Significance of Sentiment Analysis in Accessing Indian Elections Scenario

Amandeep Kaur[1]
Research Scholar, PhD (Comp. Appl.)
Guru Kashi University, Talwandi Sabo
Punjab, India.

Dr. Rajinder Singh[2]
Assistant Professor, UCCA
Guru Kashi University, Talwandi Sabo
Punjab, India.

ABSTRACT

In the ongoing general elections 2019 in India, technology has made it easier for political parties to gauge the mood of the voters and put their best message forward. Political parties are leveraging big data analytics to gather insights into voter preferences based on their socio-economic status, caste, local issues, and various other parameters. Based on the voter sentiment and segments, customized election campaigns with the most relevant messages and videos are being created and pitched to the specific target groups. Once the campaign is launched, data is gathered to analyze its effectiveness and tweak it further according to the response it generates. The research paper emphasized on collecting tweets and Facebook posts related to Indian elections and performing sentiments analysis on them. The paper illustrates the methodology adopted to perform sentiment analysis and reach a definite conclusion.

Index Terms – Elections, negative words, sentiment analysis, positive words, text mining.

I. INTRODUCTION

The first election making use of Big Data analytics on planet earth were US Presidential elections 2012. Electoral strategies involved expenses of $7 billion and about 10% of the amount was used for online advertising by both the parties. Both the contenders, Barack Obama and Mitt Romney, planned their own strategies. But team Obama managed the things more efficiently. They understood the gravity and complexities that Big Data would result to. Team Obama stored all their data into one single repository [1, 2].

Big data is defined as a combination of massive technological power and endlessly detailed voter information. It has certainly changed the way campaigns are conducted. Like corporations, campaigns now know far more about their constituents than ever before like, what they read, which movies they stream, which shows they watch, where they shop, which products they buy. This allows campaigns to identify their most likely voters and target them with ads and favorable content [3, 4]. The result, increasingly, is that candidates talk only to voters disposed to agree with them, as opposed to persuading those who don't. In the ongoing general elections 2019 in India, technology has made it easier for political parties to gauge the mood of the voters and put their best message forward. Political parties are leveraging big data analytics to gather insights into voter preferences

based on their socio-economic status, caste, local issues, and various other parameters. Based on the voter sentiment and segments, customized election campaigns with the most relevant messages and videos are being created and pitched to the specific target groups. Once the campaign is launched, data is gathered to analyze its effectiveness and tweak it further according to the response it generates [5, 6, 7].

Social media and WhatsApp have emerged as key influencers and political parties are making good use of these tools to reach out to the voters. Informal conversations between online users are enabling parties to pick up positive narratives and convert them into trending stories. The IT cells also monitor negative narratives and push counter stories in a bid to minimize the harm these negative sentiments can cause [8, 9]. Online browsing patterns of users are being tracked to pitch relevant political messages that are further fine-tuned according to the user engagement with these pitches. Depending on the uptake of the messages at a particular location, political parties are also creating narratives for location-specific rallies and the key messages they must deliver [10, 11, 12].

Not just the campaign, but poll strategy and fielding candidates have also become scientific with big data guiding the parties here as well. Analysis of data from numerous apps—that provide insights into voter perception of the MPs and MLAs, and review of their work—gives an insight into the popularity (or otherwise) of a candidate, enabling parties to make informed decisions about the best candidate for a constituency [13, 14, 15].

II. SENTIMENT ANALYSIS AND ITS APPLICATIONS

Sentiment Analysis is defined as a computational procedure of determining whether a writing is positive, negative, or neutral. It is also defined as opinion mining. The basic aim of sentiment analysis is to determine the attitude of an individual or group regarding a particular topic or overall context. The sentiment or attitude may be a judgement, evaluation or emotional reaction. Sentiment analysis can help you gauge opinion, which can in turn guide strategy and help decision-making. In the current business landscape, it's increasingly important that one know what our customers, competitors and employees think about the business, products and brand and sentiment analytics can help one do that — often relatively inexpensively[16, 17].

Some of the fields making use of sentiment analysis are mentioned as under.

- Politics: Sentiment analysis is used in political scenario to keep track of political views of general masses, to find the attachment of public to a political leader or political party, to find out the reason behind hatred of electorates against particular political leader or political party, and to detect consistency and inconsistency between statements and actions of political parties.

- Trade / Business – Sentiment analysis help organizations to develop strategies, to gain insight of customer's preferences, to know about their likes and dislikes, to know how people responded to their product launches and campaigns.

- Public actions – Sentiment analysis also enables to monitor social occurrences in order to gain prior idea of any kind of anti-social activities or situations and gaining insight of mood of general masses.

- Determine Which Customer Segments Have the Strongest Opinions - Sentiment analysis can tell political parties which electorates feel negative about their party for whatever reason, providing political party an opportunity to address their concerns. A political party could find out, say, that the most negative sentiment is coming from people who don't like their election manifesto. Likewise, sentiment analysis can help a political party to identify which electorates are their strongest brand ambassadors — electorates you definitely don't want to neglect [18].

- Plan Product Improvements **-** Sentiment analysis can use unstructured data to help political parties learn what people loved or didn't love about party's latest election campaign. It could tell a political party that people love the political leaders, but find it isn't as easy to use as earlier models [19].

- Determine the Most Effective Communication Channels - While consistency is important across communication channels, you may find that people respond more positively on some channels than others. Maybe the percentage of positive tweets about your brand is higher than the percentage of positive Facebook comments. Why might that be? Such information can help you re-evaluate how you interact with customers on various platforms, and make changes as necessary.

- Drive decisions. Sentiment analysis provides insight on any change in public opinion related to activities of political parties that will either support or negate the direction political party is heading. High or low sentiment scores help you identify ways to restructure teams or develop new creative strategies.

- Highlight competitive advantage. There are strategic benefits in knowing electorates sentiment related to political parties competitors. Sentiment analysis can help predict electorate's trends, so keeping a pulse on public opinion of other political parties and provides a control group to compare your scores against [20, 21].

- Predict product lifecycle. Data derived from sentiment analysis reveals how well a political party is faring in the market, how its performance can be improved, and if it's time to pull any particular candidate off the records.

Sentiment analysis is an essential tool for improving online reputation, because it takes largely unstructured data that would be a nightmare to analyze traditionally, and gives you actionable facts and figures that you can tactically and strategically address [22, 23, 24].

In addition to the definition problem, there are multiple layers of meaning in any human generated sentence. People express opinions in complex ways; rhetorical devices like sarcasm, irony, and implied meaning can mislead sentiment analysis [25]. The only way to really understand these devices are through context: knowing

how a paragraph is started can strongly impact the sentiment of later internal sentences. Most of the current thinking in sentiment analysis happens in a categorical framework: sentiment is analyzed as belonging to a certain bucket, to a certain degree. For example, a given sentence may be 45% happy, 23% sad, 89% excited, and 55% hopeful. These numbers don't add up to 100 – they're individual indications of how "X" a sentence's sentiment is [26].

III. RESULTS OBTAINED FROM SENTIMENT ANALYSIS REGARDING INDIAN LOK SABHA ELECTIONS 2019

As we complete the fifth phase of voting in this general election season, the declining trend in terms of tweets mentioning either BJP or Congress continued. In the first week of May 2019, the average count of tweets mentioning Congress at 21k was slightly higher than the average tweets mentioning BJP (19k). The last time Congress received more mention of tweets in comparison to BJP was during mid-March 2019, helped by Mr Gandhi's political rallies.

In terms of hashtags, we see that within the BJP tweets, hashtags such as #JeetegaToModiHi (Modi will win for sure) #PhirEkBaarModiSarkaar (Once more Modi government) #DeshKiPasandModi (India likes Modi) and #HarGharModiKeSaath (Every house is with Modi) were among the top five trending hashtags. The number of unique users for the BJP on twitter stood at 55,146 users from April 30th to May 5, 2019, whereas for the Congress party, unique users stood at 52,708. In the tweets mentioning Congress party, we see a continuing trend of #AbHogaNyay (Now is the time for NYAY-minimum income scheme), #NyayForEmpowerment and #NyayForJharkhand similar to the trend seen since mid-April, 2019. Mr Gandhi's interview to IndiaToday (mentioned as #RahulGandhiToIndiaToday) has been mentioned 403 from April 30 to May 5, 2019. In the same week, Mr. Modi's interview with another news agency (mentioned as #ModiOnIndiaTV) has seen a mention of 822 times.

As BJP started to project surgical strike in Balakot as the main political agenda, Congress responded by listing out surgical strikes carried out during UPA regime in a press conference held on May 4, 2019. Mr Modi replied to this claim by stating that Congress does not know the difference actual war and wars in video games. With these comments, we see that #ArmySeMaafiMaangoModi (Apologize to the Army) was a mention of 1143 times.

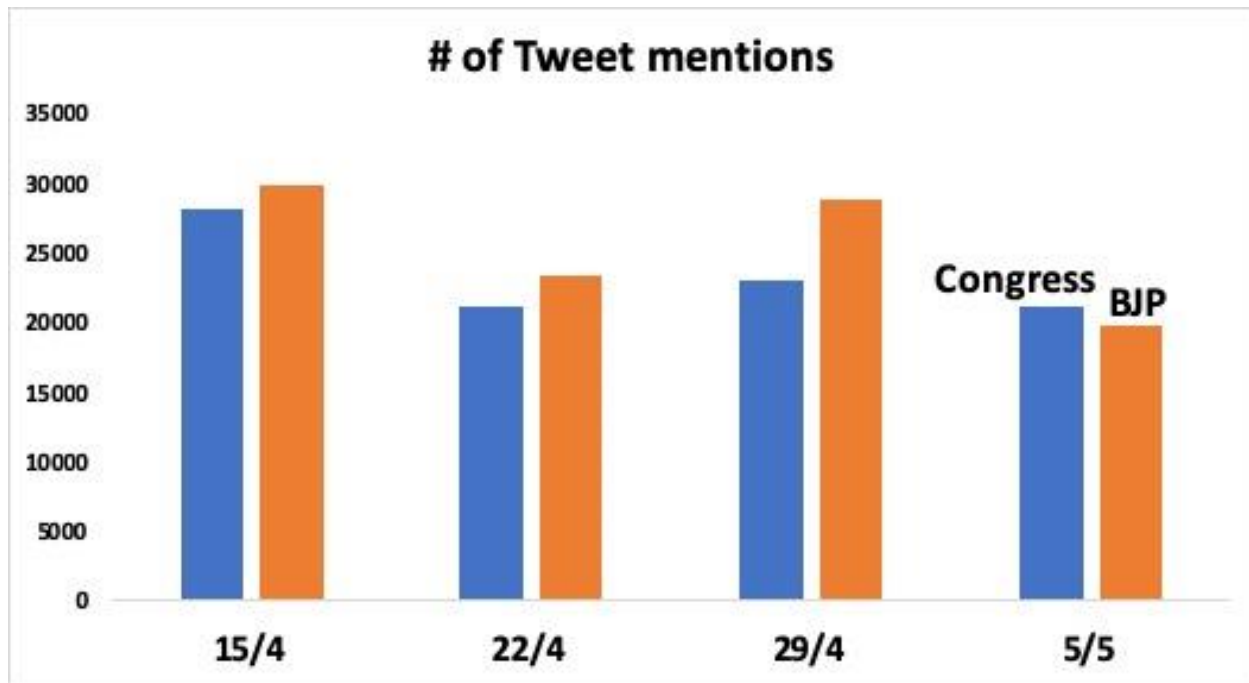Average weekly tweet mentions for Congress and BJP.

Fig. 1: Figure indicates Average weekly tweet mentions for Congress and BJP

Among the top five leaders from Congress, the Gandhi siblings continue to mention the most number of times as shown in Fig.2.
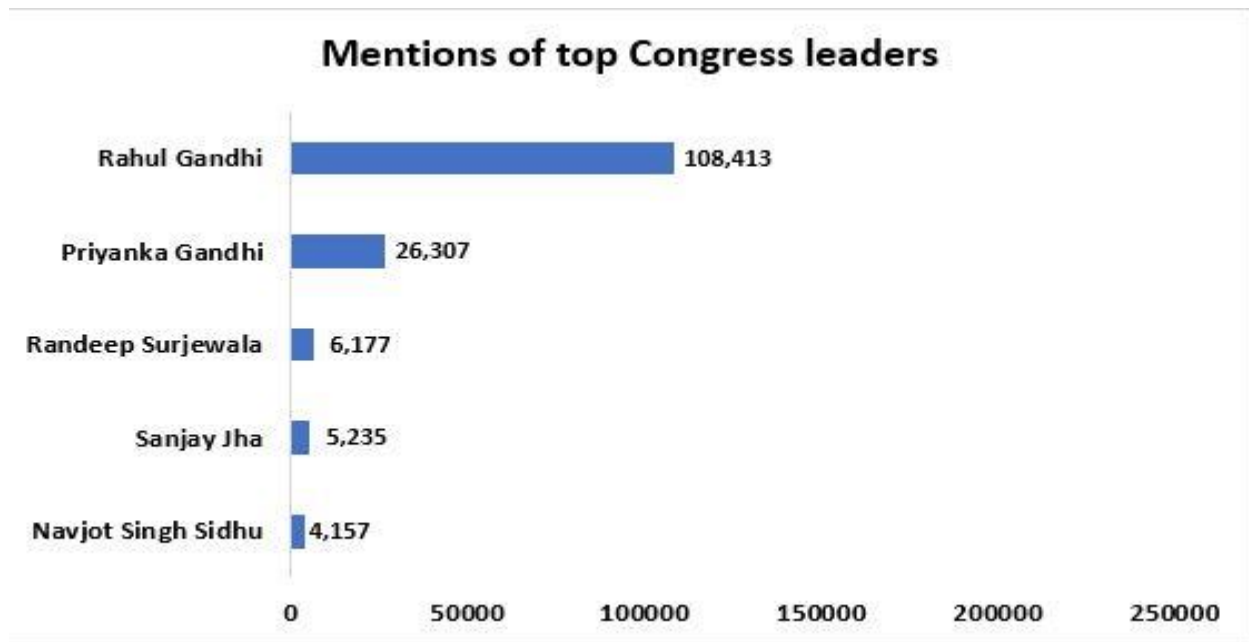

Fig. 2: Figure indicates the mention of top five Congress leaders

Within the BJP tweets, Mr. Modi and Mr Amit Shah (Party President of BJP) continued to see more mentions as shown in Fig. 3.
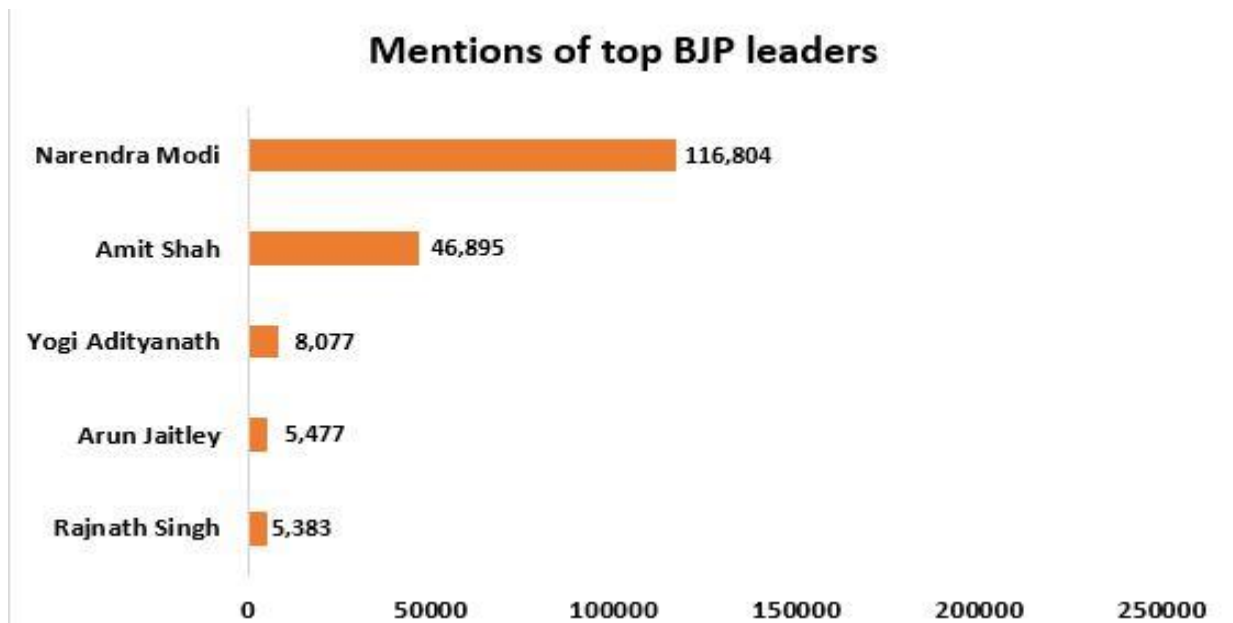
Fig. 3: Figure indicates the mention of top five BJP leaders

The following mentioned tweets are collected from #BJPVsCongress and are used to analyze the sentiment analysis of general public.

- Delhiites wiped out BJP and Congress from VS and AAP showed the will to fight the "mafias" and the Results.Citizens.
- Happy.. Hope Delhities now take this fight to the parliament to fight the "MAFIAS" .
- CM city karnal is going to be a fight between BJP vs Congress in upcoming Loksabha election
- Ok here is the data 43% bjp 42% sp + bsp. If sp and bsp loses 6% then the fight is in between 43 vs 36 dont worry about the congress it was just 7%
- It's Only AAP Vs BJP in Delhi: Congress Plays as Spoiler | Technique Today rediction by very well known Astrologer, Pandit Raj kumar Sharma. He appears very frequently on TV9 Maharashtra and gives political predictions.
- Saffron vs Saffron: Battle for Bhopal LS Seat. How Congress veteran Digvijaya Singh donned saffron robe to counter BJP's Sadhvi Pragya Thakur
- company Ke professional employee v BJP Vs Congress v bate hue hain.. Pata Nahi kal requirement me Ye question na puch le ki Tum BJP supporter ho ya Congress ke....Jai hind
- BJP because of one point Merit vs Reservation. I have no issue with Congress but Merit. Congress started reservation and is there for 75 years. Time to build merit based country. Sagar USA
- If one picture could perfectly define the "mahaul" in West Bengal, it's this picture. I clicked this in Birbhum and one can see this scene all over West Bengal. One thing's for sure, it's a clear BJP vs TMC contest and both CPM/Congress have become negligible. Advantage BJP.
- Delhi, Don't waste your vote on Congress. If you don't want to see Modi again as PM. Vote for AAP. In Delhi, Only AAP can rout BJP from power.

- Apart from @AjayMaken who's contesting from my constituency of New Delhi, I don't see any buzz around any congress candidate on ground.

- Scary 1996-98 UF scenario. Admission that Congress is fading away. Not good for an India that's a world leader today. Has to be Congress Vs BJP however much either of them has shortcomings

- Congress has not allowed SP/BSP to build a momentum. BJP narrative Modi vs Chaos is selling. Elections are run on atmospherics & perception. This mistake would evolve into a Modi Wave this May. BJP will end up winning more than half the seats. Get 50+ in UP, courtesy @INCIndia

- Sunny Deol Vs @sunilkjakhar in Gurdaspur will be a great contest. It is an old BJP seat which Congress snatched in a by-poll after Vinod Khanna passed away. Nationalism is a major issue here -- the area saw a terror attack from Pakistan (Dinanagar)

- And for info of journos cheering for Sheila Tai & claiming that it is BJP Vs Congress in North East Delhi - even in 2014 Jai Parkash Aggarwal, Congi bigwig (Ex DPCC President), ended at distant third in N.E. Delhi, although helped BJP win by cutting AAP's vote.

- BJP is treading the right path. No wonder then that it has been appreciated by Hindus across the nation. In Bhopal, BJP fields

- Malegaon blast accused Sadhvi Pragya vs Congress' Digvijaya Singh | Elections News, The Indian Express

- 2 Days back it was Borivali station. Today it was Liberty Garden in Malad where people chanting Modi- Modi in Urmila Road Show.

- What was BJP's flag doing on Election Commission's vehicle in Asansol? Now its obvious that the fight is between "BJP + TMC + Congress + Media + Election Commision vs Left Front" in Bengal. We urge all the secular and democratic forces to #Vote4Left ?

- Seat-by-seat view of UP shows: 1 Congress leaving over 50 seats for MGB 2 Better deal between Cong & SP, RLD. Main SP vs Cong battle in 2 seats 3 But BSP & Cong are harming each other in a few seats 4 If such adjustments work, BJP is staring at a rout

- I have sent word to Rajnath to revise the Manifesto on two blunder:a) Doubling of farmers income by 2022 means a 24% per year growth rate which is unbelievable given past world record, best can be 10% per year. b)India's comparable GDP is 3rd largest not 6th.

The steps adopted for performing sentiment analysis on election relevant data are mentioned as under.

1. An unstructured database has been constructed by collecting tweets, messages, and comments from social media in a text file.

2. A dictionary is constructed which comprises of positive and negative words. Each word is assigned a score depending upon the intensity of the word. For instance, abandon -2, abandoned -2, abandons -2, abducted -2, affection 3, affectionate 3, agrees   1 etc.

3. A source code in Java is developed and made to run in Eclipse. The keywords in the constructed database are extracted and on the basis of the occurrence and the score assigned to them in the dictionary, the value of each tweet is calculated.

4. The values of all the tweets are finally calculated and the final score is obtained. If the final score is a positive value, then it can be concluded that the overall impact of the collected tweets ended on a positive note and vice versa.

The obtained results after executing the constructed code developed in Java are shown below.

{Delhiites wiped out BJP and Congress from VS and AAP showed the will to fight the "mafias" and the Results.Citizens Happy.. Hope Delhities now take this fight to the parliament to fight the "MAFIAS" .=-3.0}

Current Counter score=-3.0

{CM city karnal is going to be a fight between BJP vs Congress in upcoming Loksabha election=-1.0}

Current Counter score=-4.0

{Ok here is the data 43% bjp 42% sp + bsp. If sp and bsp loses 6% then the fight is in between 43 vs 36 dont worry about the congress it was just 7%=-4.0}

Current Counter score=-8.0

{It's Only AAP Vs BJP in Delhi: Congress Plays as Spoiler | Technique Today=0.0}

Current Counter score=-8.0

{rediction by very well known Astrologer, Pandit Raj kumar Sharma. He appears very frequently on TV9 Maharashtra and gives political predictions.=0.0}

Current Counter score=-8.0

{=0.0}

Current Counter score=-8.0

{Saffron vs Saffron: Battle for Bhopal LS Seat. How Congress veteran Digvijaya Singh donned saffron robe to counter BJP's Sadhvi Pragya Thakur=0.0}

Current Counter score=-8.0

{company Ke professional employee v BJP Vs Congress v bate hue hain.. Pata Nahi kal requirement me Ye question na puch le ki Tum BJP supporter ho ya Congress ke....Jai hind=1.0}

Current Counter score=-7.0

{BJP because of one point Merit vs Reservation. I have no issue with Congress but Merit. Congress started reservation and is there for 75 years. Time to build merit based country. Sagar USA=0.0}

Current Counter score=-7.0

{If one picture could perfectly define the "mahaul" in West Bengal, it's this picture. I clicked this in Birbhum and one can see this scene all over West Bengal. One thing's for sure, it's a clear BJP vs TMC contest and both CPM/Congress have become negligible. Advantage BJP.=4.0}

Current Counter score=-3.0

{Delhi, Don't waste your vote on Congress. If you don't want to see Modi again as PM. Vote for AAP. In Delhi, Only AAP can rout BJP from power.=0.0}

Current Counter score=-3.0

{Apart from @AjayMaken who's contesting from my constituency of New Delhi, I don't see any buzz around any congress candidate on ground.=0.0}

Current Counter score=-3.0

{Scary 1996-98 UF scenario. Admission that Congress is fading away. Not good for an India that's a world leader today. Has to be Congress Vs BJP however much either of them has shortcomings=3.0}

Current Counter score=0.0

{Congress has not allowed SP/BSP to build a momentum. BJP narrative Modi vs Chaos is selling. Elections are run on atmospherics & perception. This mistake would evolve into a Modi Wave this May. BJP will end up winning more than half the seats. Get 50+ in UP, courtesy @INCIndia=4.0}

Current Counter score=4.0

{unny Deol Vs @sunilkjakhar in Gurdaspur will be a great contest. It is an old BJP seat which Congress snatched in a by-poll after Vinod Khanna passed away. Nationalism is a major issue here -- the area saw a terror attack from Pakistan (Dinanagar)=-1.0}

Current Counter score=3.0

{And for info of journos cheering for Sheila Tai & claiming that it is BJP Vs Congress in North East Delhi - even in 2014 Jai Parkash Aggarwal, Congi bigwig (Ex DPCC President), ended at distant third in N.E. Delhi, although helped BJP win by cutting AAP's vote.=5.0}

Current Counter score=8.0

{BJP is treading the right path. No wonder then that it has been appreciated by Hindus across the nation. In Bhopal, BJP fields Malegaon blast accused Sadhvi Pragya vs Congress' Digvijaya Singh | Elections News, The Indian Express=0.0}

Current Counter score=8.0

{2 Days back it was Borivali station. Today it was Liberty Garden in Malad where people chanting Modi- Modi in Urmila Road Show.=0.0}

Current Counter score=8.0

{What was BJP's flag doing on Election Commission's vehicle in Asansol? Now its obvious that the fight is between "BJP + TMC + Congress + Media + Election Commision vs Left Front" in Bengal. We urge all the secular and democratic forces to #Vote4Left ?=-1.0}

Current Counter score=7.0

{Seat-by-seat view of UP shows: 1 Congress leaving over 50 seats for MGB 2 Better deal between Cong & SP, RLD. Main SP vs Cong battle in 2 seats 3 But BSP & Cong are harming each other in a few seats 4 If such adjustments work, BJP is staring at a rout=-3.0}

Current Counter score=4.0

{I have sent word to Rajnath to revise the Manifesto on two blunder:a) Doubling of farmers income by 2022 means a 24% per year growth rate which is unbelievable given past world record, best can be 10% per year. b)India's comparable GDP is 3rd largest not 6th.=4.0}

Current Counter score=8.0

Final Counter score=8.0

Sentiment Analysis ended on a positive note

IV.   CONCLUSION

The research paper discussed the importance of sentiment analysis and its applications in different sectors alongside elections. The paper illustrated and emphasized on analyzing tweets and Facebook posts to have insight of sentiments of particular audience for particular political leader or political party. Sentiment analysis acts as a guidance to those electorates who are still confused to decide that whom they should vote for.

REFERENCES

1. Mohey El-Din Mohamed Hussein, D.: A survey on sentiment analysis challenges. J. of King Saud Univ. – Engg. Sci., 2016; 1 – 9.

2. Jagdev G. Sentiment Analysis and its Impact in Modeling Election Scenario. International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), 2018;5(2): 22-27.

3. Bermingham A., Smeaton A. F. Classifying sentiment in microblogs: is brevity an advantage? In Proc. of the 19th ACM int. conf. on Inform. and Know. Manag. 2010; 1833 – 36.

4. Sharma Y., Mangat V., Kaur M. A Practical Approach to Sentiment Analysis of Hindi Tweets. In Proc. of the 2015 1st Int. Conf. on Next Gen. Comp. Techno (NGCT). 2015; 677 – 80.

5. Jagdev G. et al. A Comparative study of Conventional Data Mining Algorithms against Map-Reduce Algorithm. International Journal of Advance Research in Science and Engineering (IJARSE), 2017; 6(5): 325-35.

6. Elghazaly T., Mahmoud A., Hefny H.A. Political Sentiment Analysis Using Twitter Data. In Proc. of the Int. Conf. on Inter. of things and Cloud Comp.2016. 1-5.

7. Sharma, P., Moh, T-S. Prediction of Indian Election Using Sentiment Analysis on Hindi Twitter. 2016 IEEE Int. Conf. on Big Data (Big Data). 2016; 1966 – 71.

8. Jagdev G. et al. Analyzing Maneuver of Hadoop Framework and MapR Algorithm Proficient in supervising Big Data. IJATES 2017; 5 (5): 372-82.

9. Kharde, V. A., Sonawane, S.S. Sentiment Analysis of Twitter Data: A Survey of Techniques. Int. J. of Comp. App. 2016; 139(11): 5 – 15.

10. Medhat, W., Hassan, A. Korashy, H.: Sentiment analysis algorithms and applications: A survey. A. S. Engg. J. 5, 2014; 1093–1113.

11. Kaur Amandeep, Jagdev G.. Exploring Application of Big Data in Elections – From Data to Action, International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), 2017; 4(4): 64-71.

12. Niyogi, M., Pal, A. K. Discovering conversational topics and emotions associated with Demonetization tweets in India. ArXiv: 1711.04115v1 [cs.CL]. 2017; 1 – 6.

13. Vijayarani, Dr. S., Ilamathi, J., Nithya, M.: Preprocessing Techniques for Text Mining - An Overview. Int. J. of Com. Sci. & Comm. Net. 2015;5(1): 7-16.

14. Mohammad S. M., Kiritchenko S. Using Hashtags to Capture Fine Emotion Categories from Tweets. Sp. issue on Semantic Analy. in Soc. Med., Comp. Intell. 2013; 1 –22.

15. Jagdev G. et al. A Study of Clustering and Classification Techniques involved in Data Mining. IJATES. 2017; 5(5): 360-71.

16. Jagdev G. et al. Analyzing and Filtering Big Data concerned with elections via Hadoop Framework. IJARSE. 2017; 6 (4): 781-90.

17. Vu Dung Nguyen, Blesson Vaghese. Royal Birth of 2013:Analysing and Visualising Public Sentiment in the UK using Twitter. Research Gate, 2013.

18. Jagdev G. et al. Comparing Conventional Data Mining Algorithms with Hadoop based Map-Reduce Algorithm considering elections perspective. International Journal of Innovative Research in Science and Engineering (IJIRSE). 2017; 3 (3): 57-68.

19. Shuri Sarah E., Yaghi Rawan I. Twitter Sentiment Analysis: Case Study In Automotive Industry. IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT). 2015. 1-5.

20. Jagdev G. et al. Big Data Proposes an Innovative concept for contesting elections in Indian Subcontinent. IJSTA. 2015. 1(3): 23-28.

21. Muherjee Subhabrata, Malu Akshat. TwiSent:A MultiStage System For analyzing Sentiment in Twitter. Proceedings of the 21st ACM International conference on Information and knowledge management. 2012. 1–4.

22. Jagdev G. et al. Excavating Big Data associated to Indian Elections Scenario via Apache Hadoop. IJARCS. 2016; 7(6):117-23.

23. Blessy Selvam, Abirami S. A Survey on OM Framework. IJARCCE. 2013; 2(9).

24. Lovelin Ponn Felciah M., Nbuselvi R.A. A Study on Sentiment Analysis of Social Media Reviews. IEEE econd Conference on Innovations in Information Embedded & Communication Systems, 2015. 1–3.