

SENTIMENTAL ANALYSIS ON TWITTER DATA

B.Ramyaashri

S.Srihareni

S.Subathra

Department of computing

Coimbatore institute of technology

ABSTRACT:

The aim of this project is to build technology to detect and summarize an overall sentiment and to fit trend line and time series forecasting for sentiments. Using twitter data, we build models for classifying “tweets” into positive, negative and neutral sentiment. The data for this project is collected on the basis of daily tweets that are based on the political ratings between actors Rajini and Kamal. By comparing all the tweets between Rajini and Kamal, we classify tweets into positive, negative and neutral sentiment and fit a trend line for these sentiments. By this we also infer who has the largest audience.

KEYWORDS:

Tweeter, Positive sentiment, Negative sentiment, Neutral sentiment, Trend line, Time series forecasting, Unigram model,

Feature based model, Tree kernel based model.

METHODOLOGY USED:

Sentimental analysis

DATA SOURCE:

<https://twitter.com/>

DATA DESCRIPTION:

The brief terminology associated with tweets is as follows:

- Emoticons: These are facial expressions pictorially represented using punctuation and letters; they express the user’s mood.
- Target: Users of Twitter use the “@” symbol to refer to other users on the micro blog. Referring to other users in this manner automatically alerts them.

- **Hashtags:** Users usually use hashtags to mark topics. This is primarily done to increase the visibility of their tweets.

These are the terminology that we came across in our research based on the tweets in twitter.

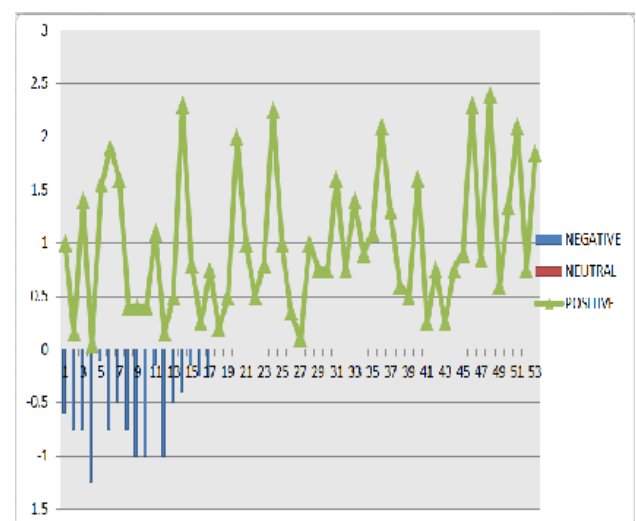
ANALYSIS:

Sentiment analysis is a type of text mining which aims to determine the opinion and subjectivity of its content. In this analysis we make use of a social media network namely TWITTER which is an online news and social networking service on which users post and interact with messages known as “tweets”. The data for this project is collected on the basis of daily tweets that are based on the political ratings between actors Rajini and Kamal. Here we have collected the information by liking all the comments that are about the politics led by Kamal and Rajini politics based tweets in twitter. Then we have extracted the twitter data into R by using the packages namely “twitterR” which helps us to extract the data from twitter and the package “roauth” which helps us to establish a connection among twitter and R which will be necessary for extracting data for analysis. We collect data separately for politics led by Kamal and politics led by

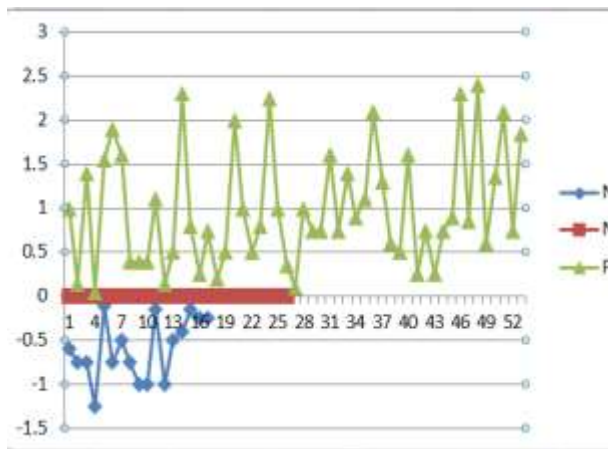
Rajini. Thereby after extracting the data into R we cleaned the data that is cleaning the data is to remove unnecessary symbols or characters to perform sentimental analysis. We experiment with three types of models: unigram model, a feature based model and a tree kernel based model. By comparing all the tweets between Rajini and Kamal, we classify tweets into positive, negative and neutral sentiment. Using this classification, we fit a trend line and time series forecasting for these sentiments. In this project, time series analysis helps us to predict who is going to rule this state in future. By this we also infer who has the largest audience.

RESULT:

PLOT FOR KAMAL POLITICS:



PLOT FOR RAJINI POLITICS:



SENTIMENT ANALYSIS FOR KAMAL POLITICS:

R OUTPUT:

```
Negative  Neutral  Positive
      17       26       53
```

SENTIMENT ANALYSIS FOR RAJINI POLITICS:

R OUTPUT:

```
Negative  Neutral  Positive
      15       32       57
```

CONCLUSION:

By applying sentimental analysis, we conclude that Rajini has the largest audience when compared to Kamal. By comparing the positive and negative sentiments between Rajini and Kamal, we infer that Rajini is going to get majority votes.

REFERENCE:

- <https://dataaspirant.com/2018/03/22/twitter-sentiment-analysis-using-r/>
- <https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>
- https://medium.com/@rohitna ir_94843/analysis-of-twitter-data-using-r-part-3-sentiment-analysis-53d0e5359cb8
- <https://github.com/jeffreybreen/twitter-sentiment-analysis-tutorial-201107/blob/master/data/opinion-lexicon-English/positive-words.txt>
- <http://rstatistics.net/extracting-tweets-with-r/>
- Bo Pang (Author), Lillian Lee (Author), Opinion mining and sentiment analysis
- Erik Cambria (author), Dipankar Das Sivaji Bandyopadhyay (author) Antonio Feraco(author), A practical guide for sentiment analysis, volume-5
- James sanger ,The Text Mining Handbook, Ronen Feldman,

- Alexander Pak and Patrick Paroubek. 2010. Twitter as a corpus for sentiment analysis and opinion mining. Proceedings of LREC.
- B. Pang and L. Lee. 2004. A sentimental education: Sentiment analysis using subjectivity analysis using subjectivity summarization based on minimum cuts. ACL.