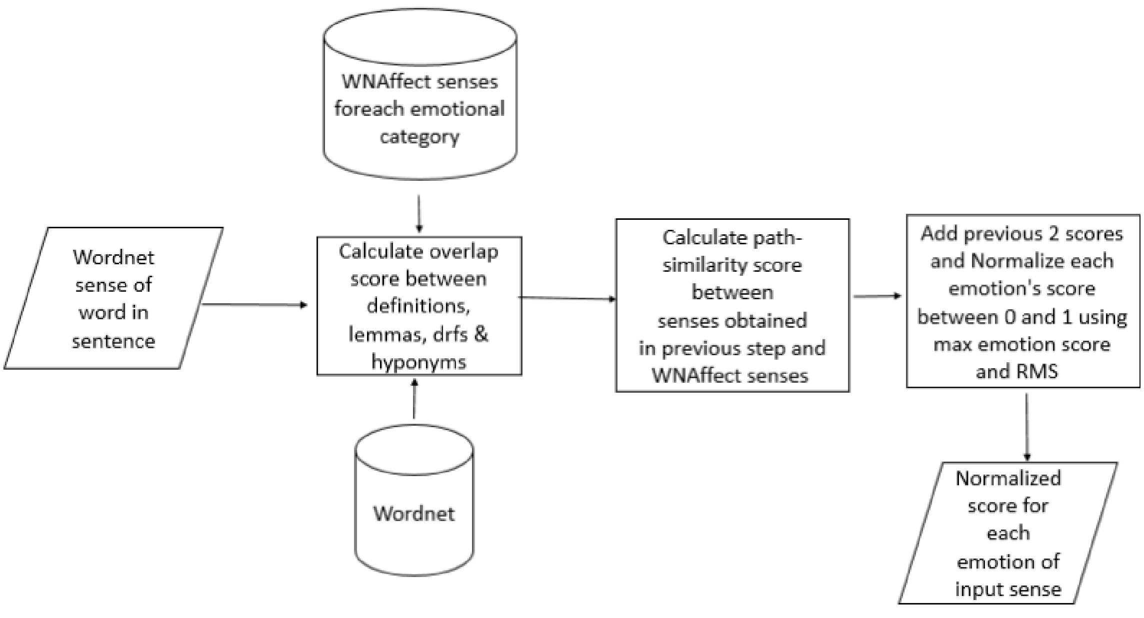

Figure 5: Flow chart for Calculating similarity scores

To calculate similarity score between two wordnet senses, we count the number of overlaps in their definitions, lemmas, derivationally related forms and hyponyms. For each senses in definition, lemmas, derivationally related forms and hyponyms, We calculate path-similarity using Wordnet to make the similarity module more robust.

Path-smilarity returns a score denoting how similar two word senses are, based on the shortest path that connects the senses in the is-a (hypernym/hypnoym) taxonomy. We add this two scores to get a final score for each emotion of corresponding wordnet sense.

The scores obtained till now will be within 0 to any finite positive number. To normalize them between 0 and 1, we divide each score with (1 + maximum of 6 emotions) and then divide each emotion with 1.25*root mean square(RMS) value of all emotions. After this scores will fall between 0 and 1 but they will always be less than 1 as we divide them with 1.25 times of RMS. This is done so that we can increase score if positive intensifier is present. This will be seen in later subsection.

3.5 Modification for negators

The words such as ‘not’, ‘no’, ‘none’, ‘never’ etc. are negators which negate the emotion. If a negator is present before a word in a sentence, we find antonym of word from wordnet and calculate similarity score of antonym word. If no antonyms are found we complement all emotional scores of word which are greater than 0.5. If two negators are present in same sentence, they cancel each other’s effect.

3.6 Modification for intensifiers

There are two types of intensifiers: positive and negative. Positive intensifiers such as ‘especially’, ‘exceptionally’, ‘excessively’, ‘extremely’, ‘extraordinarily’, etc. increase emotion level of word. Negative intensifiers such as ‘barely’, ‘moderately’, ‘slightly’, ‘less’, etc. are negative intensifiers which decrease the emotion level of the word.

We change the score of each emotion for sense if positive intensifier is present in the sentence according to equations (1) and (2).

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

$$\mu(pos\_intensifier(word)) = Score(word)^2 \forall Score(word) < 0.5 \tag{2}$$

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

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

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

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

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

Above equations are incorporated from work done in [22]. Diagram 6 shows effect of negators and intensifiers on scores of sentence *movie was not? very/hardly good*

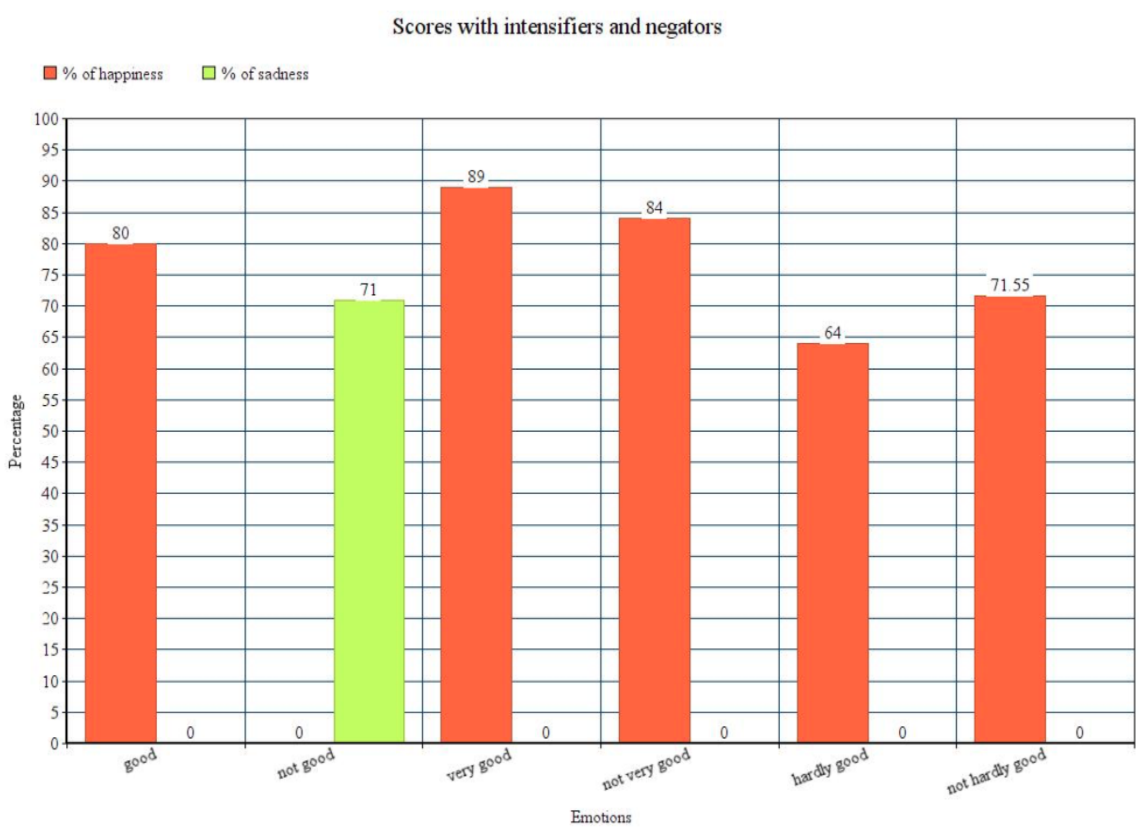

Figure 6: Scores in presence of intensifiers and negators

3.7 Fuzzy Union of similarity scores

The use of Fuzzy logic in the modeling of the emotions is a remarkably simple way to process vague, ambiguous or imprecise information. 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.

3.8 Aggregate tweets' score

At this step we have each tweet's score for six emotions within 0 to 1. 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.

We demonstrate the final result with graph as shown in next section.

4 Evaluation

We have used Aman's dataset for evaluating our model's accuracy. Saima Aman has prepared a emotion-annotated dataset [23] [24] with Ekman's six basic emotions and neutral emotion. These dataset contains 4089 gold sentences. Precision, Recall and F-Measure obtained on these gold sentences are shown in Figure 7.

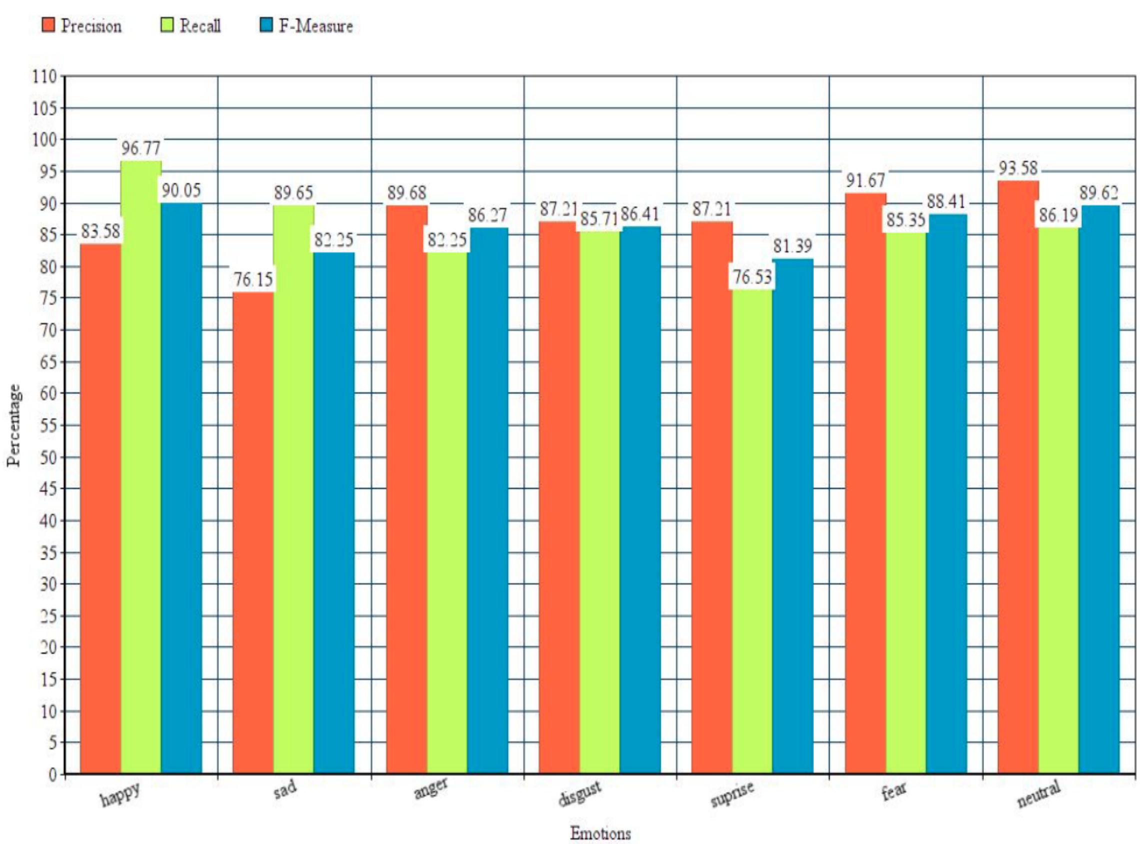

Figure 7: Precision, Recall and F-Measure on Aman's dataset

Figure 8 shows comparison of our model with other models which have used Aman's dataset for calculating accuracy.

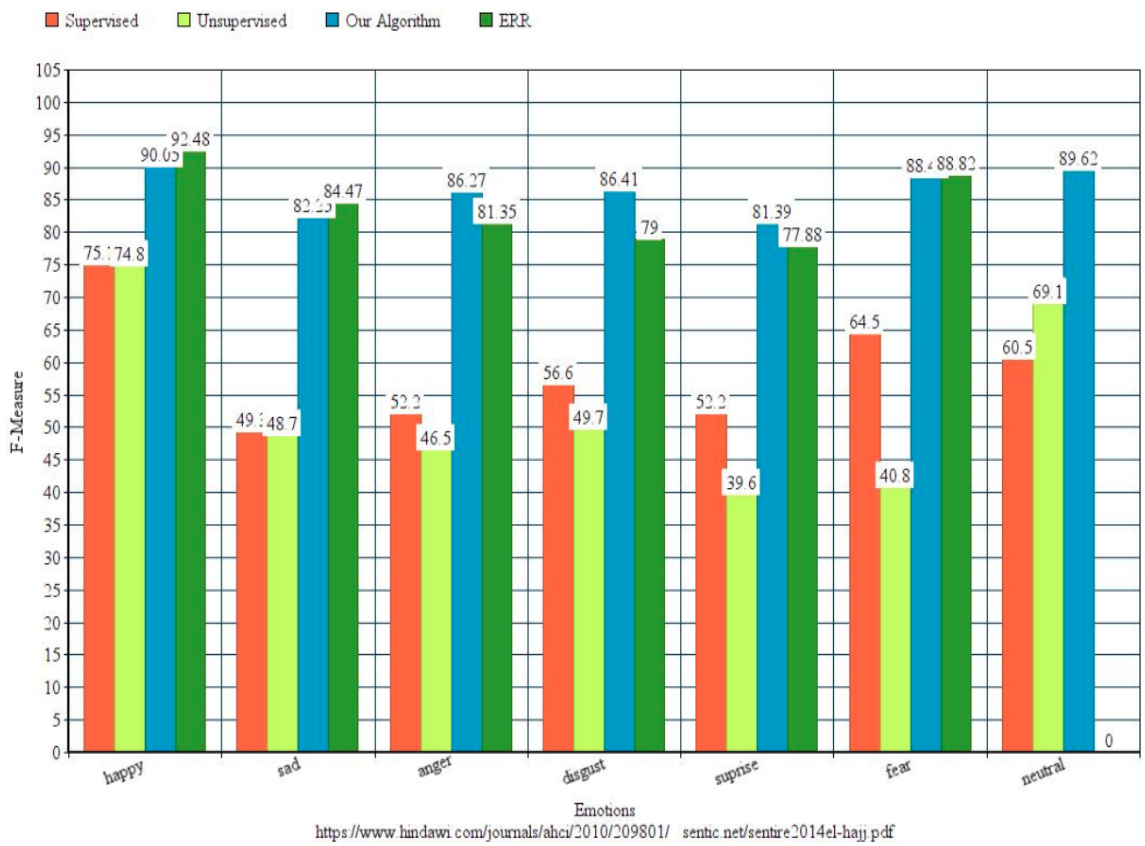

Figure 8: Comparison of F-Measures of different models on Aman's dataset

Average accuracy of our model is around 85%.

5 Results

By applying the proposed algorithm on different sets of input tweets for different Hashtags, we obtained the results below :

5.1 Outputs on different Topics of twitter

1) BoycottSnapchat

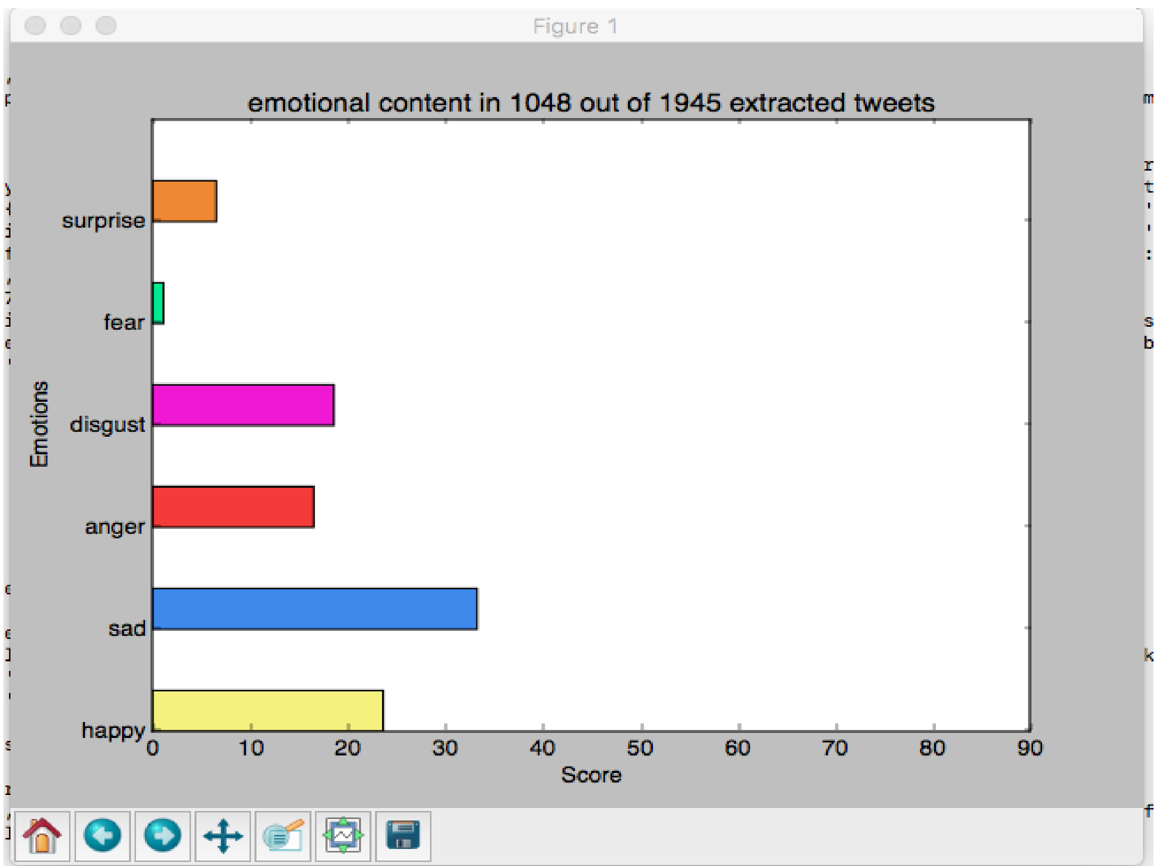

Figure 9: Graph showing emotions on tweets having BoycottSnapchat

2) GujaratWelcomesPMModi

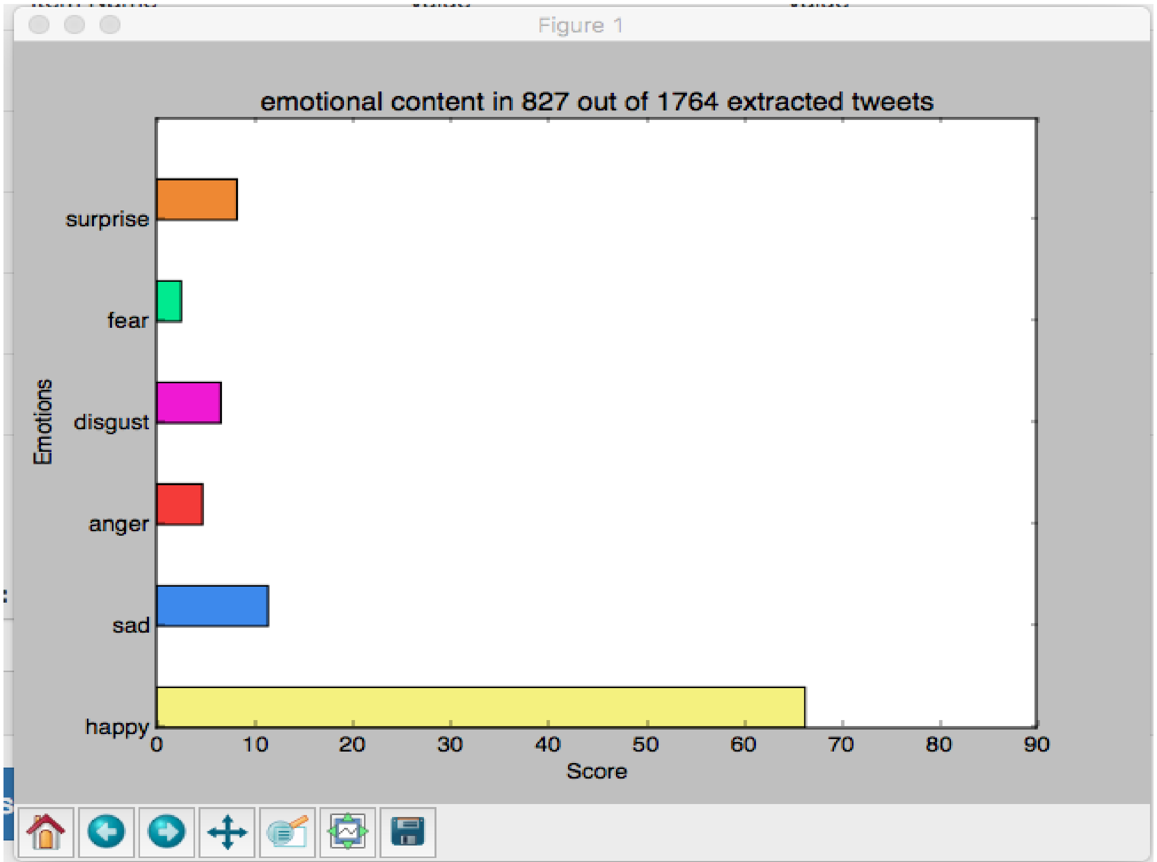

Figure 10: Graph showing emotions on tweets having GujaratWelcomesPM-Modi

3) Justice Leila Seth

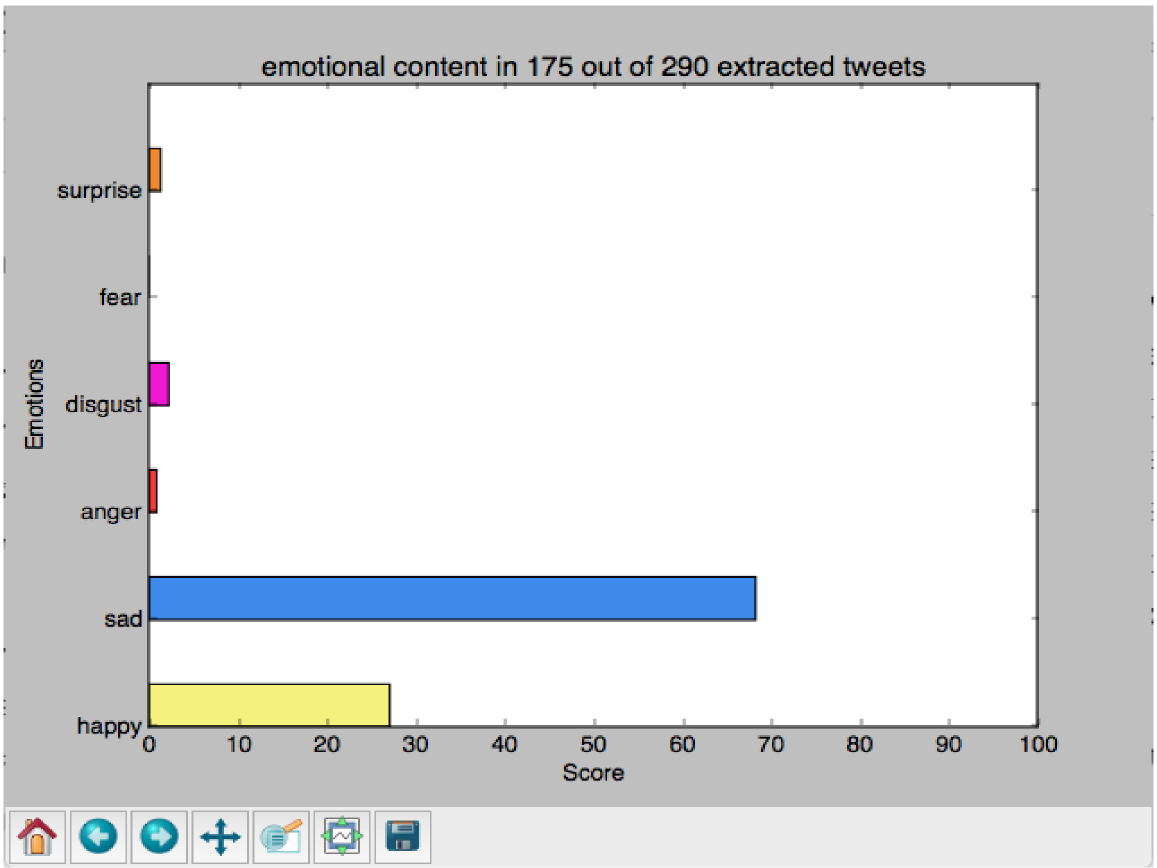

Figure 11: Graph showing emotions on tweets having "Justice Leila Seth passed away"

4) Baahubali2

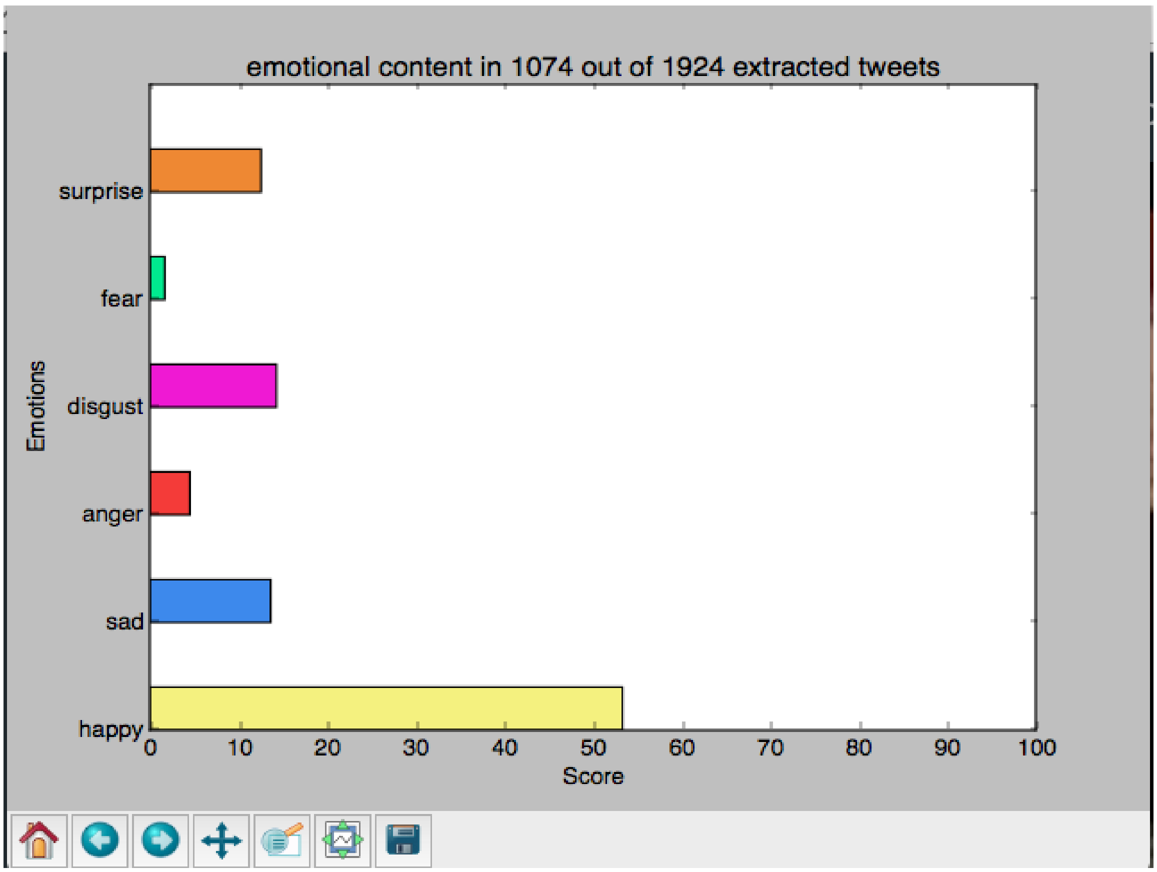

Figure 12: Graph showing emotions on tweets having Baahubali2

6 Applications

All e-platforms have customer care centers to resolve queries and complains of customers and maintain their brand value among the market by helping their customers and providing every guidance about their product. If a product fails and customer is complaining about it then currently human annotators prioritize the list in which issues of customers must be resolved by going through mails of every customer and analyzing the disappointment level in the mood of the customer through anger content in his mail. If the company uses the application developed by us on their mail box computer system and feed their customer's mail as input to the application then our algorithm will process those mails and can sort them in decreasing order of anger and disgust content. This makes prioritizing process efficient and saves human labour.

Counsellors can use this application for analyzing the problems of their patients and detecting the level of sadness and fear in those patients. Patients with high level of sadness or anger must be handled with extra care and firstly made comfortable on the other hand, patient with moderate level of sadness can be handled lightly and may be given some challenging task to enhance his skills as later person may have will power to accept the challenges as compared to the former one. The proposed method can be used by healthcare professionals or counseling agencies to monitor and track a patient's emotional states, or to recognize anxiety or systemic stressors of populations (e.g. different student groups on campus).This enhances the work in the field of counselling.

Personality of a person can be evaluated by analyzing all the posts, messages or blogs he has ever mentioned on any social networking sites or blogging sites. If a person is short tempered then anger content in his texts must be more in comparison to person with ample of patience and far sightedness.

It can also be used on a personal computer before sending any text in professional work life. For instance, an employee sending a harsh email to his colleague or superior. A tool that can analyze the email for emotions and alert the employee about its harshness before sending it comes in very handy to protect the employee's state.

Government of many countries conducts survey to gauge how happy or sad its citizens are. Happy Planet Index , Societal Well-being metrics, etc are different measures which different countries like UK, South Korea, Settle, Dubai, etc uses.

However, the system developed based on our proposed approach would be able to automatically detect what people feel about various topics from twitter tweets. For example, the system can recognize:

- percentage of people expressing higher levels of life satisfaction
- percentage of people who feel happy and cheerful,
- percentage of people who feel calm and peaceful, and
- percentage of people expressing higher levels of anxiety or depression.

7 Conclusion & 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.

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 [25] 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 [26] 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

[1] Ho, D. T., Cao, and T. H., "A high-order hidden markov model for emotion detection from textual data," 2012.

[2] Yang, H. Willis, A. D. Roeck, and A. Nuseibeh, "A hybrid model for automatic emotion recognition in suicide notes," 2012.

[3] R. BURGET, J. KARÁSEK, and Z. SMÉKAL, "Recognition of emotions in czech newspaper headlines," 2011.

[4] D. Ghazi, D. Inkpen, and S. Szpakowicz, "Hierarchical versus flat classification of emotions in text," 2010.

[5] M. Hasan, E. Rundensteiner, and E. Agu, "Emotex: Detecting emotions in twitter messages," 2014.

[6] S. Shaheen, W. El-Hajj, H. Hajj, and S. Elbassuoni, "Emotion recognition from text based on automatically generated rules," 2014.

[7] 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/

[8] R. V. M and M. Kumar, "Emotion and sarcasm identification of posts from facebook data using a hybrid approach," 2017.

[9] P. Ekman. Ekman's six basic emotions: List definitions. [Online]. Available: http://study.com/academy/lesson/ekmans-six-basic-emotions-list-definitions-quiz.html

[10] Using hashtags on twitter. [Online]. Available: https://support.twitter.com/articles/49309

[11] Twitter search api. [Online]. Available: https://dev.twitter.com/rest/public/search

[12] Twitter4j. [Online]. Available: http://twitter4j.org/en/index.html

[13] T. S. N. L. P. Group. Stanford named entity recognizer. [Online]. Available: https://nlp.stanford.edu/software/CRF-NER.shtml

[14] Natural language toolkit. [Online]. Available: http://www.nltk.org/

[15] Nltk's tweet tokenizer. [Online]. Available: http://www.nltk.org/api/nltk.tokenize.html

[16] Pyenchant, a spellchecking library for python. [Online]. Available: http://pythonhosted.org/pyenchant/tutorial.html

[17] Wordnet, a large lexical database of english. [Online]. Available: https://wordnet.princeton.edu/

[18] M. E. Lesk, "Lesk algorithm."

[19] A. Esuli and F. Sebastiani, "Sentiwordnet: A publicly available lexical resource for opinion mining."

[20] M. Hu and B. Liu, "Mining and summarizing customer reviews-proceedings of the acm sigkdd international conference on knowledge discovery and data mining," 2004.

[21] Wordnet-affect. [Online]. Available: http://wndomains.fbk.eu/wnaffect.html

[22] V. Loia and S. Senatore, "A fuzzy-oriented sentic analysis to capture the human emotion in web-based content," 2013.

[23] S. Aman and S. Szpakowicz, "Using roget's thesaurus for fine-grained emotion recognition," 2008.

[24] ——, "Identifying expressions of emotion in text," 2007.

[25] P. D. Turney, "Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews."

[26] Amazon mechanical turk. [Online]. Available: https://www.mturk.com/mturk/welcome