



ISSN Print: 2394-7500  
ISSN Online: 2394-5869  
Impact Factor: 5.2  
IJAR 2017; 3(7): 660-662  
www.allresearchjournal.com  
Received: 03-05-2017  
Accepted: 04-06-2017

**Dr. Kavita Pabreja**  
Associate Professor & Head-  
Department of Computer  
Science, Maharaja Surajmal  
Institute, (GGSIP University),  
Dwarka, Delhi, India

## GST sentiment analysis using twitter data

**Dr. Kavita Pabreja**

### Abstract

The growing popularity of social media has raised the opportunity for exploring and tracking the response of new reforms and policies in India. Social media has been used profoundly all over the world for analysis of political campaigns, stock market data, new product launch, movie release etc. Many researchers have been analyzing the tweets by citizens of a nation on Twitter which is a microblogging website where users read and write millions of tweets on a variety of topics on daily basis. In this paper, Twitter has been used as a forum to understand the sentiments of citizens of India towards recently launched Goods and Services Tax by Indian Government on 1<sup>st</sup> July 2017. The tweets originating in India on 30<sup>th</sup> June and 1<sup>st</sup> July have been analyzed. The emotions of public in terms of anger, anticipation, disgust, fear, joy, sadness, surprise have been extracted based on their live opinion.

**Keywords:** Twitter, Word cloud, Goods and Services Tax, Sentiment analysis, Indian Government.

### 1. Introduction

Goods and Services Tax (GST) is an indirect tax applicable throughout India which replaced multiple cascading taxes levied by the central and state governments. It was introduced as The Constitution (One Hundred and First Amendment) Act 2017<sup>[1, 2]</sup>, following the passage of Constitution 122nd Amendment Bill. The GST is governed by a GST Council and its Chairman is the Finance Minister of India. Under GST, goods and services is taxed at the following rates, 0%, 5%, 12%, 18%, 28%<sup>[2]</sup>. There is a special rate of 0.25% on rough precious and semi-precious stones and 3% on gold<sup>[3]</sup>.

The Goods and Services Tax (GST), India's biggest tax reform in 70 years of independence, was launched on the midnight of 30 June 2017<sup>[2a]</sup> by the Prime Minister of India Narendra Modi. The launch was marked by a historic midnight (June 30-July 1, 2017) session of both the houses of parliament convened at the Central Hall of the Parliament<sup>[4]</sup>.

Members of the Congress boycotted the GST launch altogether. They were joined by members of the Trinamool Congress, Communist Parties of India and the DMK, who reportedly found virtually no difference between the existing taxation system, and therefore claimed that the government was trying to merely rebrand the current taxation system but made it worse for common people by increasing existing rates on common items and reducing rates on luxury items<sup>[5a]</sup>. GST was initially proposed to replace a slew of indirect taxes with a unified tax and was therefore set to dramatically reshape the country's 2 trillion dollar economy. However, it has been met with sharp criticism from various fronts.

Recent years have witnessed the explosive growth of the social media content on the Internet, people now express their views on almost anything in discussion. There are many micro-blogging websites like Twitter, Facebook, and Tumbler etc. Twitter has become a very popular communication tool among Internet users and it is one of the most open and simplest platforms to share their opinions on different topics. The outgrowth of social media in expression of thoughts has resulted in the availability of huge volumes of data from masses. These social networking data and government's policies data can be combined to observe some useful sentiments of public. Availability of the voluminous data has opened an arena for conducting research in the domain that combines politics and social media both.

Twitter is one of the most popular social networking services used by today's generation. It enables users to send and read short 140-character messages called "tweets". When someone posts a new tweet on some topic, other likeminded users also join in the talk and this formulates a Twitter's network of likeminded users related to some topic. These set of tweets can provide an indicator of vote of confidence related to the topic of discussion.

### Correspondence

**Dr. Kavita Pabreja**  
Associate Professor & Head-  
Department of Computer  
Science, Maharaja Surajmal  
Institute, (GGSIP University),  
Dwarka, Delhi, India

As the number of followers is a measure of popularity, popular users are usually followed by other users, expanding the network more quickly. Twitter reflects the society's state of mind related to some topic. User's tweets are indicators of what people talk about or how they feel. It's important that we can use tweets to understand the public's opinion on different news and current topics.

The growing trends of Indian society participating in the social networking sites like Twitter, Facebook etc. to express their views regarding some topic of public domain is depicted clearly in the data collected by various statistical websites. According to statistical website number of active user in India were 11.5 million in the year 2013 and by the year 2016 number of active internet user has shown a tremendous growth of 101.75% reaching 23.2 million Twitter users<sup>[1]</sup>.

In this paper we present an analysis the twitter data for Goods and Services Tax recently introduced by Government of India. We gathered Twitter data using the streaming API to extract tweets related to GST.

## 2. Related Work

Social media has been explored to estimate the popularity of politicians, sentiments of general public towards some recently introduced policy maybe budget, tax reforms etc.<sup>[6]</sup>, to find out the sentiments of social media users<sup>[7, 12]</sup>. Social networking sites have also been used to compare people's political preferences expressed online with those observed by elections<sup>[9]</sup>. Social media can be analyzed on daily or hourly basis during an electoral campaign so as to get a detailed insight into emotions of voters<sup>[10]</sup>.

It is possible to track in real-time trends and capture any sudden change<sup>[11]</sup> by monitoring and analyzing the conversation on social networking sites and get the public opinion well before declaration of results of polls.

There are few studies that claim that analyzing social media allows a reliable forecast of the final result<sup>[12]</sup>. In a study by researchers<sup>[13]</sup>, it has been stated that the number of times a candidate is mentioned in blog posts is a good predictor of electoral success and can achieve better predictions than election polls. There are claims by some researchers that more the number of facebook supporters an electoral candidate has, better are the chances to win<sup>[14, 15]</sup>. On similar lines the authors<sup>[16]</sup>, compared party pointed out on Twitter with the results of the 2009 German election and discussed that the relative number of tweets related to each party is a good predictor of its vote share. There stands a better way to analyze tweets such that not just the count or mention of party name or candidate name is considered but the sentiment attached in tweets are also analyzed as stated by authors<sup>[17]</sup>.

A sentiment classifier based on lexical induction has been built by<sup>[18]</sup> and correlations between several polls conducted during the 2008 presidential election and the content of wall posts available on Facebook has been found. There are other studies by researchers<sup>[9]</sup> that show similar results displaying correlation between Obama's approval rate and the sentiment expressed by Twitter users. For predicting the results of both the 2011<sup>[12]</sup> and the 2012 legislative elections in the Netherlands<sup>[19]</sup>, sentiment analysis of tweets proved to perform quite well. Here the authors analyzed many different social media viz. Facebook, Twitter, Google, and YouTube.

## 3. Data collection and pre-processing

Twitter has been used to keep track of temporal nature of elections. The sentiments keep on changing based on some announcement or news by political parties. By creating a twitter API, 5,000 tweets have been collected and examined. The tweets published on Twitter's public message board one day prior to the GST implementation announcement and on the day of announcement (30-June 2017 and 1-July 2017), have been collected.

The tweets comprised of useful information related to GST besides special characters, punctuation marks and emojis. The data collected hence has been cleaned so as to remove punctuation symbols, special characters. All tweets have been converted to lowercase and finally a word corpus has been generated.

## 4. Technique applied

We have used the R language which is a powerful language used widely for data analysis and statistical computing. It was developed in early 90s. R has enough provisions to implement machine learning algorithms in a fast and simple manner. We have generated Word Cloud using R corresponding a set of tweets focusing on GST.

Sentiment analysis of these 5,000 tweets has also been done to understand the emotions of public towards introduction of Goods and Services Tax. We have used many packages viz. `syuzhet`, `lubridate`, `ggplot2`, `scales`, `reshape2`, `dplyr` to facilitate understanding of emotions and plotting the same.

## 5. Results and discussions

All tweets mentioning "GST" have been analyzed from following perspective:-

- Word cloud generation
- Sentiment analysis

The Word Cloud has been generated corresponding to tweets mentioning "gst" as shown in Fig. 1.



Fig 1: Word cloud "GST"

Finally the sentiment analysis of the set of tweets has been done to understand the emotions of public for GST. Emotions of anger, anticipation, disgust, fear, joy, sadness and surprise have been extracted using `get_nrc_sentiment()` function of "R" software and other pre-processing of tweets. From the sentiment analysis of "gst" tweets it is very clear that there is a feeling of anticipation for "GST". There is almost equal joy and sadness emotion for "GST" in India.

The tweets show that the citizens of India have hardly any emotion of anger or fear as observed from the tweets. The detailed sentiment analysis has been shown in Fig. 2.

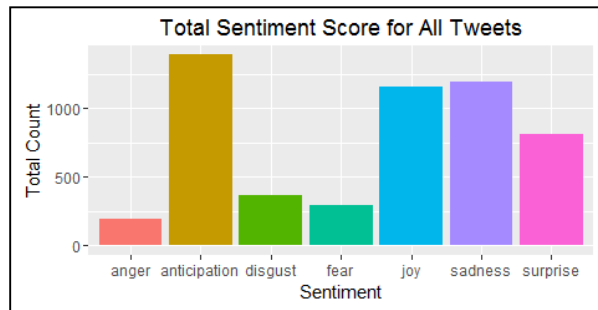


Fig 2: Sentiment analysis of tweets mentioning “GST”

## 6. Conclusions and future directions

The increasing number of social media websites by Internet users has raised the interest about the opportunity to understand the relation between people’s preferences and actual political behavior. This study focuses on the question that whether the data from social networking sites can be utilized to interpret the attitude of citizens of a nation towards various policies.

We analyzed 5,000 twitter messages mentioning keyword viz. “GST” for two days, viz. one day prior to announcement and on the day of announcement We have observed that twitter is very commonly being used as a platform for deliberation by citizens of India. It has been concluded that social media is a powerful and reliable source of public opinion as far as a nation like India is concerned. The discussions on twitter are equivalent to traditional discussions and are capable enough to give a fair idea of emotions of general public. We have done sentiment analysis of emotions of people which shows people’s acceptance for GST but with too much of anticipation feeling. In future, we plan to convert this analysis in real time corresponding to tweets arriving on temporal scale. Also we can geographically divide and analyze the tweets according to states.

## References

1. The constitution (one hundred and first amendment) act, Amendment No. 101 of Retrieved from Internet, 2016, 8.
2. India's midnight 'tryst with destiny': GST rolled out. Dynamite News.com. (ANI). Retrieved from Internet, 2017.
3. All your queries on GST answered. The Hindu. Retrieved, 2017-06-30.
4. Hindustan Times. News. Retrieved from Internet.
5. Congress To Boycott GST Launch, Arun Jaitley Suggests Broader Shoulders, NDTV, 29, 2017.
6. Gloor PA, Krauss J, Nann S, Fischbach K, Schoder D. Web science 2.0: Identifying trends through semantic social network analysis CSE. International Conference on Computational Science and Engineering, Vancouver, BC, 2009; 4:215-222.
7. Barbera P. Birds of the same feather tweet together. Bayesian ideal point estimation using twitter data, 2012.
8. Conover M, Goncalves B, Ratkiewicz J, Flammini A, Menczer F. Predicting the political alignment of twitter users. Proceedings of 3rd IEEE Conference on Social Computing Social Com, 2011.
9. O'Connor B, Balasubramanyan R, Routledge BR, Smith NA. From tweets to polls: Linking text sentiment to public opinion time series. Proceedings of the International AAAI Conference on Weblogs and Social Media, Washington, DC, 2010.
10. Xin J, Gallagher A, Cao L. The wisdom of social multimedia: Using flickr for prediction and forecast. ACM Multimedia, International Conference, New York, NY, 2010.
11. Jensen MJ, Jorba L, Anduiza E. Introduction. In E. Anduiza, M. Jensen, & L. Jorba (Eds.), Digital media and political engagement worldwide: A comparative study. New York, NY: Cambridge University Press, 2012, 1-15.
12. Tjong Kim Sang E, Bos J. Predicting the 2011 Dutch senate election results with twitter. Proceedings of SASN 2012, the EACL 2012 Workshop on Semantic Analysis in Social Networks, Avignon, France, 2012.
13. Ve'ronis J. Citations dans la presse et re'sultats du premier tour de la pre'sidentielle, 2007.
14. Upton G. Jr. Does attractiveness of candidates affect election outcomes? 2010.
15. Williams C, Gulati G. What is a social network worth? Facebook and vote share in the presidential primaries. Boston, MA: Annual Meeting of the American Political Science Association, 2008, 1-17.
16. Tumasjan A, Sprenger TO, Philipp GS, Welpe IM. Predicting elections with twitter: What 140 characters reveal about political sentiment. Social Science Computer Review, 2011; 29:402-418.
17. Chung J, Mustafaraj E. Can collective sentiment expressed on twitter predict political elections? Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, San Francisco, CA, 2011.
18. Lindsay R. Predicting polls with Lexicon, 2008.
19. Sanders E, den Bosch AV. Relating political party mentions on Twitter with polls and election results, 2013.