A Framework to Find Popularity of a Political Leader Using Emotion Mining

Chintala Kamala Kumari
M.Tech, Department of Computer Science and Engineering
S R K R Engineering College Bhimavaram, India
chkkamala@gmail.com

Dr. Kalidindi Ramaprasada Raju
Professor, Department of Computer Science and Engineering
S R K R Engineering College Bhimavaram, India
ramaprasad.kalidindi@gmail.com

***Abstract*—** Emotion mining is becoming an important part of the sentiment analysis. Online text sources are evolving into large-scale data repositories. These repositories contain valuable knowledge about human emotions like anger, disgust, fear, joy, sadness, and surprise. These emotional classes are useful in predicting the social trends. This work extracts emotions from twitter data and categorizes into six emotions. From these categories, one can identify the trending. On this basis, one can extract the popularity of a political leader in the society using tweets from his/her account. Score is calculated for each tweet of every leader and the popularity score is measured.

Keywords-*Emotion mining; Sentiment analysis; Online Text Sources; Emotions.*

*****

I. INTRODUCTION

Emotion mining is considered as a sub field of sentiment analysis that aiming for interpreting the emotions from the natural language. Emotion mining is the process that provides insight to users subconscious motivations reliably and routinely. It detects, analyzes and evaluates human's feelings towards different events, issues, services or other interest.

Emotion analysis is to detect and recognize the types of feelings through expression of texts, such as *anger, disgust, fear, joy, sadness* and *surprise* [1]. These are mainly associated with positive and negative sentiments, positive sentiments are joy and surprise as well as negative sentiments is anger, disgust, fear, sadness. Emotion mining can be classified into three types: In keyword based approach, Emotions are detected based on the related set(s) of keywords found in the input text. In machine learning based approach [5], emotions are detected based on previous training result with respect to specific statistic learning methods. In Hybrid approach, emotions are detected based on the combination of detected keyword, learned patterns, and other supplementary information.

Emotion analysis also contains two types of strategies for emotion detection. These strategies also detect emotions from social ones, they are: facial recognition and semantic analysis. Facial recognition detects facial expressions in video and photos, which determine common emotions such as surprise, joy, anger, sadness, disgust and fear. These emotions can detect expressions on a face and reading their relationships to one another, with the help of facial databases.

The second type in emotion analysis is semantic analysis. Semantic analysis contains algorithms that detect emotion in sentence, whether it be speech or in writing. This is achieved by ranking in a text are positive or negative in emotion, through which an overall expression emerges. Algorithms may detect multiple examples of expression in a single statement, offering a comprehensive look into what the speaker or writer is thinking.

An emotion is a particular feeling that characterizes a state of mind. The emotion detection from text [4] has attracted growing attention due to its potentially useful applications. For example, psychologists [1] can better assist their patients by analyzing their session transcripts for any subtle emotions; reliable emotion detection can help develop powerful human-computer interaction devices; and deep emotional analysis of public data such as tweets and blogs could reveal interesting insights into human nature and behavior.

In this paper, mainly focus our attention on social networking site twitter. On twitter, people share their views and opinions in the form of "tweet". These tweets are the main component which determines emotion of the people as their opinions. The tweets are retrieved from twitter using twitter API and perform emotion detection method on tweet messages.

The rest of the paper is organized as follows. Section two discusses the related work on emotion detection from text. Section three discusses the detailed methodology. Section four present the result analysis. Section five concludes the paper.

II. RELATED WORK

Emotion analysis can be applied to all kinds of text, but certain domains and modes of communication tend to have more overt expressions of emotions than others. Emotions are mental states accompanied by physiological changes. Ekman [1] identified six basic emotions are happiness, sadness, anger, fear, disgust and surprise. Using these emotion classes Dhavan et al. [2], Proposed a novel framework for characterizing emotional interactions in social networks. Their aim is to extract the emotional content of texts in online social networks. The interest is in to determine whether the text is an expression of the writer's emotions or not if yes then what type of emotions. For this purpose, text mining techniques are performed on comments/messages from twitter with annotated micro blog posts annotated at the tweet level with seven emotions are anger, disgust, fear, joy, love, sadness, surprise. Analyzing how emotions are distributed in the data and compare it to the distributions in other emotion annotated corpora [3]. In this also used the annotated corpus to train a classifier that automatically discovers the emotions in tweets.

Abdul Hannan [4] proposed a method to extract emotion from text at sentence level. This method detects emotion from a text by searching emotional key words from that input. To

ISSN: 2454-4248
202– 206

make the detection more accurate, emotion-affect-bearing words and phrases were also analyzed.

Chatzakou et al. [5], proceeded with an approach which permits the detection of 12 emotions. The studies conducted on different text sources which span from news headlines to Online Social Network sources, i.e., Twitter and Facebook, to ensure that our methods are valid for online texts with different structural attributes. This work also considers the up to now lack of the explicit human emotion declaration and validation in such studies and it deals with the challenging tasks for detecting people's emotions on the 'wild'. The human emotions were to detect and clip emotional scenes from movies, using natural language processing techniques on their subtitles [6]. This involves the design of a naive counting classifier, to add semantic knowledge. With each increment, analyze the change in performance and made suitable modifications to the classifier.

Nilesh M. Shelke [7] identifies the problem of emotion recognition from text focusing on the implicit emotional statements, the descriptions of emotional events. It is to provide machines with the model for emotion reasoning allowing deeper understanding of causes of specific emotions. Mainly focuses on prior approaches to this problem explaining their assumptions, employed sources of affective information, ways of event representation, and proposed classification algorithm. The proposed approach is Emotex [8], it enables to classify large amounts of short texts with no manual efforts with accuracy. Classifying short texts according to finer-grained classes of emotions provides rich and informative data about the emotional states of individuals. These data can be used by healthcare professionals for early detection of the psychological disorders such as anxiety or depression. They are also compared the accuracy of several algorithms for classifying the moods of twitter messages. It is proposed [9] for characterizing emotional interactions in social networks, and then using these characteristics to distinguish friends from acquaintances. The framework includes a model for data collection, database schemas, data preprocessing and data mining steps. More related to this work is that of Vincent S.Erich [10], perform an emotion classification system for ThemeStreams that can detect and classify multiple emotions in Dutch political tweets. Using data from different installations of ThemeStreams, labeled dataset of 399 tweets covering 8 emotion categories has been realized. Performing the methods binary relevance method and random k-label set (RAKEL) method.

Stanford NLP [11] provides most common core natural language processing (NLP) steps, from tokenization through to co-reference resolution. An example is implemented to explain problems related to opinion mining and the challenges that effects NLP system, by using a set of pre-election tweets [12]. There are many sentiment analysis tools available which classify positive, negative and neutral of a desired topic or keyword. Using one dataset [13], for finding positive and negative emotions of certain parties or politicians in tweets and their re-tweets. Kartik Singhal et al. [14] proposed novel approach based on semantics and context rules to detect the public opinion and predicting election results. Getting political tweets during the general election in India, and evaluate proposed approach against the election results. The proposed approach contains various rules based on semantic structure of the sentence.

III. PROPOSED METHOD

The proposed method is based on keyword based approach, it is also described as the problem of finding occurrences of keywords from a given set as substrings in a given string. This problem has been studied in the past and algorithms have been suggested for solving it. In the context of emotion detection, this method is based on certain predefined keywords. These words are classified into categories such as anger, disgust, fear, joy, sadness, and surprise.

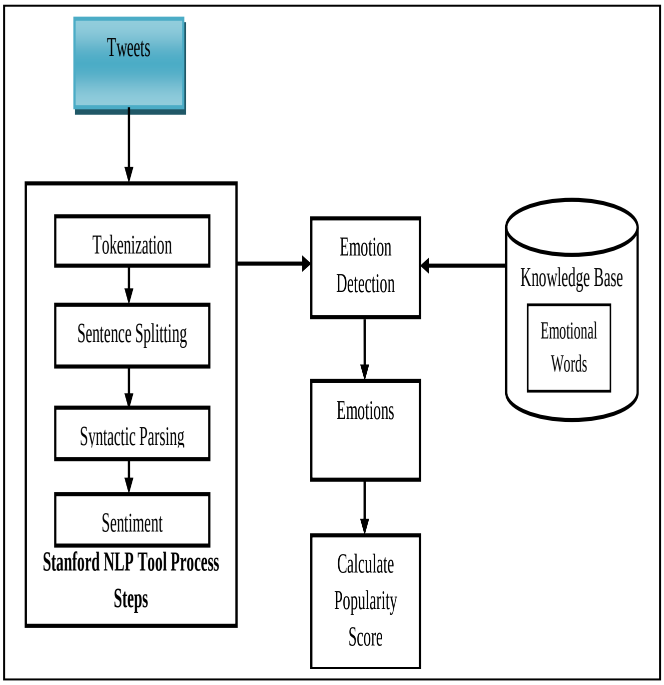

Figure 1. Proposed Architecture to Find Popularity Score

Twitter:

In this, the tweets are retrieved from the twitter based on key. The tweets retrieved from latest news on their accounts and the languages are Hindi and English. The tweets are appended to tweet list. The tweets are presented by using key and using Stanford NLP tool and the emotion detector used to find the popularity score for emotions.

Stanford NLP Tool Process Steps:

Stanford CoreNLP[11], a java annotation pipeline framework, which provides most of the common core natural language processing (NLP) steps, from tokenization through to sentiment.

Tokenization: The tokenization [11] phase of lexical analysis in Java handles breaking down the lines of Unicode source code into comments, white space, and tokens. Sometimes the tokens split words in suitable ways for NLP processing.

Sentence Splitting: In sentence splitting [11], splits a sequence of tokens into sentences.

Syntactic Parsing: Provides full syntactic analysis [11], using both the constituent and the dependency representations. The constituent-based output is saved in TreeAnnotation.

Sentiment: Attaches a binarized tree of the sentence to the sentence level CoreMap. The nodes of the tree then contain the annotations from RNNCoreAnnotations indicating the predicted class and scores for that subtree. Nodes of a binarized tree of each sentence, including, in particular, the root node of each sentence, are given a sentiment score.

Knowledge Base:

The tool knowledge base (KB) is used to store information regarding emotional words. This tool performs sentiment analysis at a sentence level. It uses the Stanford parser to analyze sentence's structure and create the dependency tree based on the words' relationships. Words known to convey emotions are spotted using the lexical resources of the knowledge base and each emotional word detected is further analyzed by the tool and the way it interacts with the sentences words are determined. Based on the words relationships, identifies specific types of emotional words interactions to specify its emotional strength. Finally, the emotion extractor unit specifies the sentences overall emotional status based on the sentence emotional parts.

Emotion Detection:

The input text data can be recognized with the help of proposed emotion detector algorithm. The algorithm calculates score for every emotion of primary level available in the emotion ontology by adding the scores of its respective secondary and tertiary levels' emotions.

IV. RESULT ANALYSIS

Emotion analysis is carried out in this paper on live twitter tweets of popular leaders in politics. The tweets are taken from nationwide leaders as well as statewide leaders and the total numbers of tweets are based on latest news from their accounts. In Table 1 percentage of nationwide leaders were shown:

Table 1. Nationwide Leaders Percentage of Tweets

Emotion	N 1	N 2	N3
Anger	48.5%	51%	58%
Disgust	17.5%	13.5%	12.5%
Fear	8%	6.5%	5.5%
Joy	13%	15%	11.5%
Sadness	8.5%	8%	5.5%
Surprise	4.5%	6%	7%

In this paper, six basic primary emotions [5] are identified on each leader. The emotions are divided into two categories, positive and negative. The positive contains only joy and surprise, whereas negative contains anger, disgust, fear and sadness. Here the emotion mining is performed on tweets of Nationwide leaders. From the Table 1, Narendra Modi as N1 contains 48% of anger, Suresh Prabhu as N2 contains 51% and Rahul Gandhi as N3 contains 58%, respectively.

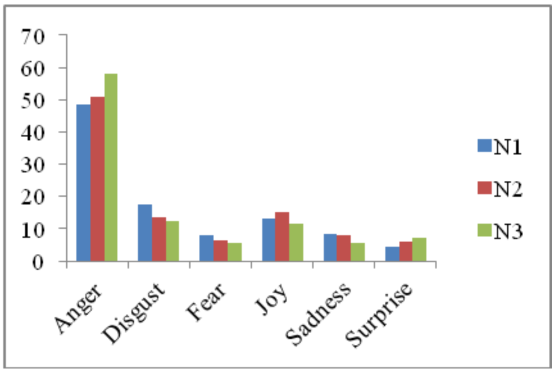

Figure 2. Percentage of Tweets for Nationwide Emotions

Based on the above values calculate the weights for all emotions to represent the P_s for popularity score of politicians using emotion mining from range -1 to 1, while comparing the geographic nation twitter has more data about politicians. Then weight is required to calculate that data. Weight of the i^{th} emotion that is taken from the tweets is denoted by W_i [0 to 1], the ratio of specific emotion count to the total number of tweets is T_i, n specifies the number of emotions and x is 1 for negative and 2 for positive emotions. Then the formula to popularity score is,

$$P_s = \Sigma^{n}_{i=1} (-1)^x (W_i T_i) \qquad (1)$$

Let the weights of emotions are anger =0.6, disgust=0.2, fear=0.1, joy=0.5, sadness=0.1, surprise=0.5 respectively. Then put the weights of all the leaders in the popularity score formula and getting the overall result for each leader. This result represents the popularity score for each politician on recent trends.

The scores for positive and negative emotions of Narendra Modi are 0.0875 and -0.2615, Suresh Prabhu is 0.105 and -0.3475 and Rahul Gandhi is 0.0925 and -0.384.

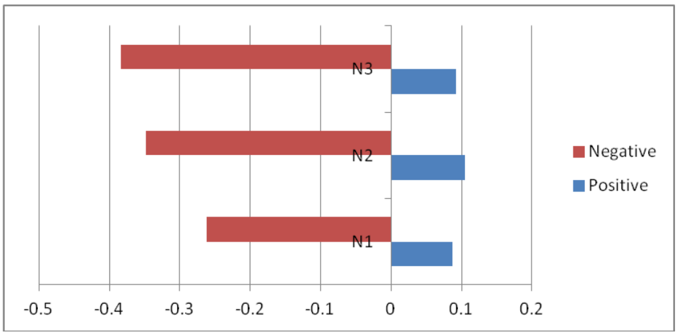

Figure 3. The Popularity Score for All Nationwide Leaders

From the above results, Narendra Modi gets the popularity score when compared to others leaders.

By all these nationwide leaders contains the emotions on negativity representation. Based on this representation we have to find more score on some of the leaders are belongs to statewide.

The below table contains statewide leaders that are represents the emotions on tweets.

Table 2. Statewide Leaders Percentage of Tweets

Emotion	S 1	S 2	S 3
Anger	50%	77.5%	72%
Disgust	10%	11%	13.5%
Fear	8%	1.5%	1%
Joy	5.5%	3%	2%
Sadness	17%	4.5%	2.5%
Surprise	9.5%	2.5%	9%

From Table 2, Nara ChandraBabu Naidu as S1 contains 50% of anger, JaganMohan Reddy as S2 contains 77.50% and PawanKalyan as S3 contains 72%, respectively. The anger emotion of Nara ChandraBabu Naidu is 50% when compared to other leaders. This represents the S1 has less anger than other leaders.

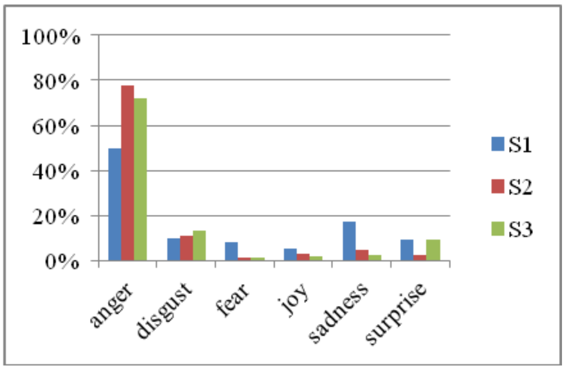

Figure 4. Percentage of Tweets for Statewide Emotions

The popularity score for positive and negative emotions from 1 to 6 in equation 1, Nara ChandraBabu Naidu has 0.075 and -0.345, JaganMohan Reddy has 0.0275 and -0.493 and PawanKalyan has 0.055 and -0.4625.

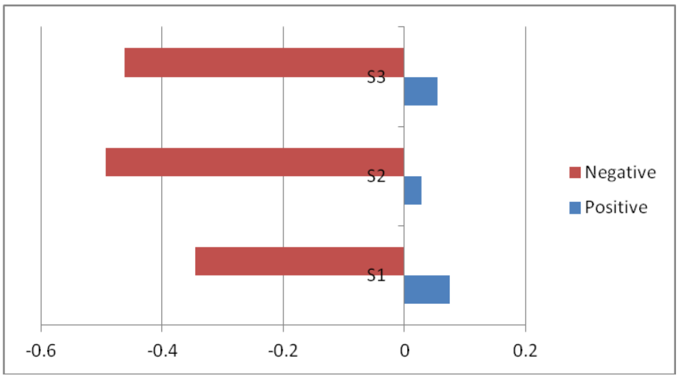

Figure 5. The Popularity Score for All Statewide Leaders

From the above results, Nara ChandraBabu Naidu gets the popularity score when compared to others leaders.

In this work, total numbers of emotion tweets are taken for the statewide and nationwide leaders. Since statewide leaders are popular in their geographical area, the popularity scores of these leaders are not compared with nationwide leaders.

V. CONCLUSION

Emotion analysis is to detect and recognize the types of feelings through expression of texts, such as anger, disgust, fear, joy, sadness and surprise. Emotion detection from text has attracted growing attention due to its potentially useful applications. In this paper, methods which are currently being used to detect emotion from text are reviewed along with their limitations and new system architecture is proposed, which would perform efficiently.

Emotion mining has a bright scope in future. The emotion mining can be implemented on various platforms using different approaches. Emotion mining can be performed on geographical analysis of political leader scope, political event budget, reviews of products for a business organization to improve their work further.

REFERENCES

[1] Paul Ekman, "An Argument for Basic Emotions" Cognition and Emotion, Vol. 6(3/4), 1992, pp. 169-200.
[2] Sanjeev Dhawan, Kulvinder Singh, Vandana Khanchi, "A Framework for Polarity Classification and Emotion Mining from Text", International Journal of Engineering and Computer Science ISSN: 2319-7242, Vol. 3(8), August, 2014, pp. 7431-7436.
[3] Kirk Roberts, Michael A. Roach, Joseph Johnson, Josh Guthrie and Sanda M. Harabagiu, "EmpaTweet: Annotating and Detecting Emotions on Twitter", In Proceedings of LREC, 2012, pp. 3806-3813.
[4] Abdul Hannan, "Emotion Detection from Text", International Journal of Engineering Research and Development, Vol. 11(7), Issue 07(July 2015), pp. 23-34.
[5] Despoina Chatzakou, Athena Vakali and Konstantinos Kafetsios, "Detecting variation of emotions in online activities", Expert Systems with Applications: Vol. (89), 2017, pp. 318-322.
[6] Chetan Kalyan, Min Young Kim, "Detecting Emotional Scenes Using Semantic Analysis on Subtitles", Jun 4 2009, https://nlp.stanford.edu/courses/cs224n/2009/fp/6.pdf
[7] Nilesh M. Shelke, "Approaches of Emotion Detection from Text", International Journal of Computer Science and Information Technology Research, Vol. 2(2), Issue 2, pp. 123-128, April-June 2014.
[8] Maryam Hasan, Elke Rundensteiner and Emmanuel Agu, "EMOTEX: Detecting Emotions in Twitter Messages", ASE BIGDATA/SOCIALCOM/CYBERSECURITY Conference, 27-31, May 2014.
[9] Mohammed Yassine, Hazem Hajj, "A Framework for Emotion Mining from Text in Online Social Networks", In Proceedings of 2010 IEEE International Conference on Data Mining Workshops, pp. 1136-1142.
[10] Vincent S. Erich, "Multi-Emotion Detection in Dutch Political Tweets", 26, Jun 2015, https://esc.fnwi.uva.nl/thesis/centraal/files/f696726091.pdf
[11] Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, David McClosky, "The Stanford CoreNLP Natural Language Processing Toolkit", In Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, June 23-24, 2014, pp. 55-60.

[12] Diana Maynard and Adam Funk, "Automatic Detection of Political Opinions in Tweets", In Proceedings of Extended Semantic Web Conference 2011 Workshops, Vol. 7117, May 29, 2011, pp. 88-99.

[13] Stefan Stiegliz and Linh Dang-Xuan, "The Role of Sentiment in Information Propagation on Twitter – An Empirical Analysis of Affective Dimentsions in Political Tweets", In Proceedings of the 22nd Australasian Conference on Information System (ACIS*)*, Vol. 38, 29th Nov - 2nd Dec, 2011.

[14] Kartik Singhal, Basant Agrawal and Namita Mittal, "Modeling Indian General Elections: Sentiment Analysis of political Twitter Data", Information Systems Design and Intelligent Applications, Vol. 339, 21st Jan, 2015, pp. 469-477.