

TWITTER SENTIMENT ANALYSIS ON 'INDIA CHINA BORDER DISPUTE'

Sangeeta Rani^{1*} and Nasib Singh Gill²

^{1,2}Department of Computer Science and Application, M D University, Haryana, India.

Article Received on 19/03/2021

Article Revised on 09/04/2021

Article Accepted on 30/04/2021

*Corresponding Author

Sangeeta Rani

Department of Computer
Science and Application, M
D University, Haryana,
India.

ABSTRACT

Social networking sites are frequently used by common men to express their opinion on different issues worldwide. Twitter is one of the social networking sites used by Twitter users for putting their comments on various topics. This information collected can be used to find people's opinion or sentiment that can further help in various analysis and

decision support. The present paper collects and analyzes 10000 real-time tweets related to 'India China Border Dispute' collected from 17 to 21 June 2020. Dictionary-based classifier is used to find the opinion of people as positive, neutral or negative on India China border issue. A lexicon having 4787 negative and 2008 positive words is used in dictionary based classification technique. RStudio tool is used for implementation and analysis. The results show that 58.65 % people are tweeting negatively, 31.37 % tweets are neutral and only 9.98 % tweets are positive. This shows the criticality of the issue and people's concern about the matter, as maximum number of people are tweeting negatively. The issue should be resolved for border peace and tranquility.

KEYWORDS: Dictionary Based Classifier, Lexicon Classifier, Opinion Mining, Twitter Sentiment Analysis, 'India China Border Dispute'.

1. INTRODUCTION

Twitter is a popular social networking site and is frequently used by Twitter users to put their opinion on a public platform on various issues. This leads to a huge repository of data available for research. Sentiment analysis or opinion mining is a way of finding the sentiment

of text or tweets as negative, positive or neutral.^[1-5] Sentiment analysis using twitter is an easy and effective way of predicting people's opinion and helps with decision making for various issues. Although the sample space collected in the form of tweets for analysis is not universal as all the people may not be using Twitter, but this can be one way of finding people's opinion that can complement other methods of sentiment analysis.

Various techniques are used for Twitter sentiment classification like dictionary-based classifier, Machine Learning classifier and ensemble classifiers. Dictionary-based classifier is unsupervised classification techniques that is easy to implement, but its efficiency is not as good as machine learning and ensemble classifiers.^[6] SVM (Support Vector Machine), KNN (K-Nearest Neighbor), NB (Naïve Bayes), Decision tree etc. are various machine learning classifiers used for tweet classification and are very effective with good accuracy.^[7] Ensemble classification techniques like Stack-based classifiers, Bagging and Boosting are used to further improve classification accuracy.^[8-9] Various Preprocessing techniques^[10] and feature selection techniques^[11-12] are used to prepare the tweets before actual sentiment analysis to optimize the results.

Twitter sentiment analysis is used in many real time applications to find people's opinion on such things as Business, Sports, Education, Election result prediction, Healthcare, tourism and many more. Twitter opinion mining is also used in disaster management for collecting people's opinion and situational analysis through real time tweets. Geospatial analysis is also possible by collecting location based tweets. In the present paper, Tweets related to 'India China Border Dispute' are analyzed to predict people's opinion. Section 1 gives the introduction of sentiment analysis. The literature survey of related work and different real-time application of twitter sentiment analysis is mentioned in section 2. The methodology and Data Set used for the analysis is mentioned in Section 3. Section 4 discusses the result and analysis part. The research work is concluded in Section 5.

2. Related Work

Twitter sentiment analysis has a long list of applications. Several researches have already been done and ongoing in this area. The motive of research is either to improve classification efficiency or to use it in different real-life applications to do valuable analysis. Twitter sentiment analysis is used in various Business prediction^[13-14], health care applications^[15-16], Tourism^[17-18] and Games^[19] etc. Twitter sentiment analysis is also used in election prediction and analysis of various government policies.^[20-23] V. Curtis et al.^[24] in his research analyzed

the 'Clean India' campaign to find people's opinion about it. Various researches have been published for the use of Twitter sentiment analysis in disaster management.^[25-26] A. Dudani et al.^[27] used Twitter sentiment analysis for Kerala Flood monitoring. A. Karami et al.^[28] used Twitter for Situational Awareness for Disaster management. Location-based Tweets can also be downloaded and analyzed for more specific analysis.^[29-32] N. Naveed et al.^[33] used geospatial location based tweets sentiment analysis for Topics Detected in Social Media. O. Nazan et al.^[34] used Sentiment Analysis on Twitter for Syrian Refugee Crisis. Vast range of applications use Twitter sentiment analysis. Currently, India-China relation and border issue are hot and critical issue. In the present research, real time tweets related to 'India China Border Dispute' are downloaded and analyzed. The paper shows the awareness and concern of people about the matter.

3. Methodology and Data Set

In the present research work, real time tweets related to 'India China Border Dispute' collected from 17 June to 21 June 2020 are downloaded by using Twitter API. The date wise tweets used in implementation are mentioned in Table 1. Total 10000 tweets related to the issue are downloaded for sentiment analysis. Tweets are cleaned, preprocessed and represented as BOW. Dictionary based classification technique is used for finding opinion of people about 'India China Border Dispute' in terms of sentiment score as negative, positive or neutral. Tweets with sentiment score above zero are categorized as positive, zero as neutral and below zero as negative. A lexicon having 4787 negative and 2008 positive words is used in dictionary based classification technique. RStudio tool is used for implementation and analysis.

Table 1: Details of Tweets.

Date	Number of Tweets
17/6/2020	581
18/6/2020	4485
19/6/2020	1889
20/6/2020	1508
21/6/2020	1537

4. RESULTS AND ANALYSIS

Sentiment score is extracted by applying dictionary-based classifiers on the tweets mentioned in Table 1. Opinion of tweets as positive, negative and neutral from 17 June to 21 June 2020 is given in Table 2. The maximum number of tweets is tweeted on 18 June 2020, and 60% of

those tweets are negative. From total tweets, 58.65 % of tweets are negative, and 31.37 % are neutral. Only 9.98 % of tweets are positive. People are very much concerned about the issue and are in a negative state of mind. Figure 1 shows the sentiment score of each tweet in a range from -4 to 3. Figure 2 shows the variation in the number of tweets and sentiment of tweets with date from 17 to 21 June 2020. Figure 3 shows the frequency of tweets v/s sentiment score in a range from -4 to 3. The bar chart shows the maximum number of negative tweets with sentiment score as -1. Visualization of tweets in the form of a word cloud is shown in Figure 4. The results show that maximum numbers of people are tweeting negatively. It is a very sensitive issue regarding the India-China relationship, and people are very much concerned about the matter. The matter is very critical, and China is a very important neighbor country in terms of business, economy and border tranquility.

Table 2: Sentiment of Tweets.

Date	Positive	Negative	Neutral
17/6/2020	36	314	231
18/6/2020	357	2964	1164
19/6/2020	180	1227	482
20/6/2020	202	709	597
21/6/2020	223	651	663

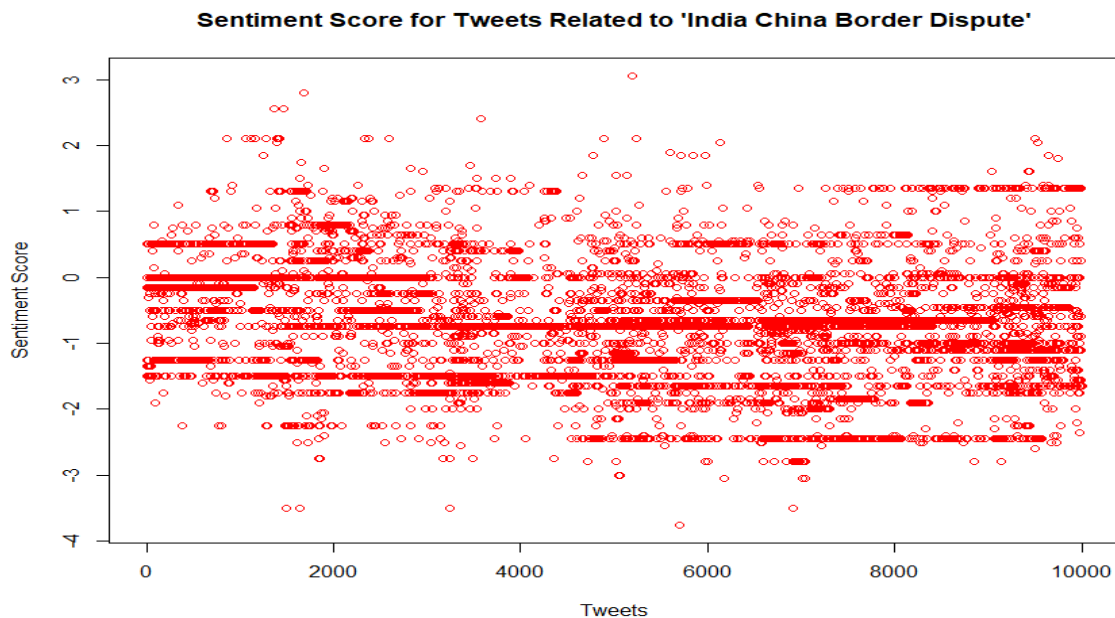


Fig 1: Tweets Sentiment Score from Date 17 to 21 June 2020.

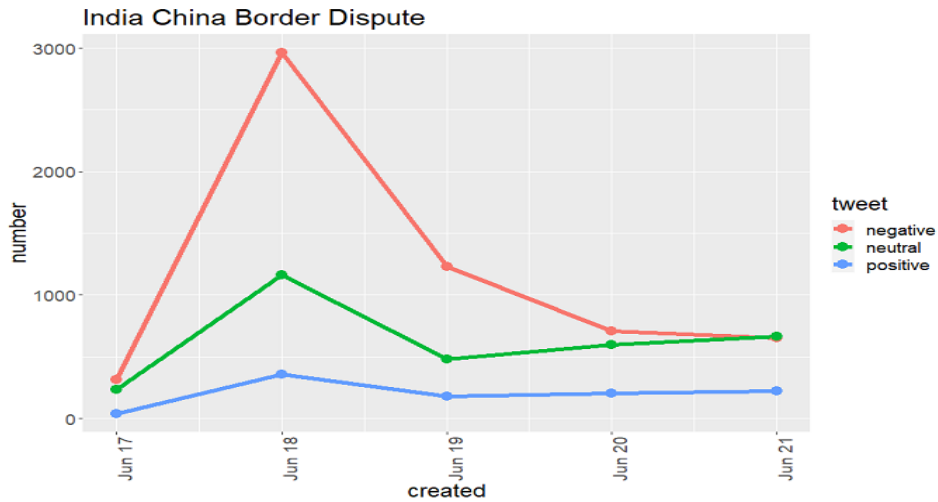


Fig 2: Comparative Analysis of Tweets Sentiment from Date 17 to 21 June 2020.

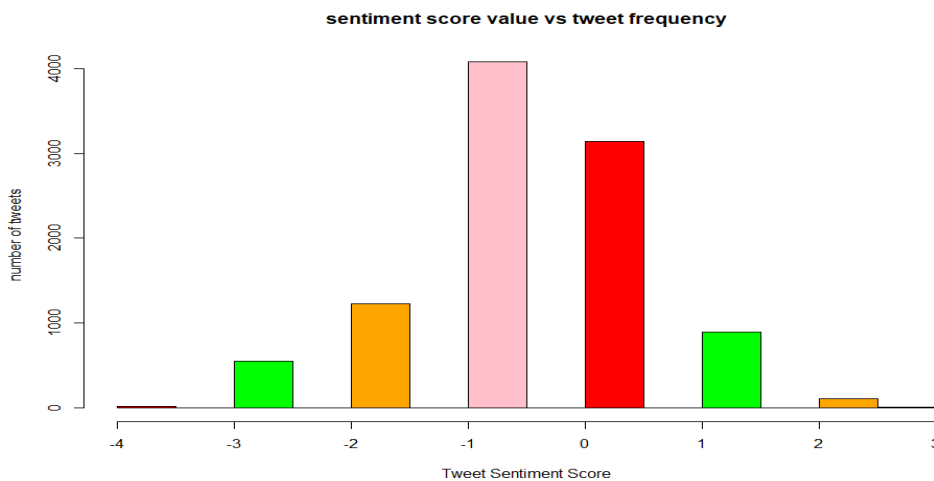


Fig 3: Sentiment Score of Tweets related to ‘India China Border Dispute’



Fig. 4: Visualization of Tweets as Word Cloud related to ‘India China Border Dispute’

5. CONCLUSION

Twitter sentiment analysis has a wide variety of real-life applications to find people's opinion about various issues. This can be very helpful to decision support and policy making. The present paper downloads real-time tweets related to 'India China Border Dispute', preprocess it and do sentiment analysis of this data. The paper analyzes date, frequency and sentiment of tweets from 17 to 21 June 2020. The results show that 58.65 % of tweets are negative, 31.37 % are neutral and only 9.98 % of tweets are positive. This shows the criticality of the issue and people's concern about the matter, as maximum number of people are tweeting negatively.

REFERENCES

1. Agarwal, B. Xie, I. Vovsha, O. Rambow, R. Passonneau, "Sentiment Analysis of Twitter Data", In Proceedings of the workshop on Languages in Social Media, Columbia University, New York, ISBN: 978-1-932432., 2011; 96(1): 30-38.
2. W. Medhat, A. Hassan, H. Korashy, "Sentiment analysis algorithms and applications: A survey", in Ain Shams Engineering Journal, Dec 2014; 5(4): 1093-1113.
3. S. Maheshwari, S. Shukla, D. Kumari, "Twitter Opinion Mining Using Sentiment Analysis", World Scientific News, An International Scientific Journal, 2019; 78-87.
4. N. Rajput, S. Chauhan, "Analysis of Various Sentiment Analysis Techniques", International Journal of Computer Science and Mobile Computing, Feb 2019; 8(2): 75-79.
5. D. Mohey E. M. Hussein, "A survey on Sentiment Analysis Challenges", Journal of King Saud University – Engineering Sciences, Oct 2018; 30(4): 330-338.
6. D. Ray, "Lexicon Based Sentiment Analysis of Twitter Data", International Journal for Research in Applied Science & Engineering Technology (IJRASET), Oct 2017; vol. 5, no. X.
7. S. Kurnaz, M. A. Mahmood, "Sentiment Analysis in Data of Twitter using Machine Learning Algorithms", International Journal of Computer science and Mobile Computing, March 2019; 8(3): 31-35.
8. O. Kolchyna, Tharsis T. P. Souza, P. Treleaven, T. Aste, "Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination", Cornell University Library, arXiv: 1507.00955, 18 Sep 2015.
9. K. Lakshmi Devi, P. Subathra, P. N. Kumar, "Tweet Sentiment Classification Using an Ensemble of Machine Learning Supervised Classifiers Employing Statistical Feature

- Selection Methods”, Proceedings of the Fifth International Conference on “Fuzzy and Neuro Computing (FANCCO - 2015)”, Nov 2015; 415: 1-13.
10. M. Khader, A. Awajan, and G. Al-Naymat, “The Impact of Natural Language Preprocessing on Big Data Sentiment Analysis”, International Arab Journal of Information Technology, 2019; 16: 506-513.
 11. R. Mansour, M. F. A. Hady, E. Hosam, H. Amr, and A. Ashour, “Feature Selection for Twitter Sentiment Analysis: An Experimental Study”, International Conference on Intelligent Text Processing and Computational Linguistics, Springer, 2015; 92-103.
 12. R. Ahujaa, A. Chuga, S. Kohlia, S. Guptaa, and P. Ahujaa, “The Impact of Features Extraction on the Sentiment Analysis”, International Conference on Pervasive Computing Advances and Applications – PerCAA 2019, Procedia Computer Science, 2019; 152: 341–348. Class Association
 13. M. Noori, and J. Maktoubian, “Business Improvement Approach Based on Sentiment Twitter Analysis: Case Study”, in EAI Endorsed Transactions on Cloud Systems, Apr 2019; 1-7.
 14. Rasool, R. Tao1, K. Marjan, T. Naveed, “Twitter Sentiment Analysis: A Case Study for Apparel Brands”, IOP Conf. Series: Journal of Physics, 2019; 1176: 1-7.
 15. M. M. Müller, M. Salathé, “Crowd breaks: Tracking Health Trends Using Public Social Media Data and Crowd sourcing”, Frontiers in public health, 12 Apr. 2019; 7: 81.
 16. Xiang Ji, Soon Ae Chun, James Geller, “Monitoring Public Health Concerns Using Twitter Sentiment Classifications”, in IEEE International Conference on Healthcare Informatics, 2013; 978-0-7695-5089-3/13, pp. 235-244.
 17. R. Alaei, S. Becken, B. Stantic, “Sentiment Analysis in Tourism: Capitalizing on Big Data”, Journal of Travel Research, Feb 2019; 58(2): 175-191.
 18. J. J. Padilla, H. Kavak, C. J. Lynch, R. J. Gore, S. Y. Diallo, “Temporal and spatiotemporal investigation of tourist attraction visit sentiment on Twitter.” PloS one, 14 Jun. 2018; 13, no 6.
 19. E. Eryarsoy, D. Delen, “Predicting the Outcome of a Football Game: A Comparative Analysis of Single and Ensemble Analytics Methods”, Proceedings of the 52nd Hawaii International Conference on System Sciences, Jan 2019; 1107-1115.
 20. L. Terán, J. Mancera, “Dynamic Profiles using Sentiment Analysis and Twitter Data for Voting Advice Applications”, in Government Information Quarterly, ELSEVIER, July 2019; 36(3): 520-535.

21. V. Rekha, R. Raksha, P. Patil, N. Swaras, G. L. Rajat, "Sentiment Analysis on Indian Government Schemes Using Twitter data", International Conference on Data Science and Communication (IconDSC), IEEE, Aug.
22. Naiknaware, Bharat R., Seema S. Kawathekar and Sachin N. Deshmukh. "Sentiment Analysis of Indian Government Schemes using Twitter Datasets.", in IOSR Journal of Computer Engineering, 2017; 70-78.
23. P. Verma, S. Jamwal, "Mining Public Opinion on Indian Government Policies using R", in International Journal of Innovative Technology and Exploring Engineering (IJITEE), Jan 2020; 9(3).
24. V. Curtis, "Explaining the outcomes of the 'Clean India' campaign: institutional behavior and sanitation transformation in India", BMJ Global Health Journals, 2019; 4(5): 1-11.
25. N. Pourebrahim, S. Sultana, J. Edwards, A. Gochanour, S. Mohanty, "Understanding communication dynamics on Twitter during natural disasters: A case study of Hurricane Sandy", In International Journal of Disaster Risk Reduction, July 2019; 37.
26. Alfarrarjeh, S. Agrawal, S. H. Kim and C. Shahabi, "Geo-Spatial Multimedia Sentiment Analysis in Disasters," 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Tokyo, 2017; 193-202.
27. Dudani, V. Srividya, B. Sneha, B.K. Tripathy, "Sentiment Analysis on Kerala Floods", In International Conference on Innovative Computing and Communications. Advances in Intelligent Systems and Computing, vol. 1087. Springer, Singapore, 2020.
28. Karami, V. Shah, R. Vaezi, A. Bansal, "Twitter Speaks: A Case of National Disaster Situational Awareness", In Journal of Information Science, 2017.
29. Z. Gong, T. Cai, J. Thill, S. Hale, M. Graham, "Measuring Relative Opinion from Location-Based Social Media: A Case Study of the 2016 U.S. Presidential Election", in arXiv, Feb 2020.
30. Agarwal, Ritu Singh & Durga Toshniwal (2018), "Geospatial Sentiment Analysis Using Twitter Data for UK-EU Referendum", Journal of Information and Optimization Sciences, 39(1): 303-317.
31. T. B. Mirani and S. Sasi, "Sentiment Analysis of ISIS Related Tweets Using Absolute Location," 2016 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, 2016; 1140-1145.
32. T. B. Mirani and S. Sasi, "Sentiment Analysis of ISIS Related Tweets Using Absolute Location," 2016 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, 2016; 1140-1145.

33. N. Naveed, M. Abbas, Z. Rauf, “Location Based Sentiment Mapping of Topics Detected in Social Media”, in Journal of Applied and Emerging Sciences, 2018; 8(2).
34. O. Nazan, A. Serkan, “Sentiment Analysis on Twitter: A text Mining Approach to the Syrian Refugee Crisis”, Telematics and Informatics, 2017.