Proceedings of the 12th INDIACom; INDIACom-2018; IEEE Conference ID: 42835
2018 5th International Conference on "Computing for Sustainable Global Development", 14th - 16th March, 2018
Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), New Delhi (INDIA)

Sentiment Analysis of Demonetization in India

Charu Gupta
Department Of Computer Science
Jamia Milia Islamic
New Delhi, India
guptacharu1996@gmail.com

Cheshta Arora
Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM)
New Delhi, India
arora.cheshta26@gmail.com

Manisha Paruthi
Department Of Computer Science
Delhi University
New Delhi, India
manisha.mca17.du@gmail.com

Rakhi Joon
Department Of Computer Science
Delhi University
New Delhi, India
rjoon30@gmail.com

Vaishali Joshi
Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM)
New Delhi, India

Abstract— Sentiment Analysis is one of the key areas of research in the field of NLP and gaining more and more popularity day by day. There are many challenges which can be explored via sentiment analysis like the political, social, and many more. The challenge considered in this paper is demonetization in India. Since from last few months demonetization affected the Indian economy as well as public in various ways and the reviews are a mixture of positive, negative and neutral response. So, this is a major problem to be dealt with and the sentiments of public on social media plays a vital role for any of these challenges discussed so far. Here, in this paper the twitter feeds are collected and analyzed further to carry out the proposed task. Twitter is one of the most popular social media used by most of the people involved in any social community and also very popular. The present paper gives an insight to collecting and analyzing the responses on suitable social networks. Further result analysis of positive, negative and neural tweets are done for accuracy.

Keywords— sentiment;positive;negative;neutral;API's;devloper twitter;python libraries

I. INTRODUCTION

Sentiment analysis is helpful to an extensive variety of issues that are important to human-computer interaction practitioners, experts and researchers, and additionally those from fields, like sociology, advertising and marketing, financial aspects, psychology and political science. The inborn nature of micro blog content, observed on Facebook and Twitter, gives serious challenges to sentiment analysis and its practical applications. Some of these starts from the sheer rate and volume of social content generated by users, combined with the contextual variations and sparseness which results from tendency to use abbreviated language conventions to express sentiments and shortness of the text used.

Sentiment analysis can be done on various challenges and one of the challenges considered in this paper is "Demonetization in India". On 8 November 2016, Prime Minister Narendra Modi declared the Government of India's choice to cancel the lawful tender character of INR 500 and INR 1,000 banknotes with impact from 9 November 2016. He likewise reported the issuance of new INR 500 and INR 2,000 banknotes in return for the old banknotes. While the declaration basically rendered the old banknotes invalid from 9 November 2016, the Ministry of Finance has been checking the usage of these measures four key contentions are being sent about the choice:

- It is guaranteed that the choice will affect the poor than others, for the most part recounted, reports recommend.
- The cash related stun can be, and will be, quickly overpowered by the use of budgetary approach instruments to restore liquidity.
- This decision will encourage the route toward making India a "cashless economy", with benefits that will make without further ado costs helpful.
- Since the decision is noticeable, it must be awesome. This raises an intriguing issue: in a greater part run government, can there be a predominant measure of conventionality of an approach than its fame?

Alternative payment techniques, for example, e wallets, online exchanges utilizing e-managing an account, charge and Visa use have expanded fundamentally because money exchanges were affected by a lessening in liquidity. This will build utilization of such installment underlying frameworks, and empower a move towards an effective cashless foundation.

With money exchanges affected by a lessening in liquidity, alternative payment techniques, for example, e wallets, online exchanges utilizing e-managing an account, charge and Visa use have expanded fundamentally. This will build utilization of such installment frameworks, and empower a move towards an effective cashless foundation.

Since cash currency is an important part of Indian economy and the sentiment analysis of public reviews can be done regarding this. In this paper the review of sentiment analysis is done along with the classification of positive, negative and neutral tweets on demonetization taken from twitter. The tweets are collected using python and further processed through file handling. The results are further analyzed for accuracy.

The paper is organized in the following manner: Section 2 discusses about various social media from that information can be collected to be further used for experimental tasks. In Section 3 we have discussed about extraction of data using twitter developer and python (textblob and tweepy libraries) its processing. Section 4 discusses about the implementation techniques. In Section 5, we present our result analysis on sentiments and their effects on demonetization. Section 6 discusses about the impact of demonetization. We conclude in Section 7.

II. SENTIMENT ANALYSIS

Sentiment Analysis is the way toward inferring the opinion or state of mind of a speaker. Sentiment Analysis is basic since encourages you see what clients like and aversion about you and your image. Customer feedback—from online networking, site,call focus specialists, or some other source—contains a fortune trove of valuable business data. In any case, it isn't sufficient to recognize what customers are discussing. One should likewise know how they feel. Sentiment Analysis is one approach to reveal those sentiments. Once in a while known as "Opinion mining," sentiment analysis can fill ones in regarding whether there has been an adjustment in general assessment toward any part of your business. Peaks or valleys in slant scores give you a place to begin in the event that you need to influence item upgrades, to prepare deals or customer mind operators, or make new promoting efforts. Twitter is an social networking and online news service where users can post and interact with messages, "tweets," limited to 140 characters. Registered users can post tweets, whereas the unregistered individuals can just read them. Users get to Twitter through its site interface, SMS or a cell phone application

Twitter Inc. is situated in San Francisco, California, United States, and has in excess of 25 workplaces around the globe.

III. DATA EXTRACTION

Twitter is a popular social blogging platform where users express opinions about different topics through small status messages called tweets. In the following paper we propose a strategy to automatically separate notion (positive or negative) from a tweet. This automation of extracting sentiments from is very useful because it allows feedback to be aggregated without manual intervention. With a specific end goal to prepare a classifier, managed adapting as a rule requires hand-marked preparing of information. With the widespread scope of points which are talked about on Twitter, it would be very hard to physically collect enough information to prepare an estimation classifier for tweets. Our solution is to extract data from twitter using python libraries.

A. Tweepy Library

Tweepy is open-sourced library, availing access key from twitter developers[8] that allows Python to communicate with Twitter platform and use its APIs.

B. NumPy Library

NumPy library is used for including support for large, multi-dimensional arrays and matrices, along with that it provides a large range of high-level mathematical functions to operate on these multi-dimensional arrays.

C. Textblob Library

Textblob library is meant for processing of textual data. It provides a simple API for converting data into natural language processing (NLP) such as part-of-speech tagging, sentiment analysis, noun phrase extraction, translation, classification and more. With the help of python it is easy to extract large amounts of tweets with specific keywords in them.

Some of the examples are:

Part-of-speech Tagging:

These tags can be accessed using the **tags** property

```
from textblob import TextBlob
#create a TextBlob object
pos=TextBlob("Big yellow Door is such an amazing cafe")
print("tags: ",pos.tags)

#expected output:
#tags: [('Big', 'NNP'), ('yellow', 'JJ'), ('Door', 'NNP'), ('is', 'VBZ'), ('such', 'JJ'), ('an', 'DT'), ('amazing', 'JJ'), ('cafe', 'NN')]
```

1. Noun Phrase Extraction
2. Sentiment Analysis:

Each word in the lexicon has scores for:

1) polarity: negative vs positive

   range= (-1.0,1.0)

2) subjectivity: objective vs subjective

   range= (0.0,1.0)

3) intensity: modifies next word?(impact n next word)

   range= (x0.5,x2.0)

IV. IMPLEMENTATION

A. *Extraction of Tweets*

Tweets are a dependable source of data on the grounds that individuals tweet about everything without exception they do including purchasing new items and checking on them. In addition, all tweets contain hash labels which make recognizing relevant tweets a simpler task. Various research works has just been done on twitter information. The majority of which for the most part shows how valuable this data is to anticipate different results. Our current research deals with analysis of tweets on demonetization in India and analyzing people reviews

1. Usage Of API:

Tweepy supports accessing Twitter the newer method OAuth to use the Twitter API. One of the main usage cases of tweepy is it monitors the tweets and performs actions when some event happens. StreamListener object is the key component which monitors tweets in real time and catches them.

2. Access Token:

It first obtains an OAuth access token for a Twitter user (or, one could give Application-only authenticated requests, when user context is not required), to make authorized access to Twitter's APIs. Once an access to token and token secret has been done, the accessed Twitter API is your oyster! By following the steps described in Authorizing a request on twitter developers, authorized requests to the Streaming APIs and REST API can be issued. If the OAuth process sounds difficult, one can use Web Intents, which do not need access tokens to work together with the Twitter API.

B. *Pre Processing of Extracted Data*

To work with collected data we have to perform some processing on it. It is a basic step before the collected data is ready for giving some sort of results or some analysis. A large portion of accessible text data is very unstructured and redundant in nature – to accomplish better insight of data or to build better algorithms, it is important to use clean data. For instance, data available on social media platforms like twitter is profoundly unstructured – it is in the form of a casual conversation – grammatical mistakes, awful punctuation, use of slang, nearness of undesirable contents like URLs, Stop words, Expressions and so forth are the typical suspects. Following are the means for data cleaning:

Escaping HTML characters: The Data collected from web contains a lot of html elements like <   > & which get mixed(embedded) with the original data while extraction of data. It is important to remove tose elements.

Disentangling information: This is the way toward changing data from complex images to straightforward and less demanding to comprehend characters**.**

Punctuation Lookup: To keep away from any word sense ambiguity in content, it is prescribed to keep up appropriate structure in it and to comply with the principles of setting free sentence structure.

Removal of Stop-words: Data analysis is data driven and is done at word level, the regularly occurring words, or stop words are removed while preprocessing the data. To remove these words either we can use predefined language specific libraries or create a long list of stop-words.

Punctuations Removal: Punctuation marks needs to be managed. For Eg.: ".", ",","?" are essential accentuations that ought to be held while others should be removed.

Split Attached Words: Many users on social media write text data, which is informal in nature, not able to understand by the system . Most of the tweets are written with many attached words like AwesomeDay,NiceFood etc. These words should be split into their original forms using some rules these rules are simple.

TABLE 1: KEYWORD CLASSIFICATION WITH POLARITY

Sentiment	Keywords	Polarity
POSITIVE	# great	1.000
	# amazing	0.8
	#good	0.90999
	#adorable	0.5
	#adequate	0.5
NEGATIVE	#alcoholic	-0.5
	#anger	-0.7
	#annoy	-0.6
	#ascetic	-0.45
	#chitchat	-0.2
NEUTRAL	#characteristic	0.0
	#certain	0.0
	#christian	0.0
	#hindu	0.0
	#crucial	0.0

V. RESULT ANALYSIS

With the help of pie chart people's reviews on demonetization are analyzed and further tweets through keywords (positive, negative, neutral) are also extracted out.

DEMONETISATION:

Positive: Shell company,blackmoney,account frozen is really government used these words in their press note or without analys… https://t.co/nmsctEDLff

Negative: No #BlackMoney recovered

Neutral: Electricity connections to #Farmers itself is not being p…

TABLE 2: SENTIMENT BASED POLARITY PERCENTAGE

Sentiments	Percentage
Positive	25.675675675675677
Negative	37.83783783783784
Neutral	36.486486486486483

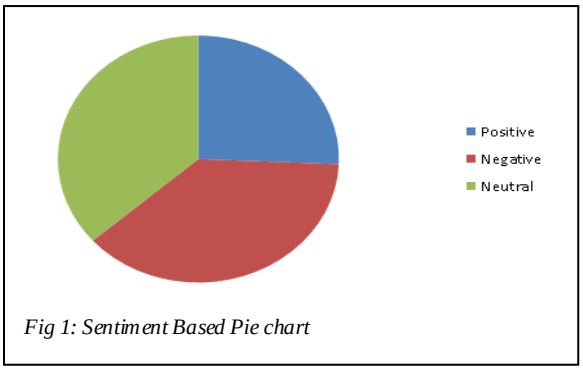

Fig 1: Sentiment Based Pie chart

VI. IMPACT

Demonetization leads people to use digital wallets to pay for their daily expenses and necessities. Well-known brands in this field are Mobikwik and Paytm, as they have the largest reach from airplane ticket bookings to local shops as well. As a result, more and more people are getting migrated to digital wallets and the services have seen a boost in user base since 8th November and a sudden increase in social media page likes is observed.

Fig 2: Graph Representing Paytm Likes On Social Media [6]

Fig 3: Graph Representing Mobikwik Likes On Social Media[7]

VII. CONCLUSION AND FUTURE SCOPE

In this research a methodology is discussed, the popularity/opinion/sentiment of people through tweets can be determined. The reason for choosing 'demonetization' for analysis, is the availability of reasonable amount of tweets based on demonetization. The number of tweets must be significant for accurate results. Initially the tweets were filtered manually. Only the tweets with selected grammatical relations involving previously chosen keywords were picked for analysis. Each of the tweets in this filtered set were then given some rating according to their sentiments.. Each individual word in every tweet was assigned its own part of speech tag . While in this study we focused only on tweets which are extracted in the text form, this method can be extended to use of emoticons to detect people's response can give positive results. Sentiment analysis is a vast field of

study with enormous possibilities .Few of them are listed as(1) ironical statements, metaphors, sarcastic comments can be interpreted according to the context.

REFERENCES

[1] Rakesh Kumar Assistant professor, Department of economics, Maharaja Agrasen college, University of Delhi*) – De-constructing demonetisation in India* .

[2] B. Liu – "*Sentiment analysis and opinion mining*," Synthesis Lectures on Human Language Technologies, vol. 5, no. 1, pp. 1– 167, 2012.

[3] Nithish et al. (2013) mainly focused on market reaction using sentiment analysis

[4] Ostrowski (2012) – Discussion on social media platform.

[5] Syed Akib Anwar Hridoy, M. Tahmid Ekram, Mohammad Samiul Islam, Faysal Ahmed and Rashedur M. Rahman – *Localized twitter opinion mining using sentiment*

[6] https://www.facebook.com/MobiKwik/about/

[7] https://www.facebook.com/pg/Paytm/community/?ref=page_internal.

[8] https://dev.twitter.com/