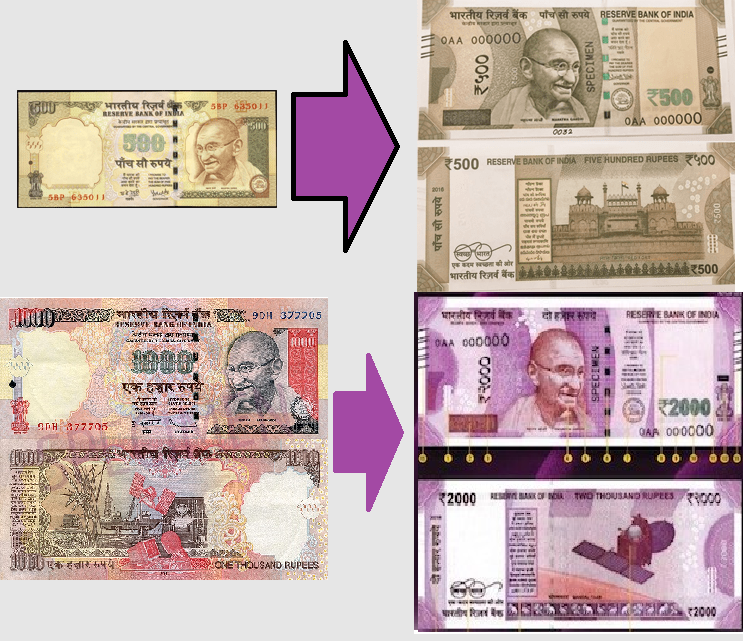
**Analysis of Demonetization in India using machine learning**



**Made by: Gursimran Singh**

**Index**

|  |  |
| --- | --- |
| **Tittle** | **Page No.** |
| 1. Introduction | 3 |
| 1. Reaction | 3 |
| 1. Problem Statement | 5 |
| 1. Available Approaches | 6 |
| 1. Implementation | 7 |
| 1. Working Code | 10 |
| 1. Result | 15 |
| 1. Screenshots of Result | 17 |
| 1. Conclusion and Improvement | 18 |
|  |  |

**Introduction**

On 8 November 2016, the [Government of India](https://en.wikipedia.org/wiki/Government_of_India) announced the [demonetization](https://en.wikipedia.org/wiki/Legal_tender#Demonetisation) of all Rs500 and Rs1000 banknotes. The government claimed that the action would curtail the shadow economy and crack down on the use of illicit and counterfeit cash to fund illegal activity and terrorism. The sudden nature of the announcement—and the prolonged cash shortages in the weeks that followed—created significant disruption throughout the economy, threatening economic output. The move was heavily criticized as poorly planned and unfair, and was met with protests, litigation, and strikes. Initially, the move received support from several bankers as well as from some international commentators. It was heavily criticized by members of the opposition parties, leading to debates in both houses of [parliament](https://en.wikipedia.org/wiki/Indian_Parliament) and triggering organized protests against the [government](https://en.wikipedia.org/wiki/Modi_ministry) in several places across India. The move is considered to have reduced the country's [GDP](https://en.wikipedia.org/wiki/Gross_domestic_product) and [industrial production](https://en.wikipedia.org/wiki/Industrial_production). As the cash shortages grew in the weeks following the move, the demonetization was heavily criticized by prominent economists and by world media.

1. **Reaction**
2. **Support**

The decision met with mixed initial reactions. Several bankers supported the move adding that it would also accelerate e-commerce. Finance Minister said that demonetization would clean the complete economic system, increase the size of economy and revenue base. He mentioned the demonetization along with the upcoming [Goods and Services Tax (GST)](https://en.wikipedia.org/wiki/Goods_and_Services_Tax_(India)) as "an attempt to change the spending habit and lifestyle."

The [Indian National Congress](https://en.wikipedia.org/wiki/Indian_National_Congress)  welcomed the move but remained skeptical on the consequences that would follow. Chief Minister of Bihar supported the move. The demonetization also got support from Chief Minister of Andhra Pradesh ,Former Chief Election Commissioner and [Anna Hazare](https://en.wikipedia.org/wiki/Anna_Hazare) . By and large, international response was positive which saw the move as a bold crackdown on corruption. Chinese state media [*Global Times*](https://en.wikipedia.org/wiki/Global_Times) praised the move and termed it as "fierce fight against black money and corruption."

1. **Criticism**

The Indian Supreme Court while hearing one among a slew of cases filed against the sudden demonetization decision in various courts observed that it "appears to be carpet bombing and not surgical strike" which government repeatedly claims it to be.

Chief Ministers of several Indian states led major protests against the decision in their states and in parliament. Initially, the move to demonetize and try to hinder black money was appreciated, but the manner in which it was carried out by causing hardships to common people was criticized.

**Hence, having a bunch of mixed reactions, finding an overall opinion of the country based upon sentiment analysis techniques provide a method to judge demonetization in India**.

**Problem Statement**

To analyze Reviews on twitter about Demonetization in India and processing the reviews using machine learning and thus analyzing the overall reaction of the country by generating an analysis score.

**Available Approaches**

There are various ways of analyzing twitter data and generating sentiments. Some of them are mentioned below:

1. Sentiment Analysis of Twitter Data using NLTK

Natural Language Toolkit (NLTK) is library in Python, which provides a base for building programs and classification of data. NLTK is a collection of resources for Python that can be used for text processing, classification, tagging and tokenization. This toolbox plays a key role in transforming the text data in the tweets into a format that can be used to extract sentiment from them. NLTK provides various functions which are used in pre-processing of data so that data available from twitter become fit for mining and extracting features. NLTK support various machine learning algorithms which are used for training classifier and to calculate the accuracy of different classifier

1. Sentiment Analysis using TextBlob

[TextBlob](https://github.com/sloria/textblob) is an open source text processing library written in Python. It can be used to perform various natural language processing tasks such as part-of-speech tagging, noun-phrase extraction, sentiment analysis, text translation, and many more. The basic steps for using TextBlob are:

* Create a TextBlob object, passing a string with the text we want to work with.
* Use various methods and attributes of the resulting object to get at various parts of the text.

TextBlob can also tell us what part of speech each word in a text corresponds to. It can tell us if a word in a sentence is functioning as a noun, an adjective, a verb, etc.

**For the purpose of the project, second method i.e. Sentiment Analysis using TextBlob is used.**

**Implementation**

Following Steps are followed for sentiment analysis of the tweets related to demonetization.

### Authentication with twitter account

### Before analyzing reviews related to demonetization, authentication of twitter account has to be done using personal customer key, consumer secret, access token and access token secret for accessing twitter API. This authentication is carried using ‘tweepy’ package.

key = '\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*'

secret = '\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*'

token = '\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*'

token\_secret = '\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*'

# authenticating

try

self.auth = OAuthHandler(key, secret)

self.auth.set\_access\_token(token, token\_secret)

# creating tweepy API object to fetch tweets

self.api = tweepy.API(self.auth)

except

print("Authentication Error")

1. Cleaning of Tweets

Clean tweet text is obtained by removing links, special characters

using simple re package statement

join(re.sub("(@[A-Za-z0-9]+)|([^0-9A-Za-z \t])|(\w+:\/\/\S+)", " ", tweet).split())

1. Access Tweets

Tweets are accesses using API.search function

all\_tweets = self.api.search(q = query, count = count)

1. Sentiment Analysis of Tweets

Tweets are classified as positive, negative and neutral based on sentiment polarity using TextBlob function found in textblob package. TextBlob is a Python (2 and 3) library for processing textual data. It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification and translation.

def sentiment\_analysis(self, tweet)

score = TextBlob(self.clean\_tweet(tweet))

if score.sentiment.polarity > 0:

return 'positive'

elif score.sentiment.polarity == 0:

return 'neutral'

else:

return 'negative'

1. Calculating Analysis Score

TextBlob gives the polarity for a tweet between (-1 ,1). -1 refers the most negative tweet and 1 refers to most positive tweet. The worked out formula for calculating analysis score is as follows:

**For positive tweets : score (in %) = ((polarity\*0.5) + 0.5 )\*100**

**For neutral tweets : score (in %) = 50**

**For negative tweets: score (in %) = ((polarity + 1)\*50)**

def calculate\_score(self,tweet):

polar = TextBlob(self.clean\_tweet(tweet))

#calculating overall score

if polar.sentiment.polarity > 0:

c = polar.sentiment.polarity;

c = c\*10 + 10;

#print int(c);

self.graph[int(c)] = self.graph[int(c)] + 100;

return ((polar.sentiment.polarity\*0.5) + 0.5 )\*100

elif polar.sentiment.polarity == 0:

self.graph[11]+=1;

return 50

else:

c = polar.sentiment.polarity;

c = c\*10 + 10;

#print int(c);

self.graph[int(c)] = self.graph[int(c)] + 100;

return ((polar.sentiment.polarity + 1 )\*50)

1. Plotting Graph

Finally, a graph is plotted between no. of tweets and polarity.

def plot\_graph(self):

base = [-1,-0.9,-0.8,-0.7,-0.6,-0.5,-0.4,- 0.3,0.2,0.1,0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0]

plt.plot(base, self.graph)

plt.ylabel('no. of tweets')

plt.xlabel('Analysis score')

plt.show()

### Working Code

import re

import tweepy

from tweepy import OAuthHandler

import matplotlib.pyplot as plt

from textblob import TextBlob

class Demonetization(object):

overall\_score=0;

total=0;

graph=[0]\*21;

def authentication(self):

key = '\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*'

secret = '\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*'

token = '\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*'

token\_secret = '\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*'

# attempt authentication

try:

# create OAuthHandler object

self.auth = OAuthHandler(key, secret)

# set access token and secret

self.auth.set\_access\_token(token, token\_secret)

# create tweepy API object to fetch tweets

self.api = tweepy.API(self.auth)

except:

print("Place")

print("Error: Authentication Failed")

def plot\_graph(self):

base = [-1,-0.9,-0.8,-0.7,-0.6,-0.5,-0.4,-0.3,-0.2,- 0.1,0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0]

plt.plot(base, self.graph)

plt.ylabel('no. of tweets')

plt.xlabel('Analysis score')

plt.show()

def clean\_tweet(self, tweet):

return ' '.join(re.sub("(@[A-Za-z0-9]+)|([^0-9A-Za-z \t])|(\w+:\/\/\S+)", " ", tweet).split())

def get\_tweet\_sentiment(self, tweet):

'''

Utility function to classify sentiment of passed tweet

using textblob's sentiment method

'''

# create TextBlob object of passed tweet text

score = TextBlob(self.clean\_tweet(tweet))

# set sentiment

if score.sentiment.polarity > 0:

return 'positive'

elif score.sentiment.polarity == 0:

return 'neutral'

else:

return 'negative'

def calculate\_score(self,tweet):

polar = TextBlob(self.clean\_tweet(tweet))

#calculating overall score

if polar.sentiment.polarity > 0:

c = polar.sentiment.polarity;

c = c\*10 + 10;

#print int(c);

self.graph[int(c)] = self.graph[int(c)] + 1;

return ((polar.sentiment.polarity\*0.5) + 0.5 )\*100

elif polar.sentiment.polarity == 0:

self.graph[11]+=1;

return 50

else:

c = polar.sentiment.polarity;

c = c\*10 + 10;

#print int(c);

self.graph[int(c)] = self.graph[int(c)] + 1;

return ((polar.sentiment.polarity + 1 )\*50)

def get\_tweets(self, query, count = 1000):

tweets = []

try:

# call twitter api to fetch tweets

all\_tweets = self.api.search(q = query, count = count)

# parsing tweets one by one

for tweet in all\_tweets:

# empty dictionary to store required params of a tweet

parsed\_tweet = {}

# saving text of tweet

parsed\_tweet['text'] = tweet.text

# saving sentiment of tweet

parsed\_tweet['sentiment'] = self.get\_tweet\_sentiment(tweet.text)

self.total = self.total + 1;

self.overall\_score= self.overall\_score + self.calculate\_score(tweet.text);

# appending parsed tweet to tweets list

if tweet.retweet\_count > 0:

# if tweet has retweets, ensure that it is appended only once

if parsed\_tweet not in tweets:

tweets.append(parsed\_tweet)

else:

tweets.append(parsed\_tweet)

# return parsed tweets

return tweets

except tweepy.TweepError as e:

# print error (if any)

print("Error : " + str(e))

def main():

api = Demonetization()

api.authentication()

tweets = api.get\_tweets(query = 'Demonetization', count = 10000)

print ("Overall Success of Demonetization =")

print api.overall\_score/api.total

print len(tweets);

api.plot\_graph();

ptweets = [tweet for tweet in tweets if tweet['sentiment'] == 'positive']

print("Percentage of Positive Tweets:%".format(100\*len(ptweets)/len(tweets)))

ntweets = [tweet for tweet in tweets if tweet['sentiment'] == 'negative']

print("Percentage of Negative tweets%".format(100\*len(ntweets)/len(tweets)))

#print("Neutral tweets percentage: {} % ".format(100\*len(tweets - ntweets - ptweets)/len(tweets)))

print("\n\nPositive tweets:")

for tweet in ptweets[8:18]:

print(tweet['text'])

print("\n\nNegative tweets:")

for tweet in ntweets[:10]:

print(tweet['text'])

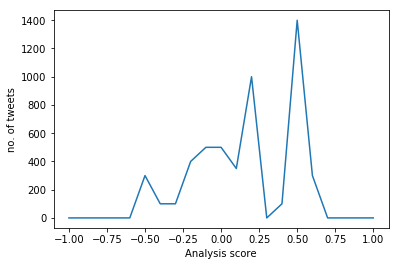
if \_\_name\_\_ == "\_\_main\_\_":

# calling main function

main()

**Result**

1. **Overall Success of Demonetization = 54.55%**
2. Plot:



1. Percentage of Positive Tweets = 30%
2. Percentage of Negative Tweets = 19%
3. First 5 Positive Tweets:

@interceptor4h @raggedtag Basically entire IAC gang supports everything CIA wants. Aadhaar, EVMs, demonetization, etc ;)

A glance on successful implementation of a decision that revealed corrupt elements.

#DeMonetisation #GoT

https://t.co/Wh6DedxgLb

@yasirbadoo @timesofindia How long will you live in denial? Demonetization proved that these jokers are working ONL… https://t.co/5QZg1DIW6b

RT @thesuniljain: I'm not saying we don't need cash, indeed I'm a critic of demonetization, but believe govt has done an amazing job in dig…

RT @imeerage: Record global harvest of cereals, India too heads for a new high

No effect of #Demonetization or Cash crunch

https://t.co/…

RT RadhaGiri: 811 is the day when india changed with demonetization - kotak has launched an innovative product with the same name! Smart mo…

RT @RadhaGiri: 811 is the day when india changed with demonetization - kotak has launched an innovative product with the same name! Smart m…

811 is the day when india changed with demonetization - kotak has launched an innovative product with the same name! Smart move! #TieConPune

The latest The Demonetization Daily! https://t.co/PdD6fKHf91 #news

Anyone wants to guess how they got the loans and why they can't be recovered? Demonetization was one way to shore u… <https://t.co/yojjJHJO1m>

1. First 5 Negative Tweets:

Glorifying Indian economic reforms after the horror of demonetization - is ECRI a pale horseman of the apocalypse? https://t.co/gbghnzSIvb

RT @maratheshail: Demonetization has killed the rural economy &amp; the farmers, they said, as India heads for a record cereal crop

https://t.c…

Gold imports gain since demonetization: Black money turned white, then yellow https://t.co/uBqVJMbfFo via @wordpressdotcom

@snsachinnandu So you accept that demonetization is failed.

RT @Stupidosaur: @interceptor4h @raggedtag Useless Kejri also didn't speak up correct thing which would have defeated demonetization.I told…

@interceptor4h @raggedtag Useless Kejri also didn't speak up correct thing which would have defeated demonetization… https://t.co/Ja2TWMWtFJ

@narendramodi hi modiji..i want to share my experience which i was felt today. im not able to understand what is purpose of Demonetization

@sardesairajdeep that's y BJPeee did a demonetization. To break the money power of rest of the parties.but poor bhakt think it's to curb BM😂

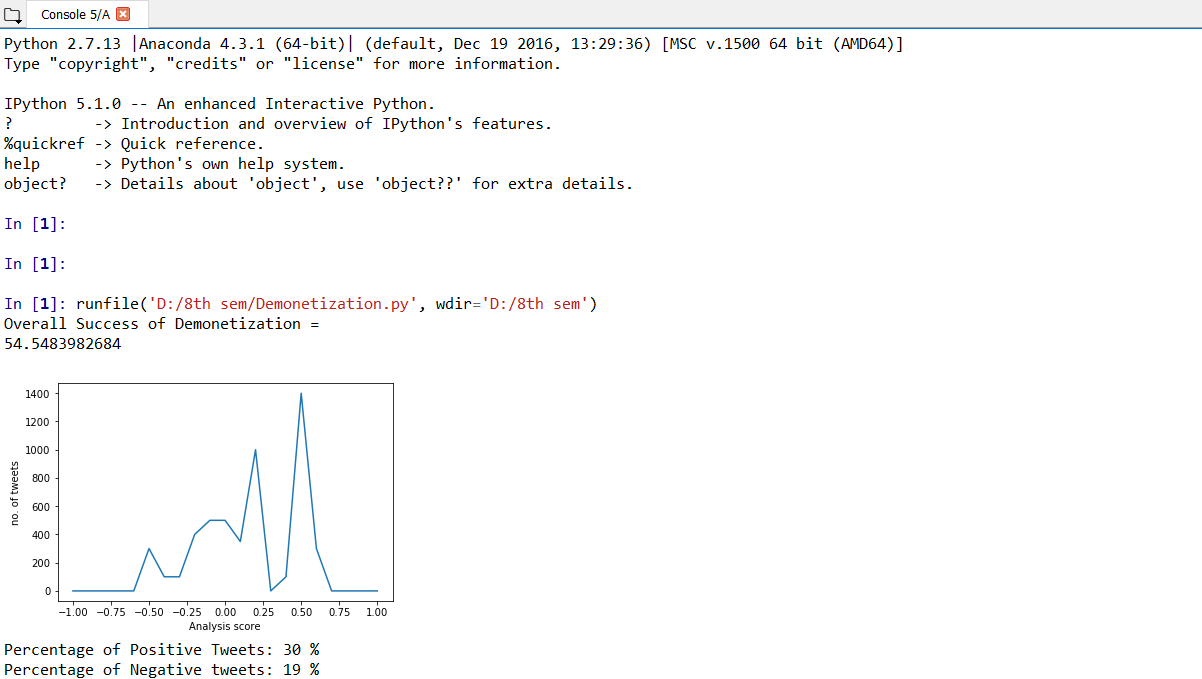
@narendramodi demonetization was a blunder mistake.but for fool's u become a HERO.

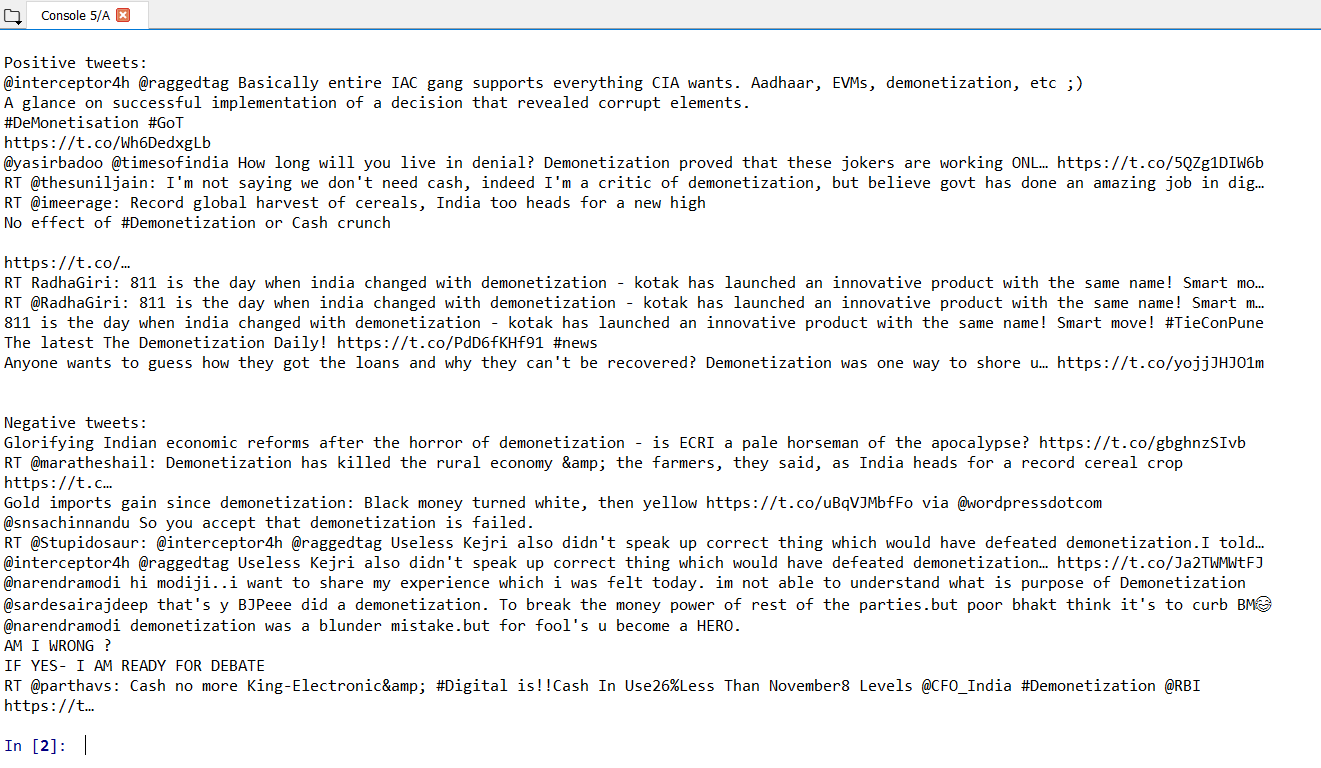
AM I WRONG ?

IF YES- I AM READY FOR DEBATE

RT @parthavs: Cash no more King-Electronic&amp; #Digital is!!Cash In Use26%Less Than November8 Levels @CFO\_India #Demonetization @RBI

**Screenshots of Result**

****

****

**Conclusion and Improvements**

Overall Success rate of approximately 55% reveals a serious demarcation between the opinion of the people on the decision of demonetization of old currency. From the technical point of view, the label of positive and negative tweet is given according to the polarity score generated. Within the positive tweets, the tweets are further classified as more or less positive. Higher the polarity score, more it has to contribute towards the success of demonetization. Similarly, within negative tweets, lower the score (Higher in modulus value) contributes more towards the failure of demonetization.

While generating polarity scores double negative words are not included. For example- if a tweet is “Demonetization is not bad for the country”, then in such case it is regarded as a negative tweet as ‘not bad’ isn’t translated into ‘good’ according to the algorithm. Thus Further analysis is to be carried out for improving the code and overcoming such ambiguity. Furthermore, for plotting the graph, currently polarity scores are rounded off to tenth place. Such rounding off gives a step nature to the curve.