

The Emotion Analysis of Indian Political Tweets using Machine Learning

Parth Pitram Sharma, Mansi Trambaklal Vegad

Parul Institute of Engineering and Technology-Diploma Studies, Vadodara, India

ARTICLE INFO

Article History:

Accepted: 25 March 2024

Published: 12 April 2024

Publication Issue

Volume 10, Issue 2

March-April-2024

Page Number

462-468

ABSTRACT

In this day and age web-based entertainment is a major region for information examination and exploration work. For Feeling Examination, I select Tweeter handle. I use Tweepy for getting to tweeter information. I perform opinion examination on Indian Political information. I got 117545 tweets of 2019 Indian Political race. I use SVM (Backing Vector Machine) Classifier for feeling Examination. Feeling assessment oversees recognizing and portraying evaluations or sentiments conveyed in source message. Electronic diversion is creating an enormous proportion of feeling rich data as tweets, sees, blog sections, etc. Feeling examination of this client made data is especially useful in knowing the appraisal of the gathering. Twitter feeling assessment is problematic stood out from general assessment examination on account of the presence of work related conversation words and erroneous spellings. The most outrageous limitation of characters that are allowed in Twitter is 140. Data base philosophy and AI approach are the two frameworks used for separating suppositions from the text. In this paper, we endeavor to analyze the twitter posts about electronic things like mobiles, workstations, etc using AI approach.

Index Terms - Sentiment Analysis, Machine Learning, Twitter Data Analysis, Classification, Political Analysis, Emotion Analysis

I. INTRODUCTION

Incline Analysis is one of the game plans of Machine learning and NLP (Natural language handling). Assessment Analysis truly do explore and distinguishing proof of the data which is given by the researchers and clients. Journalists offer the viewpoints about any item or any point and Sentiment Analysis can dissect and distinguish the feelings. Feeling

Analysis ordinarily worried about the assessment and feelings from the text/information. It dissects and demonstrates the feeling of essayist as for an information. Feeling Analysis is utilized to remove and order sentiments from various substance structures, including news, surveys, reviews and articles. It arranges contents that are positive, negative or impartial.

Twitter

Twitter gives a phase which partners people to each other paying little mind to how far they are. Twitter gives the message office and this message called "tweet". This credit goes to virtual amusement. Virtual amusement is the groundwork of sharing and getting information, data, as well as correspondence among people. They share their thinking, moving news, considerations, approaches to acting and suppositions related any subject. It is solid weapon of growing composition, and investigation. There are different critical web based life stage yet twitter is the most grounded stage for evaluation assessment considering how there are an enormous number of overall novel clients, millions bit by bit strong clients and millions posts consistently. They show their sentiments and they are taken an interest on different topics through the twitter posts. Tweets are important for Feeling Investigation enlightening assortments. The Tweets (twitter data) data can be gotten from Twitter in a very protected and basic way. We can without a very remarkable stretch help the tweets through twitter Programming interface. (Application Programming Point of interaction). Tweepy is an open source Python group that gives you an incredibly supportive technique for getting to the Twitter Programming interface with Python. Tweepy integrates a lot of classes and methods that address Twitter's models and Programming interface endpoints, and it directly handles different execution nuances, for instance, Information encoding and disentangling

II. LITERATURE REVIEW

In this examination, we study the various papers about Homophily Analysis, and Sentiment Analysis about Indian policy driven issues. For Sentiment Analysis there are numerous techniques for AI like Naive bayes, Rule based grouping, C.5 strategy, Support Vector Machine (SVM), Dictionary based strategy, and so on.

In this paper they gather 4.9 billion clients' information for political Analysis. They gather information utilizing the authority Twitter API (Application Programming Interface). In America there are two contender for the political race. They first use rule based order for Sentiment Analysis. In rule based arrangement they partition the six classes for Sentiment examinations are Political and Non-Political. In Political there are five classes Trump ally, Hillary ally, Positive, Neutral, and Negative and In Non-political only one class is whatever. They have done Sentiment Analysis with the utilization of Homophily Analysis. They have done Homophily Analysis by their Follow association, Retweet association and Mention association. They likewise examine unidirectional and corresponding association between clients. In this paper they get the Homophily result pretty much every one of the six classes of follow, retweet, and notice [6].

In this paper they had gather the information from Twitter Archiver Tool. They gather tweets in Hindi language as it were. They separate political and nonpolitical tweets and perform Sentiment Analysis of political tweets. They Performed Sentiment Analysis utilizing Naive bayes calculation, Support Vector Machine (SVM), and Dictionary based calculation. They analyze this strategies and get the precision of Sentiment Analysis. They ascertain the tweets of all the party that are positive, unbiased or negative. They demonstrated that the precision of credulous bayes is 62.1%, the exactness of SVM (Support Vector Machine) is 78.4% and the precision of Dictionary based calculation is 34% [13].

In this paper, experts propose the construction that makes sense of the development of the collection, Sentiment Analysis, and request Twitter sentiments. They predict the consequence of a political race result by using the twitter data. They prompted predict political race in US, UK, Spain, French and Indonesia itself. Taking into account these data, the makers proposed one more strategy to expect the political race result that focuses on tweet considering and Sentiment

Analysis the preprocessing task. They focused particularly on tweet counting, Sentiment Analysis and pre getting ready undertaking. They give one more technique which is straightforward stood out from various systems [20].

In this paper the principle targets are the aftereffect of Sentiment mining in each region, to look at the expectation consequences of forecasts of guileless bayes and C.5 techniques, and to break down the connection between Sentiments in every area (relationship Analysis). They utilized twitter creeping process (TCP) for gathering the information. They utilized vocabulary approach for Sentiment mining. They demonstrated that C.5 strategy is superior to guileless bayes technique regarding exactness, accuracy and review. C.5 technique has improved outcome [1].

In this paper, they used Merging of frameworks like data mining with various procedures (content mining, NLP and computational knowledge). Creators are mixing Support Vector Machine (SVM) with Decision Tree. In this, show that crossbreed model superior the overall plan assessed in precision and f-measure. Opinion figure is better when appeared differently in relation to standard systems for gathering. They used STACKNET HYBRID Technique for creation of the cross variety computation. SVM (support vector machine) framework collects the request model that gives out happy advisers for predefined class groupings. This technique making a non-probabilistic direct classifier. They utilized TF-IDF (term recurrence - converse report recurrence) technique for convert string input into numeric structure. They performed first adaboosted choice tree calculation and the precision is 67%, second SVM calculation and its exactness is 82%, and third they perform crossover calculation which is the mix of choice tree calculation and SVM calculation and its precision is 84%. Half breed calculation have the higher exactness contrasted with both [7].

In this exploration, they utilize three characterization strategies Naive Bayes, Support Vector Machine, and calculated relapse. They utilized Tweepy (twitter API)

which is utilized in python to give admittance to the twitter information. They demonstrate gullible bayes is better among three methods. They involved gullible bayes calculation with Sentimental score computation in the proper proportion for better exactness of the model. The exactness of guileless bayes calculation is 77% and the precision of feeling score count is 95%. They decided the accuracy of joined both the computation is 90%. They used Sentiment Analysis dataset from Kaggle as data source. The limitation of this assessment is that the model is language express [17].

In this examination, they use Sentiment Analysis utilizing Multinomial Naive Bayes and Decision tree calculations. They utilized Apache Spark bunch (Apache Spark system) for quick handling. In this paper the outcomes show that Decision tree performs incredibly well appearance exactness, accuracy, review and F1-Score [11].

In this paper, the precision of innocent bayes calculation was not exactly the exactness of help vector machine. They made last assumption utilizing SVM, since the precision of calculation is higher [21].

In this paper, the aftereffects of order expectation in the feeling class show that C5.0 strategy is a preferred technique over the Naive Bayes strategy. This is shown by the worth of exactness, accuracy, and review strategy C5.0 which is higher than the Naive Bayes technique [22].

In this paper, Sentiment Analysis turn out to be good for separating a large number scrutinizes going with a solitary remark [23].

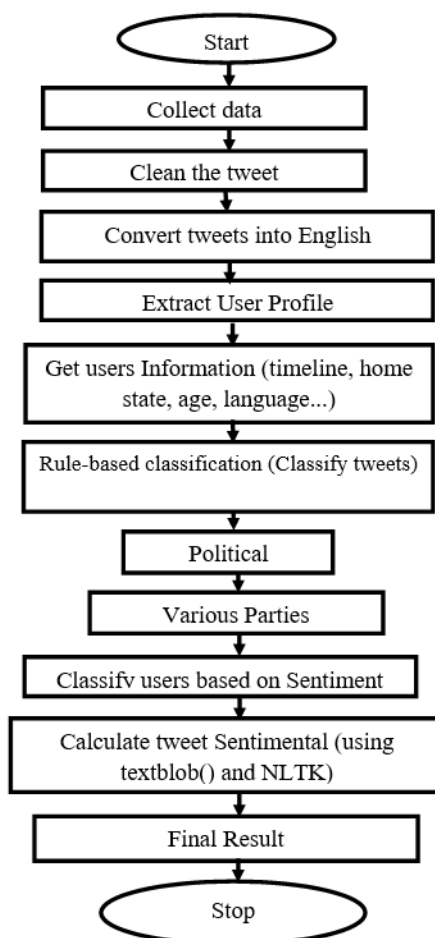
In this paper, the system can stall shared online life substance uninhibitedly used, in light of everything, dynamically to find events related to disasters or setbacks, and inspect and predict insightful ways, the structure can be utilized for fiasco notice organization for seismic quakes and waves or persistent auto collision instructing organization to rapidly recognize catastrophes and incident thusly diminishing damage [24].

PROPOSED METHOD

Feeling Analysis is helpful in each kind of business, web-based entertainment or in private use. Opinion Analysis for Indian policy centered issues by utilizing the twitter information is the troublesome errand since there are many gatherings. .

We likewise group parties by utilizing rule based characterization method. In India there is no main English tweets established so we need to change over the tweets also. The research proposes to the Analysis of Indian ideological groups. In India there are many gatherings in political decision so we partition them in two section one in current decision party and second is resistance groups. In India, there are numerous dialects tweets are there. So we convert all tweets into English.

Proposed Flow of System



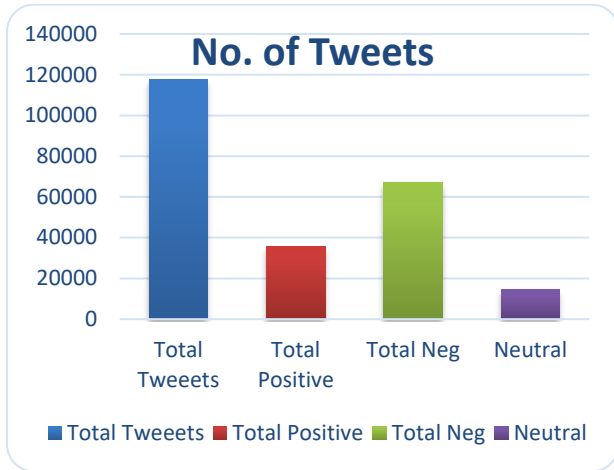
Proposed Algorithm

- Step 1:
Read the Tweets.
- Step 2:
Clean the Tweets.
- Step 3:
Convert Tweet to English.
- Step 4:
Extract User Profile.
- Step 5:
Get Gender, Home State etc Information.
- Step 6:
Get User TimeLine.
- Step 7:
Classify Tweets into political, non-political
- Step 8:
Calculate Tweet Sentimental (using textblob() and NLTK).
- Step 9:
Classify Users based on Sentimental score
- Step 10:
Final Result.

The proposed strategy gives the improved outcome and precision contrasted with existing technique. I will perform Homophily Analysis with extra elements. In existing strategy they were perform Sentiment Analysis utilizing Homophily Analysis with the utilization of follow association, re tweet association, and hash labels.

IMPLEMENTATION

First, we have done Sentiment Analysis on collected Twitter political Data. We get the Positive, negative and neutral tweets.



Total no. of Tweets with Sentiment Analysis Result

Conditions for Sentiment Analysis:

Score = no of positive - no of negative

In the event that Score > 0, the sentence has a by and large 'positive assessment'

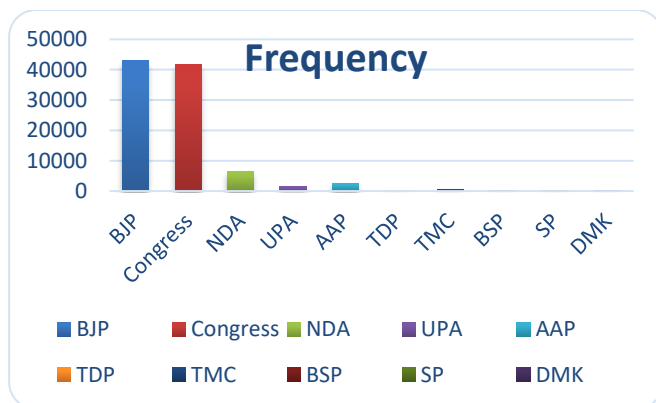
In the event that Score < 0, the sentence has a by and large 'negative assessment'

In the event that Score = 0, the sentence is viewed as a 'impartial assessment'

Total no. of Tweets, Positive, Negative, Neutral

Type of Tweets	No. of Tweets
Total Tweets	117545
Total Positives	35830
Total Negatives	67240
Neutral	14475

This is the frequency of all parties during 2019 Indian election.

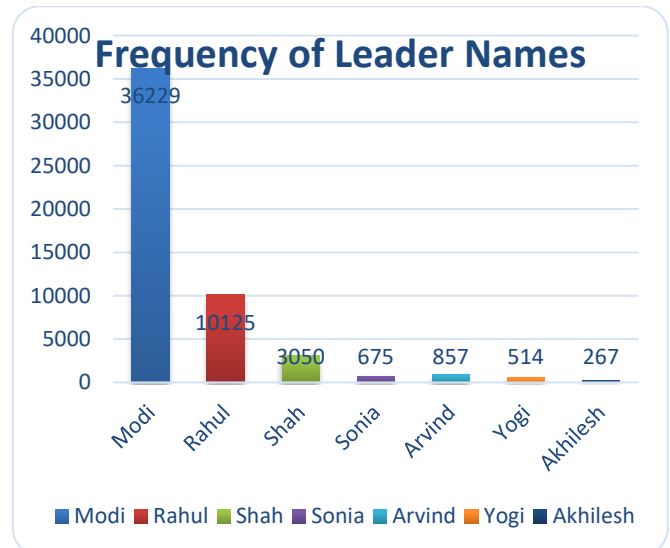


Frequency of all Parties

Frequency of all Parties

Keyword	Frequency
BJP	42957
Congress	41673
NDA	6492
UPA	1422
AAP	2617
TDP	221
TMC	617
BSP	264
SP	191
DMK	234

This is the frequency of Leaders in 2019 Indian election.

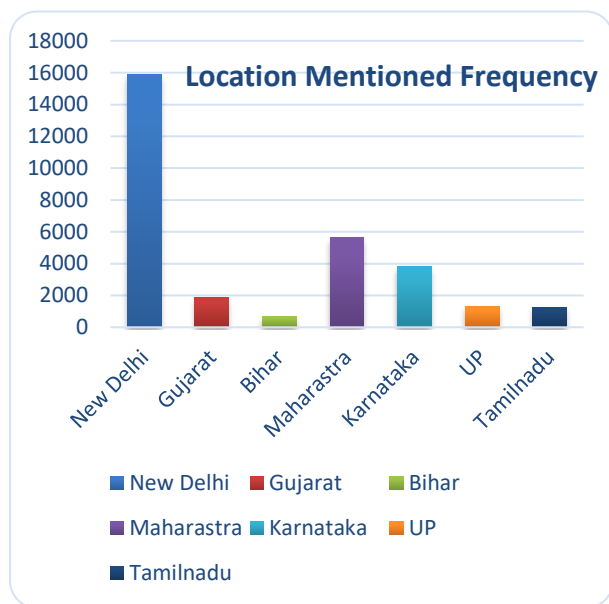


Frequency of Leader Names

Frequency of Leaders

Leader	Frequency
Modi	36228
Rahul	10125
Shah	3050
Sonia	675
Arvind	857
Yogi	514
Akhilesh	267

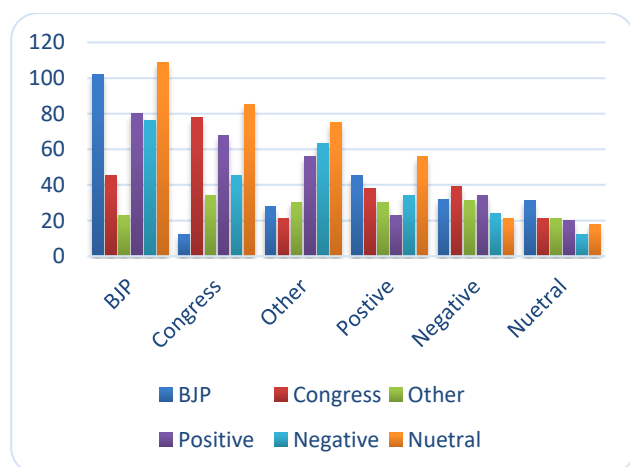
Location based Frequency of Indian Twitter Political Users twitter users.



Location Mentioned Frequency

Location	Frequency
New Delhi	15899
Gujarat	1887
Bihar	677
Maharashtra	5617
Karnataka	3812
UP	1298
Tamilnadu	1272

The result of Followers and Following Parties of Indian election



Followers | Following no. of users in Classes

CONCLUSION

Tremendous proportion of data are available on twitter for analyze. Opinion Analysis gives the Sentiment of clients about any point or item or character. In now daily's political decision is a moving, hot and intrigued point for online entertainment clients. There are numerous procedures performed for tweeter Sentiment Analysis like Naive bayes, SVM (support vector machine), choice tree, C.5, dictionary based approach thus many AI based methods. We done Sentiment Analysis and Homophily Analysis on Twitter information.

- We Performed the Sentiment Analysis using rule based classification technique.
- We have Learn Python for opinion mining.

REFERENCES

- [1] E. S. Saputro, K. A. Notodiputro, and I. A, "Study of Sentiment of Governor's Election Opinion in 2018," Int. J. Sci. Res. Sci. Eng. Technol., vol. 4, no. 11, pp. 231–238, 2018.
- [2] Patil, "Restaurant 's Feedback Analysis System using Sentimental Analysis and Data Mining Techniques," 2018 Int. Conf. Curr. Trends Toward Converging Technol., pp. 1–4, 2018.
- [3] F. A. Pozzi, E. Fersini, E. Messina, and B. Liu, Challenges of Sentiment Analysis in Social Networks: An Overview, vol. 1. Elsevier Inc., 2017.
- [4] F. Poecze, C. Ebster, and C. Strauss, "Social media metrics and Sentiment Analysis to evaluate the effectiveness of social media posts," Procedia Comput. Sci., vol. 130, pp. 660–666, 2018.
- [5] H. P. Patil and M. Atique, "Sentiment Analysis for social media: A survey," 2015 IEEE 2nd Int. Conf. Information Science Secur. ICISS 2015, 2016.
- [6] J. A. Caetano, H. S. Lima, M. F. Santos, and H. T. Marques-Neto, "Using Sentiment Analysis to define twitter political users' classes and their

- Homophily during the 2016 American presidential election,” *J. Internet Serv. Appl.*, vol. 9, no. 1, 2018.
- [7] K. Ganagavalli, A. Mangayarkarasi, T. Nandhinisri, and E. Nandhini, “Sentiment Analysis of twitter data using machine learning algorithm,” *J. Comput. Theor. Nanosci.*, vol. 15, no. 5, pp. 1644–1648, 2018.
- [8] M. Kamyab, R. Tao, M. H. Mohammadi, and A. Rasool, “Sentiment Analysis on Twitter,” vol. 9, no. 4, pp. 14–19, 2018
- [9] M. Miller, S.-L. Lynn, and M. C. James, “Birds of a Feather: Homophily in Social Networks,” *Annu. Rev. Sociol.*, vol. 27, pp. 415–444, 2001.
- [10] M. S. M. Vohra and P. J. B. Teraiya, “Journal of Information, Knowledge and Research in Computer Engineering a Comparative Study of Sentiment Analysis Techniques,” *J. Information, Knowledge Res. Comput. Eng.*, pp. 313–317, 2013.
- [11] P. Deacon, “Application of Machine Learning Techniques to Mineral Recognition,” *Computer (Long. Beach. Calif.)*, no. October, pp. 628–632, 2001.
- [12] P. Seth, A. Sharma, and R. Vidhya, “Sentiment Analysis of Tweets Using Hadoop,” *Int. J. Eng. Technol.*, vol. 7, no. 3.12, p. 434, 2018.
- [13] P. Sharma and T. S. Moh, “Prediction of Indian election using Sentiment Analysis on Hindi Twitter,” *Proc. - 2016 IEEE Int. Conf. Big Data, Big Data 2016*, pp. 1966–1971, 2016.
- [14] R. Singh and R. Kaur, “Sentiment Analysis on Social Media and Online Review,” *Int. J. Comput. Appl.*, vol. 121, no. 20, pp. 44–48, 2015.
- [15] S. A. El Rahman, F. A. Alotaibi, and W. A. Alshehri, “Sentiment Analysis of Twitter Data,” *2019 Int. Conf. Comput. Inf. Sci. ICCIS 2019*, 2019.
- [16] S. Geetha and V. K. Kaliappan, “Tweet Analysis Based on Distinct Opinion of Social Media Users,” *ICSNS 2018 - Proc. IEEE Int. Conf. Soft-Computing Netw. Secur.*, pp. 1–6, 2018.
- [17] S. Goel, M. Banthia, and A. Sinha, “Modeling Recommendation System for Real Time Analysis of Social Media Dynamics,” *2018 11th Int. Conf. Contemp. Comput. IC3 2018*, pp. 1–5, 2018.
- [18] S. Y. Yoo, J. I. Song, and O. R. Jeong, “Social media contents based Sentiment Analysis and prediction system,” *Expert Syst. Appl.*, vol. 105, pp. 102–111, 2018.
- [19] V. A. and S. S. Sonawane, “Sentiment Analysis of Twitter Data: A Survey of Techniques,” *Int. J. Comput. Appl.*, vol. 139, no. 11, pp. 5–15, 2016.
- [20] W. Budiharto and M. Meiliana, “Prediction and Analysis of Indonesia Presidential election from Twitter using Sentiment Analysis,” *J. Big Data*, vol. 5, no. 1, pp. 1–10, 2018.
- [21] Neethu M, S, Rajasree R, ‘Sentiment Analysis in Twitter using Machine Learning Techniques’, *4th ICCCNT*, 2013.
- [22] U. Kursuncu, M. Gaur, U. Lokala, K. Thirunarayan, A. Sheth, and I. B. Arpinar, “Predictive Analysis on Twitter: Techniques and Applications,” pp. 67–104, 2019.
- [23] W. Budiharto and M. Meiliana, “Prediction and analysis of Indonesia Presidential election from Twitter using Sentiment Analysis,” *J. Big Data*, vol. 5, no. 1, pp. 1–10, 2018.
- [24] Y. Mejova, ‘Sentiment Analysis: An overview’, *ymejova/publications/CompsYelenaMejova*, vol. 2010–02–03, 2009, 2009.
- [25] <https://monkeylearn.com/sentiment-analysis>