

Sentiment Analysis in Twitter (French Language)

Contributed by:
Tabassum Tahera – 1413986
Md Asir Faisal - 1421837

Course: Natural Language Processing, IIUM
Instructor: Madam SURIANI BT. SULAIMAN

What is Sentiment Analysis?

- It is classification of the polarity of a given text in the document, sentence or phrase
- The goal is to determine whether the expressed opinion in the text is positive, negative or neutral.

Negative



Praval Singh @Praval · 8m

Young techies leaving Infosys in droves | Attrition rate of 18.7% - bit.ly/1kwei68

Expand



Reply



Retweet



Favorite



Buffer



Pocket



More



David Pierce @piercedavid · Apr 14

The **Galaxy S5** is a very good (and very waterproof) smartphone that left me wanting more theverge.com/2014/4/14/5608... pic.twitter.com/x5SYQ1pcZe



View photo



Reply



Retweet



Favorite



Buffer



Pocket



More

Positive

Neutral



NDTV Gadgets @NDTVGadgets · 13h

Twitter buys social data provider Gnip ndtv.in/1hI5j1Y



View summary



Reply



Retweet



Favorite



Buffer



Pocket



More

Why is Sentiment Analysis Important?

- Microblogging has become popular communication tool
- Opinion of the mass is important
 - Political party may want to know whether people support their program or not.
 - Before investing into a company, one can leverage the sentiment of the people for the company to find out where it stands.
 - A company might want find out the reviews of its products

Using Twitter for Sentiment Analysis

- Popular microblogging site
- Short Text Messages of 140 characters
- 240+ million active users
- 500 million tweets are generated everyday
- Twitter audience varies from common man to celebrities
- Users often discuss current affairs and share personal views on various subjects
- Tweets are small in length and hence unambiguous

Our Case Study

- We have chosen a specific case study. We are collecting tweets related to French Parliament Debate and its candidates.
- Our target users are the 7 candidates of that particular debate. They are : 'Sarkozy', 'Kosciusko', 'Cope', 'Juppe', 'Fillon', 'Le Maire', 'Poisson'.
- We are just focusing on the French Language. In order to do that for French Language we are importing 2 libraries which are: PatternTagger, PatternAnalyzer from TextBlob FR (For French)

Approach

Tweeter Authentication



Prepare Query Features



Get Sentiment Label of each Tweets



Retrieve Tweets and Save them



Output the result in csv files

Twitter Authentication

- Before getting any tweets, we need to set up our authentication module.
- We need 4 keys with twitter to get access. They are : consumer key, consumer secret, access token, access secret

```
# Step 1 - Authenticate
consumer_key = 'rB2HkYfxus3YJVNWd0yPKIjQf'
consumer_secret = '6dufH1dm0kVNGVw0545Xdy0ZdI5h2DiggT4IFgax02FpvVPQgE'

access_token = '784619060302585856-MIa0tQ5Hhh1zT4sJn1GmbYMwrcTYG9M'
access_token_secret = '0Nka8ycHFBIP7RUTKt08gahtf5kK0LsmamndrBcDNBsM9'

auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)

api = tweepy.API(auth)
```


Prepare Query Features

- 7 Target Users. Candidates Names.
- Hashtag related to the debate : PrimaireLeDebat
- Since Date: 2018-05-08
- Until Date: 2018-05-09

```
#List of candidates to French Republicans Primary Elections
candidates_names = ['Sarkozy', 'Kosciusko', 'Cope', 'Juppe', 'Fillon', 'Le Maire', 'Poisson']
#Hashtag related to the debate
name_of_debate = "PrimaireLeDebat"
#Date of the debate : 1st of May
since_date = "2018-05-08"
until_date = "2018-05-09"
```

Getting the sentiment label

- We are using the textblob python library to get the sentiment.
- Threshold Value > 0 = Positive, Threshold Value < 0 = Negative

```
#Step 2b - Function of labelisation of analysis
def get_label(analysis, threshold = 0):
    if analysis.sentiment[0]>threshold:
        return 'Positive'
    else:
        return 'Negative'
```

Getting Tweets and Output them

- We use api search functions to process our query features and get the relevant tweets
- Finally we output all the results in a csv file according to the candidates' name to show the sentiment result.

```
#Step 3 - Retrieve Tweets and Save Them
all_polarities = dict()
for candidate in candidates_names:
    this_candidate_polarities = []
    #Get the tweets about the debate and the candidate between the dates
    this_candidate_tweets = api.search(q=[name_of_debate, candidate], count=100, since = since_date, until=until_date)
    #Save the tweets in csv
    with open('%s_tweets.csv' % candidate, 'w') as this_candidate_file:
        this_candidate_file.write('tweet,sentiment_label\n')
        for tweet in this_candidate_tweets:
            analysis = TextBlob(tweet.text, pos_tagger=PatternTagger(), analyzer=PatternAnalyzer())
            #Get the label corresponding to the sentiment analysis
            this_candidate_polarities.append(analysis.sentiment[0])
            this_candidate_file.write('%s,%s\n' % (tweet.text.encode('utf8'), get_label(analysis)))
#Save the map for final results
```

Mean Sentiment Polarity

- We also try to get the mean sentiment polarity for each candidate.
- Mean Sentiment Polarity in descending order :
- Poisson : 0.180
- Fillon : 0.113
- Juppe : 0.098
- Sarkozy : 0.057
- Cope : 0.036
- Le Maire : 0.007
- Kosciusko : 0.007

```
sorted_analysis = sorted(all_polarities.items(), key=operator.itemgetter(1), reverse=True)
print ('Mean Sentiment Polarity in descending order :')
for candidate, polarity in sorted_analysis:
    print ('%s : %0.3f' % (candidate, polarity))
```

Outputs

Clipboard		Font		Alignment		Number		Styles		Cells		Editing							
A16		fx		RT @vandemeulebrou: #PrimaireLeDebat FormatÃ© #Juppe. Pugnace #Cope. Moderne #NKM.Combatif #Sarkozy. Surprenant #Poisson. SÃ©rieux #Fillonâ€¦															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	tweet	sentiment_label																	
2	RT @Cuad	Negative																	
3	RT @publi	Negative																	
4	RT @NRor	Negative																	
5	RT @LCI: .	Negative																	
6	RT @PVergezHonta: #PrimaireLeDebat																		
7	#JuppÃ© convaincant																		
8	#Poisson surprenant																		
9	#BLM redondant																		
10	#Sarkozy Ã©nervant																		
11	#CopÃ© dÃ©sespÃ©rÃ©																		
12	#Fillon fat Positive																		
13	RT @publi	Fillon	JuppÃ©	Le Maire	NKM	Câ€™	Negative												
14	RT @Anto Negative																		
15	RT @franc Negative																		
16	RT @vand Positive																		
17	RT @Jeun Positive																		
18	RT @PVergezHonta: #PrimaireLeDebat																		
19	#JuppÃ© convaincant																		
20	#Poisson surprenant																		
21	#BLM redondant																		

Figure : Cope Tweets

Conclusion

- We successfully implemented twitter sentiment analysis in French language in this project. This work can be extended to other languages as well.
- Through our analysis of the relevant topic, we can get valuable insights and opinions of that particular topic.