

Mining Public Opinion on Indian Government Policies using R.



Pankaj Verma, Sanjay Jamwal

Abstract: Nowadays, social media monitoring has burgeoned at a very rapid pace, so analyzing social data plays a crucial role in knowing people's behavior. People share their views regarding trending topics that are occurring around the world. Opinion mining is used for extracting the sentiments from the textual data that are shared by peoples. In this work, we are analyzing Twitter Tweets using sentiment analysis which checks the opinion of people regarding various policies that were announced by the Indian Government. The main objective of the paper is to analyze sentiments of various Indian Government policies (namely Article370, New Motor Vehicles Act 2019, Triple Talaq, Jal Shakti Abhiyan, NRC Assam) on Twitter so that public opinions and views are analyzed. Emotions (anger, trust, fear, anticipation, disgust, sadness, joy, surprise) are analyzed using Emotion-based lexicon technique. Sentiments are classified into two categories (positive and negative) from the emotion lexicon EmoLex. The paper provides a comparative analysis of these policies and this work can act as feedback from people to the government. R programming is used for implementation and visualization.

Keywords: Sentiment Analysis, Article370, Triple Talaq, Jal Shakti Abhiyan, NRC, etc.

I. INTRODUCTION

Twitter is one of the best platforms for sharing information and expressing one's views regarding various events that are occurring in the world [1]. Twitter is the ninth largest social network platform in the world, and it has 328 million active users per month, and Lakhs of tweets tweeted per day. In this way, information and reaction convey to the top level to the bottom level of the people at very fast. Recent year's people posted more tweets and retweets about the new products launched the government policy executions and international talks. Twitter is the best platform for the opinion retrieval system in social networks. Various methods are used for extracting and analyzing the Opinion of People regarding these events, Sentiment Analysis is one of the approaches that is used for analysing the opinion of people. With the help of Sentiment analysis using twitter data, we can quickly identify the positives and negatives opinions of the tweets

[2]. In our work, we analyzed various Indian Government projects, schemes and laws by extracting the data from social networks.

Revised Manuscript Received on January 30, 2020.

* Correspondence Author

Pankaj Verma*, Department of Computer Sciences, BGSB University, Rajouri, J&K, India. E-mail: pankajvermajmu@gmail.com

Sanjay Jamwal, Department of Computer Sciences, BGSB University, Rajouri, J&K, India. E-mail: sanjayjamwal@bgsbu.ac.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](#) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Our work can be beneficial for evaluation of government performance monitoring from People's perspective instead of making Peoples survey manually which is very expensive and time-consuming. In this paper, a system is proposed for analysing the opinion of people on various government Schemes and laws, we analysed the following laws and schemes Article370, New Motor Vehicles Act 2019, Triple Talaq, Jal Shakti Abhiyan, and NRC Assam using twitter data and implementing it by R programming language [3, 4].

A. Article370

Article 370 is in Part XXI of the constitution of India, which gave a special status to Jammu and Kashmir, one of the states of India that is located in the Northern part of the Indian subcontinent. Jammu and Kashmir is the only Indian state which has its own separate state constitution and owns state flag. Article 370 came into force on 17 November 1952 which has been the subject of dispute between Indian, Pakistan, and China since 1947[5]. According to this article the parliament of India needs the Jammu and Kashmir government's approval for applying laws in the state except laws in case of foreign affairs, finance, communications, and defense. On 5 August 2019 Government of India presented the bill in the Parliament to abrogate the clauses of Article 370 and 35 A that gave them special status. They proposed to reorganize Ladakh and Jammu and Kashmir into two separate union territories, instead of full-fledged state. The bill was passed in both houses of parliament with full majority. Social media reacted enormously about this step that was taken by the Indian Government. There is lot of hue and cry with mixed confused opinions and views from different political parties and peoples around the world. They used various social networking platforms like Twitter, Facebook, and News portal for expressing their views. We extracted the data from twitter using the keyword #Article370, we analyzed the tweets using R programming and the result of our analysis is shown in result section.

B. New Motor Vehicle Act 2019

Due to the increase in accidents and traffic breakups across the country, the New Motor Vehicle Act was passed by the parliament of India on 1st September 2019.

According to this act there was a huge hike in penalties for violating the traffic rules. By introducing this act there was a huge dent in monthly budget of common person and some of them may have to pay his/her full monthly salary. After implementation of this act there are lot of reaction, opinions, views, and mimics which were posted by peoples on various social networking platforms like Twitter, Facebook, and WhatsApp, etc.



We extracted the twitter data regarding this act such that we can be able to analyze the people's reaction or views about this step of Indian government, #MotorVehicleAct keyword was used for extracting the data and R language was used for implementation. The Analysis is shown in result section.

C. Triple Talaq

Talaq is an Arabic word and is a Synonym of Divorce. Various Divorce methods are used by different religions. Triple talaq, a form of Islamic divorce used by Muslims communities especially by Hanafi Sunni Islamic schools of jurisprudence. Triple talaq is also known as talaq-e-biddat, instant divorce [6]. It allowed Muslim married men to legally divorce his wife by repeatedly uttering the word talaq three times in oral, written or, more recently, electronic form. On 22 August 2017, the Supreme Court of India deemed instant triple talaq unconstitutional [7, 8, 9] effected from 19th September 2018. On 30th July 2019 parliament of India declared that the practice of Triple Talaq as illegal and from 1st August 2019 who deny this will face severe punishment. There were various opinions, views, and comments shared by different communities, sects of people and various political parties of the world after this Act was passed. They used various social networking sites for expressing their views, opinions, etc. The result section gives the opinion of people regarding this act.

D. Jal Shakti Abhiyan

Water, a lifeline for bio kingdom. Earth consists of about 70% of water, only 0.4% is used for drinking purpose. The climate change has affected almost every country across the world, India one of the affected nations has a severe scarcity of drinking water, the government of India initiated "Jal Shakti Abhiyan" (JSA) a time-bound mission campaign run by Indian prime Minister for Conservation of drinking water. Jal is a Hindi synonym of drinking water. The Jal Shakti Abhiyan consists of two phases: the first phase started from 1st July to 15th September 2019 and the second phase will be kicked from 1st October to 30th November for all states and Union territories across India. We analyzed the social network reaction about the first phase of this mission, the analysis was performed using R language on Twitter dataset, and result section gives a detailed analysis of the ongoing Mission.

E. NRC Assam (National Register of Citizens)

For identification of Indian citizens, a register is maintained by the government of India that contains the relevant information about citizens of Assam and this register is known as National Register of Citizen (NRC). Recently it gained attention on Social networks because the Government of Assam published the list of Citizens that were registered in NRC. There were many great personalities that were excluded from the list. People use Social Networks like Twitter on which they react to this step of government. We collected the data using the keyword #NRCASSAM, we analyzed the data such that we can find the reaction of people about NRC. In our work, we employ an open-source approach for sentiment analysis and text mining. We used R-studio, a free and open-source IDE for developing and deploying R applications that can be installed on top of the Linux/Macintosh/Windows system. R language is a scripting

language for conducting statistical data computing and big data analytics and it has more than 10000 packages [10].

II. SENTIMENT ANALYSIS

The term Sentiment Analysis is used for calculating the sentiments of people regarding a particular context like product reviews, Laws, Acts, projects, policies, and other sponsored schemes [11, 12, 16]. Sentiment analysis is one of the approach of data mining [17]. In our work, we analyzed the opinion of people regarding different Indian government laws, acts and missions, etc. Different emotions can be extracted from the textual data using sentiment packages like anger, disgust, fear, joy, sadness, anticipation, surprise and trust. In this work, we use NRC emotion-based dictionary from the R package titled "Syuzhet".

III. METHODOLOGY

For Analysing the above-mentioned Acts, Laws, and Mission the various steps are used like for data extraction, data pre-processing, lexical analysis and classification which are discussed below from A to E.

A. Steps to extract the tweets

1. Create a twitter application.
2. twitteR package for interface to the twitter API.
3. ROAuth package for authentication.
4. Generation of Twitter authentication credential.
5. 7-digit OTP is generated [13].

B. Data Extraction and Collection of People's tweets

Tweets area collection of characters, on which Sentiment Analysis can be performed. For the Collection of tweets from twitter API, we use twitter API credentials such as consumer key, consumer secret key, access token, access token secret and then from these keys we connect to twitter API. In this work, we extracted tweets related to various Government Acts, Laws, and Missions, etc. searchTwitter() function is used for extracting tweets related to the above-mentioned events and dataset is stored in .csv files.

Example: searchTwitter("#Article370",n=10000,lang="en").

C. Data Pre-processing and Data Cleaning

The extracted data is inconsistent, noisy and filthy. The raw data must be pre-processed such that the resultant data set contains only vital information for analysis. The "tm" package of R library is used for text mining. Various functions can be used for removing unwanted strings from tweets [14]. Some of the functions used for performing this task are mentioned below:

- removePunctuation() eliminates punctuation marks from tweets.
- removeNumbers() eliminates numbers as they don't have any underlying sentiment.
- tolower() converts the entire corpus content into lower case.
- Stopwords removes English words without sentiments like articles, conjunction, etc.
- removewordseliminates some specific words.
- StemDocument reduces a word to its original root word.
-

For example, the word “walking” will be changed to its root form “walk”.

- stripWhitespace eliminates extra white spaces.

D. Sentiment Classification

“Syuzhet” package from the R library is used for finding sentiment score and it classifies pre-processed tweets on the basis of polarity (positive or negative) and categorized tweets into eight different types of emotions (fear, joy, anticipation, anger, disgust, sadness, surprise, trust). “Syuzhet” used NRC emotion-based dictionary for classification of tweets. The polarity can be positive or negative depending on the reaction of people regarding the respective events. The polarity is positive if people are satisfied by the step taken by the government and Vice versa. The analysis performed in this paper can be pretended as feedback from the people regarding these events. The feedback can be fruitful for future events.

E. Graphical Representation

The sentiment analysis of the extracted dataset can be represented graphically. There are various graphical packages in RStudio for the representation of data graphically. In this paper bar charts are used to represent the sentiments of people regarding the various Indian government events with the help of ggplot2, plotly, and RColorbrewer packages.

IV. EXPERIMENTS AND RESULTS

System Requirements: Windows 10 operating system, Intel i3, 1.90 GHz core processor with 4 GB RAM, R studio version 1.1.463 and R version 3.5.3.

Dataset: live tweets from twitter using TwitteR package from the R library and ROAuth, Rcurl packages for authentication and to access tweets [15].

Table No: 1 Twitter Dataset used for Experimental work

Name of the Schemes	No. of tweets	Size
Article 370	14927	2391KB
New Motor Vehicle Act 2019	10614	3307KB
Triple Talaq	1731	540KB
Jal Shakti Abhiyan	5906	1828KB
NRC Assam	10507	680KB

A. Article 370

Figure 1. gives detailed emotions on the abrogation of article 370. We used #Article370 for extraction of data from twitter, we collected around 15K tweets and the result shows that most of sentiments were positive regarding it.

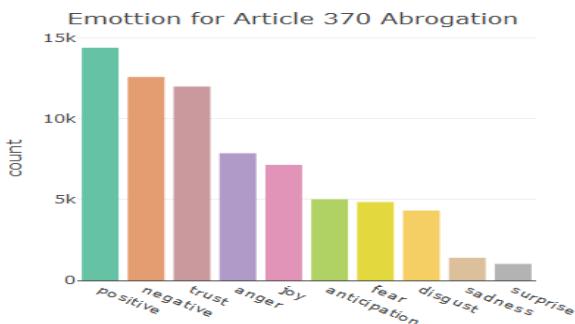


Figure 1: Emotions on abrogation of Article370

Sentiment Scores for each emotion on Abrogation of Article 370 is given in table 2.

Table 2. Sentiment score of each emotion on the abrogation of Article370.

Emotion-Based	Count	Percentage
Anger	7535	10.78%
Anticipation	6544	9.36%
Disgust	2815	4.02%
Fear	5632	8.06%
Joy	5220	7.47%
Sadness	3424	4.90%
Surprise	1450	2.07%
Trust	10309	14.75%
Negative	12176	17.42%
Positive	14755	21.12%

B. New Motor Vehicle Act

Figure 2. Gives detailed emotions on the New Motor Vehicle act. We used #MotorVehicleAct2019 for extraction of data from twitter, we collected around 10.5K tweets and the result shows that most of sentiments were positive regarding the act.

sentiment scores for New Motor Vehicle Act :

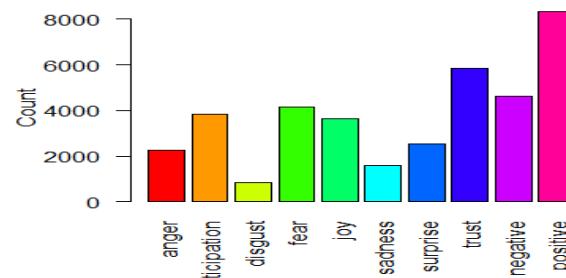


Figure 2. Emotions of the New Motor Vehicle Act.

Sentiment Scores for each emotion on New Motor Vehicle Act is Shown in table 3.

Table3. Sentiment score of each emotion for the New Motor Vehicle Act 2019.

Emotion-Based	Count	Percentage
Anger	2282	6.04%
Anticipation	3855	10.21%
Disgust	864	2.29%
Fear	4157	11.01%
Joy	3644	9.65%
Sadness	1582	4.19%
Surprise	2534	6.71%
Trust	5864	15.54%
Negative	4628	12.26%
Positive	8317	22.04%



C. Triple Talaq

Figure 3. Gives detailed emotions on Triple Talaq. We used #TripleTalaq for extraction of data from twitter, we collected around 1.8K tweets and the result shows that most of the sentiments were positive regarding it.

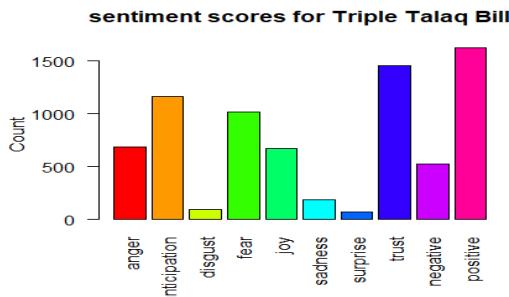


Figure 3: Emotions on Triple Talaq.

Sentiment Scores for each emotion on Triple Talaq Bill is given in table 4.

Table 4. Sentiment score of each emotion on Triple Talaq Bill.

Emotion-Based	Count	Percentage
Anger	685	9.13%
Anticipation	1161	15.48%
Disgust	98	1.30%
Fear	1022	13.62%
Joy	670	8.93%
Sadness	187	2.49%
Surprise	71	0.94%
Trust	1458	19.44%
Negative	524	6.98%
Positive	1624	21.65%

D. Jal Shakti Abhiyan

Figure 4. Gives detailed emotions on Jal Shakti Abhiyan. We used #JalShaktiAbhiyan for extraction of data from twitter, we collected around 6K tweets and the result shows that most of the sentiments were positive regarding it.

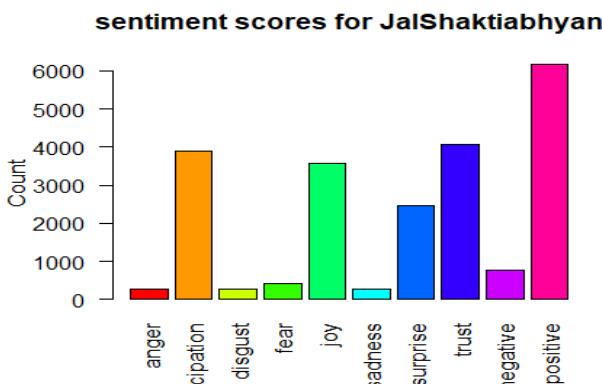


Figure 4: Emotions on Jal Shakti Abhiyan

Sentiment Scores for each emotion on Jal Shakti Abhiyan is Shown in table 5.

Table 5. Sentiment score of each emotion on Jal Shakti Abhiyan.

Emotion-Based	Count	Percentage
Anger	287	1.29%
Anticipation	3884	17.49%
Disgust	286	1.28%
Fear	413	1.86%
Joy	3583	16.13%
Sadness	287	1.29%
Surprise	2461	11.08%
Trust	4077	18.36%
Negative	762	3.43%
Positive	6162	27.75%

E. National Register of Citizens (NRC)

Figure 5. Gives detailed emotions on NRC. We used #NRCAssam for extraction of data from twitter, we collected around 10.5K tweets and the result shows that most of the sentiments were positive regarding it.

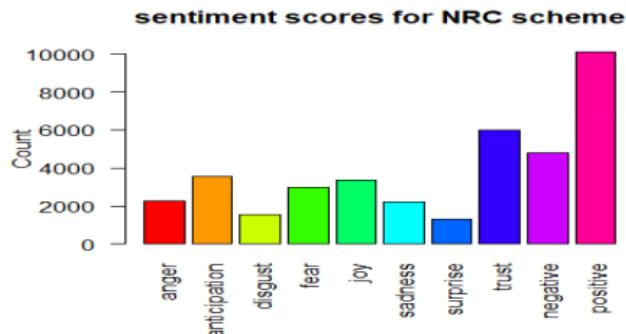


Figure 5: Emotions of National Register of Citizens.

Sentiment Scores for each emotion on NRC Assam is given in table 6.

Table 6. Sentiment score of each emotion on the National Register of Citizens of Assam.

Emotion-Based	Count	Percentage
Anger	2252	5.91%
Anticipation	3538	9.28%
Disgust	1552	4.07%
Fear	2990	7.84%
Joy	3350	8.79%
Sadness	2202	5.77%
Surprise	1328	3.48%
Trust	6010	15.77%
Negative	4790	12.57%
Positive	10092	26.48%

V. COMPARATIVE ANALYSIS

In this section, we analyze and compare different Indian government policies on the basis of emotions by extracting data from the social network platform namely twitter as given in Table 7. We calculate the number of counts for each emotion regarding different Indian government policies in-depth. The percentage of different emotions of community is given in detail in Figure 7. By comparing different emotions, we find the maximum and minimum count percentage of each emotion of various Indian government policies. From Table 8, it was observed that most of the people have fear and anger emotion on Article 370, New Motor Vehicle Act and NRC Assam. The negative count was highest for Article 370 abrogation which is equal to 17.42%. The table also shows that there are anger emotions for Article 370 and NRC Assam which is equal to 10.78% and 10.21% respectively. Other policies like Jal Shakti Abhiyan, Triple Talaq showed positive sentiments. It was observed that most of the peoples appreciate the Indian government policies.

Table 7. Sentiment score of emotions on different Indian Government schemes.

Scheme/Law/Act	Article370	Vehicle	New Motor	Triple Talaq	Jal Shakti	NRC
Anger	7535	2282	685	287	2252	
Anticipation	6544	3855	1161	3884	3538	
Disgust	2815	864	98	286	1552	
Fear	5632	4157	1022	413	2990	
Joy	5220	3644	670	3583	3350	
Sadness	3424	1582	187	287	2202	
Surprise	1450	2534	71	2461	1328	
Trust	10309	5864	1458	4077	6010	
Negative	12176	4628	524	762	4790	
Positive	14755	8317	1624	6162	10092	

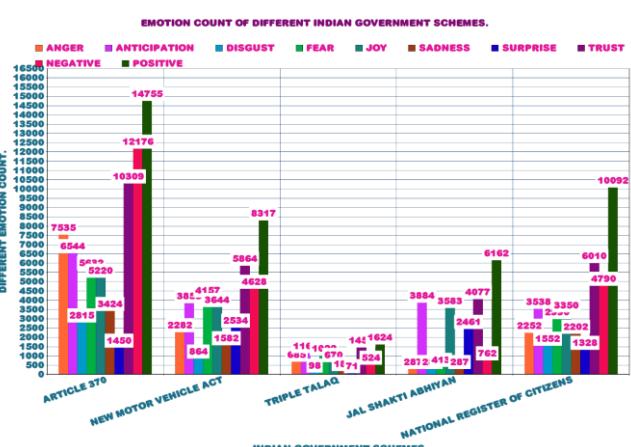


Figure6. The emotion of Article370, Motor Vehicle Act, Triple Talaq, Jal Shakti Abhiyan, National Register of Citizens.

Table 8. Percentage count of different Indian Government Schemes.

Scheme/Law/Act	Article370	Vehicle	New Motor	Triple Talaq	Jal Shakti	Abhiyan	NRC
Anger	10.78%	6.04%	9.13%	1.29%	5.91%		
Anticipation	9.36%	10.21%	15.48%	17.49%	9.28%		
Disgust	4.02%	2.29%	1.30%	1.28%	4.07%		
Fear	8.06%	11.01%	13.62%	1.86%	7.84%		
Joy	7.47%	9.65%	8.93%	16.13%	8.79%		
Sadness	4.90%	4.19%	2.49%	1.29%	5.77%		
Surprise	2.07%	6.71%	0.94%	11.08%	3.48%		
Trust	14.75%	15.54%	19.44%	18.36%	15.77%		
Negative	17.42%	12.26%	6.98%	3.43%	12.57%		
Positive	21.12%	22.04%	21.65%	27.75%	26.48%		

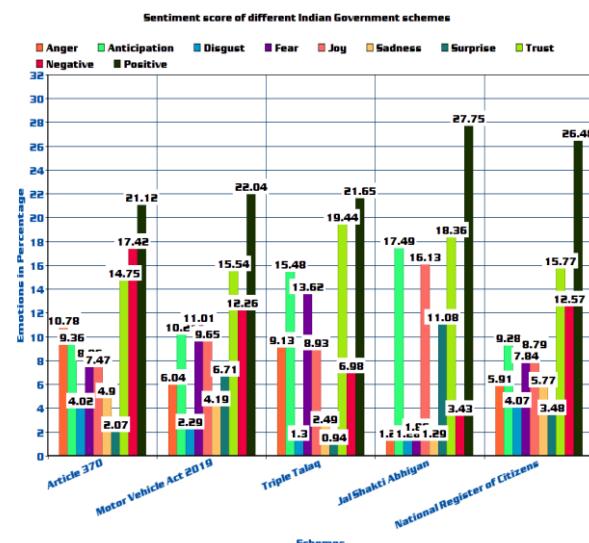


Figure 7. The emotion of Different Indian Government Schemes in Percentage.

VI. CONCLUSION

Sentiment Analysis is inferred as one of the most alluring domains that encourage to study and pertain to various sectors, a proficient practiced tool to analyze the user behavior. Whenever government policy is enforced it always has some marks specifically for the common peoples. The aim of this paper was to analyze various policies implemented by the Indian government using sentiment analysis approach. This work analyzed the emotions of people regarding the different Indian government policies,

laws and amendments on social network platforms via emotion-based sentiment analysis techniques. In this method, we compute and plot sentiment score and emotion count percentage of peoples regarding various government policies. The emotion-based method reflects in detail the sentiments of peoples. We Analyzed following policies, laws and schemes: Article370, New Motor Vehicle Act, Triple Talaq, Jal Shakti Abhiyan, NRC Assam. The result showed that most of the people welcomed all the policies of Indian government. The emotion-based sentiment analysis is very helpful for the evaluation of government initiatives and monitors the policies from people's perspective. Machine learning and Deep learning can be used for the classification of sentiments in future.

REFERENCES

1. G. Vinodhini and RM. Chandrashekharan, "Sentiment Analysis and Opinion Mining: A Survey" – International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 6, June 2012.
2. K. Arun, A. Srinagesh, M. Ramesh, "Twitter Sentiment Analysis on Demonetization tweets in India Using R language" International Journal of Computer Engineering In Research Trends, Volume 4, Issue 6, June-2017, pp. 252-258.
3. Vivek Sharma, Apoorv Agarwal, "Sentiments Mining and Classification of Music Lyrics using SentiWordNet", Symposium on Colossal Data Analysis and Networking 2016.
4. Tran Duc Chung, Rosdiazli Ibrahim, Sabo Miya Hassan, "Fast Approach for Automatic Data Retrieval using R Programming Language", Mediterranean Conference on Embedded Computing 2015.
5. Osmańczyk, Edmund Jan (2003), Encyclopedia of the United Nations and International Agreements: G to M, Taylor & Francis, pp. 1191–, ISBN 978-0-415-93922-5 Quote: "Jammu and Kashmir: Territory in northwestern India, subject to a dispute between India and Pakistan. It has borders with Pakistan and China.
6. "Triple Talaq verdict: What exactly is instant divorce practice banned by court". Hindustan Times. 22 August 2017. Retrieved 18 September 2017.
7. "The Triple Talaq Case - Supreme Court Observer". Supreme Court Observer. Retrieved 27 February 2018.
8. "Supreme Court scraps instant triple talaq: Here's what you should know about the practice". Hindustan Times. 22 August 2017. Retrieved 18 September 2017.
9. Pratap Bhanu Mehta (23 August 2017). "Small step, no giant leap". The Indian Express. Retrieved 18 September 2017.
10. <http://blog.revolutionanalytics.com/2017/01/cran10000.html>
11. Spencer J, Uchyigit G. Sentimentor: Sentiment analysis of Twitter Data. CiteSeerX 10M; 2012.
12. Rao NP, Srinivas SN, Prashanth CM. Real time opinion mining of Twitter Data. International Journal of Computer Science and Information Technologies. 2015; 6(3):2923–7.
13. Bharat R.Naiknaware, Seema Kawathekar, Sachin N.Deshmukh "Sentiment Analysis of Indian Government Schemes Using Twitter Datasets", IOSR Journal of Computer Engineering, e-ISSN: 2278-0661, p-ISSN: 2278-8727, PP 70-78.
14. Raju Ranjan, Sumana Gupta, "Supervised Texture Identification Using Dictionary Based Data Modelling", 978-1-4799-1812-6/14/\$31.00 ©2014 IEEE.
15. Cataldo Musto, Giovanni Semeraro, Marco Polignano."A comparison of Lexicon-based approaches for Sentiment Analysis of microblog posts".8th International Workshop on Information Filtering and Retrieval Pisa (Italy) December 10, 2014.
16. Verma, P., Khanday, A. M. U. D., Rabani, S. T., Mir, M. H., & Jamwal, S" Twitter Sentiment Analysis on Indian Government Project using R" October 2019.
17. Khanday, A. M. U. D., Amin, A., Manzoor, I., & Bashir, R. "Face Recognition Techniques: A Critical Review" 2018.

AUTHORS PROFILE



Pankaj Verma, Scholar, Department of Computer Sciences, BGSBU, Rajouri, India
pankajvermajmu@gmail.com

He has two research publications to his credit. His Area of research is Data Analytics.



Sanjay Jamwal, Sr. Assistant Professor, Department of Computer Sciences BGSBU Rajouri, India.

sanjayjamwal@bgsbu.ac.in

He around 55 publications to his credit and his area of research is Data Analytics. He is permanent member of CSI, ISC, IETE etc.