

# Evaluating Sentiment Analysis Techniques for Political Tweet Data

Kirti M. Jain<sup>1</sup>, Dr. Avinash J. Agrawal<sup>2</sup>

<sup>1</sup>M. Tech Scholar, Shri Ramdeobaba College of Engineering & Management, Nagpur, India,  
jainkm\_1@rknc.edu

<sup>2</sup>Associate Professor, Shri Ramdeobaba College of Engineering & Management, Nagpur, India,  
agrawalaj@rknc.edu

## Article Info

Volume 83

Page Number: 7348 - 7351

Publication Issue:

May - June 2020

## Abstract:

Twitter and other micro-blogging websites have become an epitome for analysis for user sentiments. In order to perform effective sentiment analysis, researchers need to analyze the data published by users, and then apply different language processing models on the same. A large number of tweets these days are based on political scenarios, wherein twitterati engage themselves in debates over the political decisions, and their impact on daily user's life. Various methods have been proposed in recent years, and each of them has their distinctive advantages and drawbacks. In this text, we reviewed the state-of-the art methods for sentiment analysis, and tried to keep the work for the political domain. This is done because many cases have come up in recent times, where political campaigns are run by checking out the general user's sentiments. For example, if a particular party is not-strong in a particular area, then the sentiment analyzer can help in identifying areas of improvement based on the tweets/social media data analysis for that party. Thereby, increasing the chances of winning the elections in that particular area. In this paper, we also recommend some improvements in the existing work, via which researchers can design better sentiment analysis systems.

## Article History

Article Received: 19 November 2019

Revised: 27 January 2020

Accepted: 24 February 2020

Publication: 18 May 2020

**Keywords:** Politics, tweets, social, media, sentiment

## 1. INTRODUCTION

In today's advanced period, individuals know each other significantly by means of web-based life refreshes. Gone are the days, when individuals used to telephonically examine about their own existence with loved ones. These days, individuals go to places just to show the world how pleasant their life is. Because of this structural move from individual correspondence to economic wellbeing-based perception, individual correspondence has endured a big cheese. Individuals purchase things to dazzle others whom they don't have the foggiest idea, and discover comfort in doing as such. A few people are simply eyewitnesses, who simply scroll internet-based life pages, while the other extraordinary clients ensure that their moment-to-minute updates are posted, and that their companions/devotees read and like their posts. The observational class clients are influenced by social frailty, which offers ascend to social discouragement,

and coordinates their purchasing and excursion designs. The moderate class clients, who post refreshes less as often as possible attempt to keep a low social profile, yet get discouraged when different clients are replicating their examples. While the dynamic social clients continue posting refreshes, and get discouraged when their posts don't get enough consideration via web-based networking media.

Aside from these web-based social networking-based miseries, individuals around us are commonly discouraged by things which have nothing to do with their own or expert life. Besides, because of individual and expert issues individuals get will in general get discouraged. Also, in this discouraged state of mind, when they have nobody to share their request, they go to online life for consideration and solace. In this discouraged state of mind, individuals will in general post refreshes which feature their disposition, and the moves they make while being discouraged. As Facebook and WhatsApp refreshes are turning

out to be increasingly more business nowadays, individuals are going to smaller scale blogging stages like Twitter and Snapchat for displaying their feelings. The following section manages the survey of various sadness examination calculations, and their subtleties. In the survey we look at different ongoing systems for examination and observe their performance. Besides, we have additionally given proposals with respect to which methods are better, and which ones can be improved dependent on their usage. At long last, we finish up this content with some intriguing perceptions about the reviewed algorithms, and recommend techniques to improve the same.

## 2.LITERATURE REVIEW

Already different specialists have moved in the direction of improving the presentation of wretchedness proposal. One of these works is referenced in [1], wherein the creators have utilized both discourse and literary information so as to perform estimation investigation via different review analysis techniques. They have studied help vector machines (SVM) so as to play out the given undertaking. They use WordNet Affect and SentiWordNet for language preparing and pitch, vitality, formants, power and zero intersection rate (ZCR) highlights for sound handling. They guarantee to have accomplished 81% exactness for feeling investigation, which can be additionally improved utilizing profound learning calculations like profound nets and Q-learning. Their work can be utilized for dependency examination by annexing their outcomes to our fluffy standard base. Another fascinating work with regards to this field is done in [2], wherein analysts have utilized election audits so as to perform notion investigation. They have utilized a multi-modular examination framework and can accomplish an exactness of over 60% for an enormous dataset. Their examination can along these lines be utilized for moderate to enormous datasets so as to foresee assumption esteems.

Indian sentiment examination can be joined with printed includes so as to additionally improve the nature of conclusion investigation. The work done in [3] utilizes Indian users' tweet data and joins it with printed includes so as to improve the exactness of notion investigation. They have utilized convolutional neural systems (CNN) so as to perform order. The CNN further uses a multi-part model so as to perform successful arrangement. The work further uses sound information so as to break down the client's enthusiastic state, and consolidates that data with content and video to precisely distinguish the opinions of the client. Their work can identify feelings like furious, cheerful, miserable and nonpartisan with a high exactness of over 70%. This precision can be additionally improved by utilizing spiking neural systems, wherein spikes from sound, video and content information can be joined so as to additionally improve notion location exactness. An intriguing survey about such unique exactness improvement strategies can be perused from [4], wherein methods like Page

rank, Gradient drop, Linear relapse, Link mining, Collective arrangement, AdaBoost.HM, DBA, Regression, Random walk, SVM, Cohen's K coefficient, Fuzzy grouping, PMI, Association Miner CBA, Markov-Chain Monte Carlo (MCMC), and so on have been looked at. From their survey it is intrinsic that spiral premise work neural system joined with restrictive arbitrary field hypothesis can be a successful blend for arranging any sign source. In this manner, it tends to be utilized for high precision conclusion examination. This can be seen from the work in [5], wherein different profound learning models are compared so as to improve the exactness of supposition investigation. Analysts have researched Deep Boltzmann Machine (DBM) so as to order literary information, sound information and visual information. These 3 investigations are then consolidated and a conclusive outcome is created. The outcomes are contrasted and Lexicon, Visual, SentiBank, AF SentiBank + Lexicon, MDBM, E-MDBM-V and E-MDBM-T. It is seen that the proposed technique can outflank every one of these strategies and a precision of over 75% is acquired. It is suggested that creators of supposition investigation/gloom examination frameworks use DBM strategy so as to structure their frameworks.

Just content-based election results investigation frameworks additionally give great exactness, when utilized in legitimate settings. This is demonstrated by [6], wherein content based measurements are utilized so as to investigate election results. Specialists have utilized Python's Natural API so as to perform conclusion investigation. Utilizing a very good quality instrument doesn't add to any examination, yet accelerates the advancement procedure, and ensures high exactness. Because of the API-based usage, the outcomes are over 80% exact across both continuous and static datasets. We would prescribe to utilize this motor just as an approval motor for any sort of opinion investigation framework. Like the work done in [7], wherein specialists have utilized Distributional Semantic Models for successfully breaking down discouragement from input information. They have utilized bi-grams, tri-grams, and N-grams so as to improve the exactness of supposition investigation. The framework yielded over 80% precision, which is sufficiently high for any continuous arrangement and opinion investigation.

Combination based frameworks are additionally pervasive in supposition examination. For example, the work done in [8], uses a mix of content and discourse for assessment of offline results for voting. They likewise use Naïve Bayes based classifier, and can accomplish over 45% exactness. The work is exceptionally fundamental, yet can be utilized as a beginning point for any examination around there. Another multi-modular framework with significant level of precision is portrayed in [9], wherein highlights like content ensemble approach is consolidated together so as to prepare a convolutional neural system. Highlights like N-grams, mel-recurrence parts, shading maps are assessed for content, sound and video separately. The

outcomes grandstand a precision of over 85% now and again, while the most pessimistic scenario exactness is around 70%. The framework is exceptionally powerful and along these lines can be utilized in blend with DBN so as to additionally improve the most pessimistic scenario execution. A predictive work is done in [10], wherein hybrid topic-based methods are explored for political review analysis. The framework utilizes a subjective and a discernment module so as to accomplish a comparative exactness. Application-explicit frameworks additionally make ready for notion examination. Like the work done in [11] which uses tweets of the citizens so as to identify burdensome qualities. They [11] have utilized the comparative calculations like in [10], and gained from these calculations so as to enhance the precision of the framework. The framework has utilized tweets, with sound and visual information for feeling investigation and is seen as over 90% powerful at times. A point by point study about adoptable frameworks is introduced in [12], wherein strategies like neural systems, SVMs, Naïve Bayes, include extraction procedures, pre-handling methods, and so on are considered. From their exploration unmistakably profound learning and AI models make ready for additional advancement in the territory of opinion investigation.

Deep learning-based investigation is an extremely intricate framework. Comparable work is done [13] wherein multi-modular frameworks and application explicit frameworks are concentrated in subtleties for political, climatic and other kinds of tweet data. Sentiment and social media combination-based methodologies are talked about in [14], wherein multi-modular examination is done based on intra-methodology and between methodology elements. The aftereffects of such a framework take into account continuous applications in conclusion investigation with an exactness of over 65%. Further, the tensor-based frameworks can be improved with the assistance of DBN and CNNs for a superior precision. Comparable investigations are done in [15], wherein techniques like tension examination, profound learning and AI are proposed. From the audit, we can distinguish that profound learning and AI strategies like DBN and CNN are best for sadness examination, yet restricted work has been done on fluffy standard based frameworks. In this manner, this work proposes an AI based fluffy standard based framework. The framework and its subtleties are portrayed in the following section.

### 3.RESULT EXAMINATION AND CONCLUSION

So as to perform result assessment, we took information from various number of clients, and broke down their tweets when they were in various states of mind. In view of this investigation, we assessed the supposition of the clients by means of various calculations, and assessed the exactness and review esteems. These qualities were arrived at the midpoint of, and contrasted and other standard calculations so as to assess the presentation of every calculation.

Parameters like precision and recall were evaluated using the following formulas,

$$Precision = \frac{\text{Number of correct outputs}}{\text{Total number of outputs}}$$

$$Recall = \frac{\text{Number of correct outputs}}{\text{Total number of obtained results}}$$

These are standard formulae which are utilized by us for assessment of the presentation of various techniques. We assessed the mean estimations of the exactness and review parameters, and contrasted it and the current calculations and got the accompanying outcomes,

**Table 4.** Performance comparison with other algorithms

Algorithm Name	Precision	Recall
Multinomial Naïve Bayes	0.77	0.77
Random Forest	0.81	0.8
Gradient Boosting	0.79	0.77
Ensemble Vote	0.85	0.85
Regressive Naïve Bayes	0.836	0.83
Support Vector Machine	0.8	0.79
CNN-based classifiers	0.93	0.93

From the results it is clear that the CNN algorithm performs at least 10% better than the existing state-of-the-art techniques, and thus can be used for real-time social media-based depression analysis.

#### 4.FUTURE WORK

Nearness of AI and AI (ML) in the present work will enhance the general exactness and accuracy esteems. The ML based strategies like Q-learning and re-requirement learning will take into consideration better execution because of the impetus based arrangement component, while profound nets will show signs of improvement exactness and review esteems because of their unrivalled example investigation and characterization execution.

#### REFERENCES

1. MitaliDesai ,Mayuri A. Mehta **“Techniques for Sentiment Analysis of Twitter Data: A Comprehensive Survey”** International Conference on Computing, Communication and Automation (ICCCA2016).
2. JyotiRamteke, Darshan Godhia, Samarth Shah, AadilShaikh**“Election Result Prediction Using Twitter sentimentAnalysis”**2016 International Conference on Inventive Computation Technologies (ICICT)**Date of Conference: 26-27 Aug. 2016**
3. Parul Sharma, Teng-Sheng Moh**“Prediction of Indian Election Using Sentiment Analysis on Hindi Twitter”** 2016 IEEE International Conference on Big Data (Big Data).
4. RasikaWagh , PayalPunde**“Survey on Sentiment Analysis using Twitter Dataset”**Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA 2018)IEEE Conference Record # 42487; IEEE Xplore ISBN:978-1-5386-0965-1 .
5. Ayeena Malik<sup>1</sup>, Divya Kapoor<sup>2</sup>, AmitPrakash Singh<sup>3</sup>**“Sentiment Analysis on Political Tweets”** Vth International Symposium on **“Fusion of Science & Technology”**,New Delhi, India, January 18-22, 2016.
6. Ms. FarhaNausheen, Ms. SayyadaHajera Begum **“Sentiment Analysis to Predict Election Results Using Python”**Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018)IEEE Xplore Compliant - Part Number:CFP18J06-ART, ISBN:978-1-5386-0807-4; DVD Part Number:CFP18J06DVD, ISBN:978-1-5386-0806-7.
7. PriteeSalunkhe, SachinDeshmukh**“Twitter Based Election Prediction and Analysis”**International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 04 Issue: 10 | Oct -2017.
8. Pragyajuneja, Uma Ojha **“Casting Online Votes:To Predict Offline Results Using Sentiment Analysis by machine learning Classifiers”** IEEE – 40222.
9. Rincy Jose, Varghese S Chooralil**“Prediction of Election Result by Enhanced Sentiment Analysis on Twitter Data using Classifier Ensemble Approach”** 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)**Date of Conference: 16-18 March 2016.**
10. TarekElghazaly, Amal Mahmoud, Hesham A. Hefny**“Political Sentiment Analysis Using Twitter Data”**Conference Paper · March 2016DOI: 10.1145/2896387.2896396
11. Charles Chang, Michael Masterson **“Using Word Order in Political Text Classification with Long Short-term Memory Models”**University of New England, on 02 Jan 2020 at 16:05:59, subject to the Cambridge Core terms of use, available at <https://www.cambridge.org/core/terms>.
12. Barkha Bansala, SangeetSrivastavaa, **“On predicting elections with hybrid topic based sentiment analysis of tweets ”** 3rd International Conference on Computer Science and Computational Intelligence 2018
13. Ashima Yadav<sup>1</sup> · Dinesh Kumar Vishwakarma<sup>1</sup>**“Sentiment analysis using deep learning architectures:a review”**© Springer Nature B.V. 2019
14. WidodoBudiharto, MeilianaMeiliana**“Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis”**Budiharto and MeilianaJ Big Data (2018) 5:51 <https://doi.org/10.1186/s40537-018-0164-1>