Twitter Sentiment Analysis (French Language)

Tabassum Tahera (1413986) MD Asir Faisal (1421837)
Department of Computer Science, KICT, IIUM
ttahera96@gmail.com

Abstract

In this project we tried to demonstrate the implementation of Sentiment Analysis in Twitter. We tried to find out the positive, negative and neutral sentiment polarity of tweets related to French Parliament Debate among the seven candidates from twitter.

1. Introduction

Sentiment Analysis 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. Microblogging sites are the best sources to find people opinion on different issues and matters. In this project we specifically use twitter because it is the best choice among the microblogging sites. It has short text messages of 140 characters, 240+ million active users. 500 million tweets are generated every day. Tweets are small in length and hence tend to be unambiguous. 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.

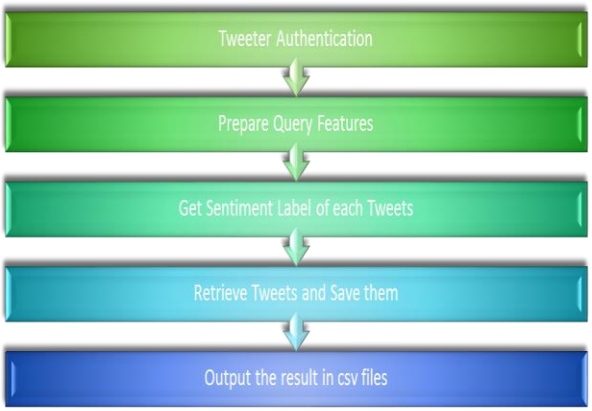
2. Literature Review

There has been lots of research done on sentiment analysis recently. In the paper named Towards Tracking Political Sentiment through Microblog Data, the researchers tried to demonstrate the changes in political sentiment related to the U.S. Supreme Court (SCOTUS) decisions. People are expressing about their feelings regarding all topics in microblogs nowadays. Based on that, the dataset of this paper is gathered form microblogs such as Twitter and it's used to show how people feel regarding political issues in US. Later on, this idea was extended to be used in a website that's meant to show people's opinions regarding similar matters such as "voting right act" and "tax cases". Their work provides an important step towards robust sentiment analysis in the political domain, and the data collected in our study is expected to serve as a stepping stone for subsequent exploration [1].

In another research paper named "Twitter Sentiment Analysis: The Good the Bad and the OMG!", the researchers tried to evaluate the usefulness of existing lexical resources as well as features that capture information about the informal and creative language used in microblogging [2]. Their experiments on twitter sentiment analysis show that part-of-speech features may not be useful for sentiment analysis in the microblogging domain. More research is needed to determine whether the POS features are just of poor quality due to the results of the tagger or whether POS features are just less useful for sentiment analysis in this domain. Features from an existing sentiment lexicon were somewhat useful in conjunction with microblogging features, but the microblogging features (i.e., the presence of intensifiers and positive/negative/neutral emoticons and abbreviations) were clearly the most useful.

3. Methodology

In order to implement our twitter sentiment analysis model, we follow few steps or procedures. The approach is given below in the diagram.



1. **Twitter Authentication**: In order to collect the tweets we first made connection with the twitter client thorugh twitter api
2. **Preparing Query Features**: To get the relevant data for our case study, we had to make custom query features where we provided the candidates name, debate name with hashtag, number of tweets to collect (100), since date and until date.
3. **Getting Sentiment Label:** We fix the threshold of the sentiment polarity. So if the polarity is greater than 0 then the label generated is positive on the other hand when it is less that 0, the label generate is negative.
4. **Retrive Tweets:** We collected 100 relevant tweets from the custom built query from our twitter client and save them.
5. **Output the result:** Finally we output the results of the tweets and their sentiment label in a separate excel file according the candidate names.
6.

4. Results

We tried to find out the negative and positive tweets related to the candidate. We also figured out the average sentiment polarity of the candidates. Mean Sentiment polarity is given in the following table in ascending order. In our project we calculate only positive tweets from the 7 candidates.

Poisson	0.180
Fillon	0.113
Juppe	0.098
Sarkozy	0.057
Cope	0.036
Le Maire	0.007
Kosciusko	0.007

5. 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.

6. References

[1] Yu Wang, Clark Tom, and Staton Jeffrey, “Towards Tracking Political Sentiment through Microblog Data”,

[2] Efthymios Kouloumpis, Wilson Theresa, Moore Johanna, Twitter Sentiment Analysis: The Good the Bad and the OMG