Insight of Public Opinion Regarding Political Perception in India Using Sentiment Analysis on Twitter

Sayali Jori, Shraddha Phansalkar

Abstract: *Election plays an important role in democracy. People select the candidate by giving votes. The Twitter data is used to predict the outcome of an election by analyzing the sentiment of tweets about the candidates. Many agencies and media companies conduct a pre-poll survey to predict the result of an election.*

Index Terms: *Emoticons, Sentiment Analysis, Twitter.*

I. INTRODUCTION

Social media is usually formed by daily and continuous communication between participants. Different candidate's information posted on social media is analyzed to find the sentiment. The sentiment will then be used to predict the outcome of the election. The popular micro-blogging service is Twitter. Micro-blogging allows the publication of short text messages, used to share all kinds of information on Twitter and many millions of them are posted every day. Prominent in Tweets is the huge use of emoticons.

Sentiment analysis aims to detect opinions expressed regarding a given subject or topic from a text. Different classification levels in the sentiment analysis are document-level, sentence-level, and aspect-level sentiment analysis. A sentence-level sentiment analysis classifies sentiment expressed in the sentence. A document-level sentiment analysis classifies the whole document as positive or negative.

Different sentiment classification techniques are a machine learning approach, lexicon-based approach, and a hybrid approach. The machine learning approach applies the famous machine learning algorithms. A collection of known and precompiled sentiment terms known as sentiment lexicon is used in the lexicon-based sentiment analysis. It is divided into a dictionary-based approach and corpus-based approach which use statistical or semantic methods to find sentiment increases system performance and accuracy.

Emoticons act as evidence feature to enhance sentiment classification. Emoticons are ASCII art. They are sometimes called Smileys. Emoticons are formed through the creative use of letters, numbers, and punctuation symbols. Emoticons represent facial features. Emoticons square measure introduced as communicatory, non-verbal components into

the written language. They give a positive or negative sense to written sentences by a visual expression. The sentiment orientation of emoticons and messages are related to each other. These emoticons are an important source of information for polarity classification. Social media has a high percentage of positive and negative emoticons in messages. Shorter emoticons are easier and quicker to type.

Consider an associate example, "I'm going to bed early tonight".

I'm going to bed early tonight   :-)   Happy smile

I'm going to bed early tonight   :-‖   Angry

It can be seen from the above example that, the addition of an emoticon suffix greatly changes the context of the message. Emoticons imply three sentiments. Positive Sentiment, such as the happy face :). Negative Sentiment, such as the sad face :(, and Neutral Sentiment, such as undecided: /.

There is the challenge of sarcasm consider an example,

"I love Monday mornings!" Is this a positive or negative sentiment?

Consider an example "I love my dog".

What is different about these two?

With different contexts, they could mean different things.

There is research going on in the field of sentiment analysis. Instead of looking at the presence of emoticons, they parse the entire message. This is very challenging work [16].

II. LITERATURE REVIEW

Mauro Dragoni et al. [1] used NeuroSent tool for analyzing the sentiment of a document using domain specific information. It makes use of deep learning architecture where words are represented by distributed vectors. The performance of the system is evaluated by using the Dranziera protocol.

Jyoti Ramteke, Samarth Shah, Darshan Godhia, Aadil Shaikh [2] proposed a scalable machine learning model to predict the election results using two-stage frameworks. They used two-stage frameworks to create training data from twitter data without compromising on features and context.

Rui Xia, Jie Jiang, and Huihui He et al. [3] proposed a distantly supervised lifelong learning approach, for large-scale social media sentiment analysis for addressing distant supervision work in terms of continuously increasing and constantly changing topics. The continuous sentiment learning in social media can learn the knowledge from past tasks and continuously updates the knowledge as new tasks appear. The

lifelong sentiment classifier is evaluated on nine benchmark data sets.

Pritee Salunkhe, et al. [4] use twitter for making the prediction of an election result by analyzing the sentiment of political parties and candidates.

Waghode Poonam B, Prof. Mayura Kinikar [5] proposed a model that can predict public opinions using emoticons.

Suman Rani, Jaswinder Singh [6] uses a system that by mining sentiments or emotions of tweets about Indian politicians using Support Vector Machine. Unigram and TF-IDF are used as feature extractors. Measuring the performance in terms of accuracy, precision, recall, and f-measure.

Nimit Bindal, Niladri Chatterjee [8] proposed a two-step framework in which the lexicon is used to score tweets with a positive or negative polarity, tweets with low polarity strength are passed to a Support Vector Machine (SVM) classifier.

Kabir Ismail, Fatima Chiroma [7] proposed a system to predict the outcome of the 2016 US presidential candidate's election using the polarity lexicon model.

Vikash Nandi, Suyash Agrawal [9] proposed a hybrid approach of sentiment analysis; it combines both the Lexical Dictionary based approach with the features of Support vector machine learning classifier.

Ayeena Malik, Divya Kapoor, Amit Prakash Singh [10] proposed a system to analyse the tweets of various users regarding a political leader, which showcases the responses or opinions of voter's pre and post-election using AFFIN-111.

"Akshat Bakliwa" et al. [11] discuss a series of 3-class sentiment classification experiments on a set of 2,624 tweets produced during the run-up to the Irish General Elections on February 2011. The highest accuracy we achieve is 61.6% using supervised learning, and a feature set consisting of subjectivity-lexicon-based scores, Twitter-specific features, and the top 1,000 most discriminative words. This is superior to varied naive unsupervised approaches that use sound judgement lexicons to cypher Associate in nursing overall sentiment score for a try.

"Rincy Jose", et.al. [12] added semantics in feature vectors to enhance sentiment classification and then using ensemble methods for classification.

"Ali Hasan" et al. [13] introduces a hybrid approach that involves a sentiment analyser that includes machine learning. Moreover, this paper conjointly provides a comparison of techniques of sentiment analysis within the analysis of politics by applying supervised machine-learning algorithms like Naïve mathematician and support vector machines (SVM).

"Abhishek Bhola" [11], analysed the knowledge set to verify if the trivial things were additionally evident within the data collected. It was evident from our information that the political behaviour of the politicians affected their followers count and therefore quality on Twitter.

"Gayatri P. Wani", et al. [14] introduces the general user tweets from the election point of view. Here the System will study the user view of the Indian election. Based on the user's tweets system analyses if there exists a pattern between the tweets and to analyse and draw meaningful inferences from the collection of these tweets collected over a certain period; "Fuhai Chen, RongrongJi, Jinsong Su, Donglin Cao, and YueGao", et al. [15] discusses a Weakly Supervised Multimodal Deep Learning (WS-MDL) scheme towards robust and scalable sentiment prediction. The system learns convolutional neural networks iteratively and selectively from "weak" emoticon labels, which mostly contains noise.

III. APPLICATION

Political Exit Poll Prediction: Analysing messages of either political parties or politician for the prediction of election results. The tweet content is considered as a valid indicator of political sentiment. The prediction could be made by comparing the number of tweets that have positive sentiments towards each political candidate using different sentiment analysis techniques.

IV. CONCLUSION

As people increasingly use emoticons in a text in order to express their sentiment, it is a crucial piece of information and this is essential for political parties in order to better understand people feelings. The precision of recognizing emotions can increase and improve with the analysis of emoticons. With sentiment analysis, emotions are classified into positive, negative, and neutral. The results are displayed by giving the percentage of sentiment categories so that it can be concluded that a selected Twitter post get positive or negative responses based on all comments received from users. For future work, we plan to use emoticons for sentiment analysis of tweets in different domains.

REFERENCES

1. Mauro Dragoni,Giulio Petrucci, "A Neural Word Embedding's Approach For Multi-Domain Sentiment Analysis", IEEE Transactions on Affective Computing June 2017.
2. Jyoti Ramteke ; Samarth Shah ; Darshan Godhia ; Aadil Shaikh, "Election result prediction using Twitter sentiment analysis", 2016 International Conference on Inventive Computation Technologies (ICICT).
3. RUI XIA ; JIE JIANG ; HUIHUI HE, "DISTANTLY SUPERVISED LIFELONG LEARNING FOR LARGE-SCALE SOCIAL MEDIA SENTIMENT ANALYSIS", IEEE TRANSACTIONS ON AFFECTIVE COMPUTING (VOLUME: 8 , ISSUE: 4 , OCT.-DEC. 1 2017).
4. Pritee Salunkhe, Avinash Surnar, Sunil Sonawane, "A Review: Prediction of Election Using Twitter Sentiment Analysis", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 06, Issue 05, May 2017, ISSN: 2278 – 1323.
5. Waghode Poonam B,Prof. Mayura Kinikar, "Twitter Sentiment Analysis with Emoticons", International Journal Of Engineering And Computer Science ISSN:2319-7242 Volume 4 Issue 4 April 2015, Page No 11315-11321.
6. Suman Rani, Jaswinder Singh, "SENTIMENT ANALYSIS OF TWEETS USING SUPPORT VECTOR MACHINE", International Journal of Computer Science and Mobile Applications, Vol.5 Issue. 10, October- 2017.
7. Kabir Ismail Umar, Fatima Chiroma, "Data Mining for Social Media Analysis: Using Twitter to Predict the 2016 US Presidential Election", International Journal of Scientific & Engineering Research, Volume 7, Issue 10, October-2016 ISSN 2229-5518.
8. NIMIT BINDAL; NILADRI CHATTERJEE, "A TWO-STEPMETHOD FOR SENTIMENT ANALYSIS OF TWEETS", 2016 (ICIT).
9. Vikash Nandi, Suyash Agrawal, "Political Sentiment Analysis using Hybrid Approach", International Research Journal of Engineering and Technology (IRJET) Volume: 03 Issue: 05 | May-2016.
10. Ayeena Malik, Divya Kapoor, Amit Prakash Singh, "Sentiment Analysis on Political Tweets", Vth International Symposium on "Fusion of Science &

Technology", New Delhi, India, January 18-22, 2016.

11. Sayali Jori, "Prediction of Political Election Results in India using Sentiment Analysis on Twitter Reviews: Survey", IJSRD International Journal for Scientific Research & Development| Vol. 6, Issue 10, 2018 | ISSN (online): 2321-0613.

12. Rincy Jose, " Prediction of election result by enhanced sentiment analysis on twitter data using classifier ensemble Approach", Conference: 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE).

13. Ali Hasan "Machine Learning-Based Sentimental Analysis for Twitter Accounts", Math. Comput. Appl. 2018, 23, 11.

14. Gayatri P. Wani, "Analysis of Indian Election using Twitter", International Journal of Computer Applications (0975 – 8887) Volume 121 – No.22, July 2015.

15. FUHAI CHEN, "PREDICTING MICRO BLOG SENTIMENTS VIA WEAKLY SUPERVISED MULTIMODAL DEEP LEARNING" JOURNAL IEEE TRANSACTIONS ON MULTIMEDIA VOLUME 20 ISSUE 4, APRIL 2018 PAGES 997-1007.

16. http://datagenetics.com/blog/october52012/index.html.

AUTHORS PROFILE

Sayali Jori has completed BE (Computer) from Govt. College Awasari. She is currently pursuing MTech (Computer Science and Engineering) from Symbiosis Institute of Technology. Her research interest includes data mining.

Shraddha Phansalkar received her M.E (Computer science) with gold rank from Mumbai University, Ph.D. from Symbiosis International University, Pune in the area of Cloud computing. She is specialized in Distributed Systems. She works as Professor in Symbiosis Institute of Technology under Symbiosis International University.