Imperial Journal of Interdisciplinary Research (IJIR)
Vol-2, Issue-7, 2016
ISSN: 2454-1362, http://www.onlinejournal.in

An Implementation the Concept of Analysis of Microblogging with Twitter for Evaluating the Branding Trends

Anil Ahir[1] & Santosh Tamboli[2]
[1]Mumbai University, Vidyalankar Institute of Technology, Mumbai , India
[2]Assistant Professor, Vidyalankar Institute of Technology, Mumbai , India

***Abstract**: We here proposing an automated method of classifying the sentiments in the tweets. We examine the tweets structure and arrangement and we examined the sentiment extract feasibility in our theoretical paper [1]. In this particular paper we shall be putting into practise the theory that we have proposed. The contributions of this paper are: (1) We introduce -specific prior polarity features. (2) We explore the use of a simple classification model. And to draw a conclusion that the parsing of the natural language.*

***Keywords**: Branding, Tweets, polarity, Microblogging, emoticons.*

1. Introduction

Microblogging websites have evolved as a customer service portal which people have started using in the daily life Identifying the pain area of a consumer is critical for the timely acknowledgement of events, like failures in service or defamation that can affect the brands performance and image.
In this paper we look at the implementation aspect of Twitter as a platform for checking peoples sentiments like positive and negative. We had proposed a classification tasks: a binary task of classifying sentiment into positive and negative classes. We use manually annotated Twitter data for our experiments. A benefit of this data, over earlier used data-sets, is that the tweets are together in a flooding style and therefore represent a exact sample of real tweets in terms of language use and content. [1] [2]

2. Literature Survey

A number of studies have focused on developing sentiment analysis detection system. In provides a good characterization of different types of sentiments expressed by the users of the particular brand on Twitter. I have tried to introduce a general framework that aims to simplify the sentiments expressed on Twitter. They are successful in identifying sentiments that are generated by the users.
Keyword/pattern-based techniques are different from other techniques with respect to the measurement metric. The metrics used in a keyword/pattern-based technique is extracted from the tweet, and commonly includes phrases, abbreviations, smileys or as they are called emoticons/ emojis. Thus the total extraction of the information from the data is more accurate this gives the user a better vantage to make a call by seeing the information and its flow.
This vantage for the user is the key towards survival in the social media world as the word about a brand and its image spreads like wild fire and can wreak havoc or create a considerable damage which is irreversible to the brand entity.
Another involvement of this proposal is that we report fallouts on manually explained data that does not undergo from any known biases. Our data is a random sample of streaming tweets unlike data collected by using specific queries. The size of our hand labeled data allows us to perform cross-validation experiments and check for the variance in performance of the classifier across folds.

2.1 Data Description

Twitter is a social networking and microblogging service that allows users to post real time messages, called tweets. Tweets are short messages, restricted to 140 characters in length. Due to the nature of this microblogging service (quick and short messages), people use acronyms, make spelling mistakes, use emoticons and other characters that express special meanings. Following is a brief terminology associated with tweets. 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 microblog. 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.[3]

3. Resources and Pre-processing of data

In this implementation we two resources for pre-processing twitter data, first an emoticon dictionary and Second an acronym dictionary. We prepare the emoticon dictionary by labelling 170 emoticons listed on Wikipedia[1] with their emotional state. For example, ":)" is labelled as positive whereas ":=(" is labelled as negative. We assign each emoticon a label from the following set of labels: Extremely-positive, Extremely-negative, Positive, Negative, and Neutral. We compile an acronym dictionary from an on-line resource. The dictionary has translations for 5,184 acronyms. For example, lol is translated to laughing out loud.

We pre-process all the tweets as follows:

a) Replace all the emoticons with a their sentiment polarity by looking up the emoticon dictionary,

b) Replace all negations (e.g. not, no, never, n't, cannot) by tag "NOT", and

c) Replace a sequence of repeated characters by three characters, for example, convert coooooooool to cool. We do not replace the sequence by only two characters since we want to differentiate between the regular usage and emphasized usage of the word [4]

Acronym	English expansion
gr8, gr8t	great
lol	laughing out loud
rotf	rolling on the floor
bff	best friend forever

Table 1: Example acronym and their expansion in the acronym dictionary.

Emoticon	Polarity
:-) :) :o) :] :3 :c)	Positive
:D C:	Extremely-Positive
:-(:(:c :[	Negative
D8 D; D= DX v.v	Extremely-Negative
: j	Neutral

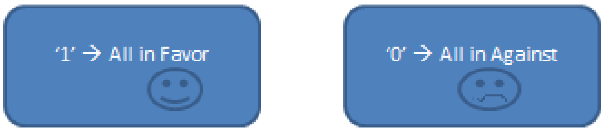
Table 2: Part of the dictionary of emoticons

4. Polarity

A number of our features are based on prior polarity of words. We use Dictionary of Affect in Language (DAL) (Whissel, 1989) and extend it using Dictionary.

We then assign the polarity 1 for Positive and 0 for negative.

Thus the polarity assigned basis the words used in the tweets shall give out the picture as to how the brand is being received. [4]



5. Implementation

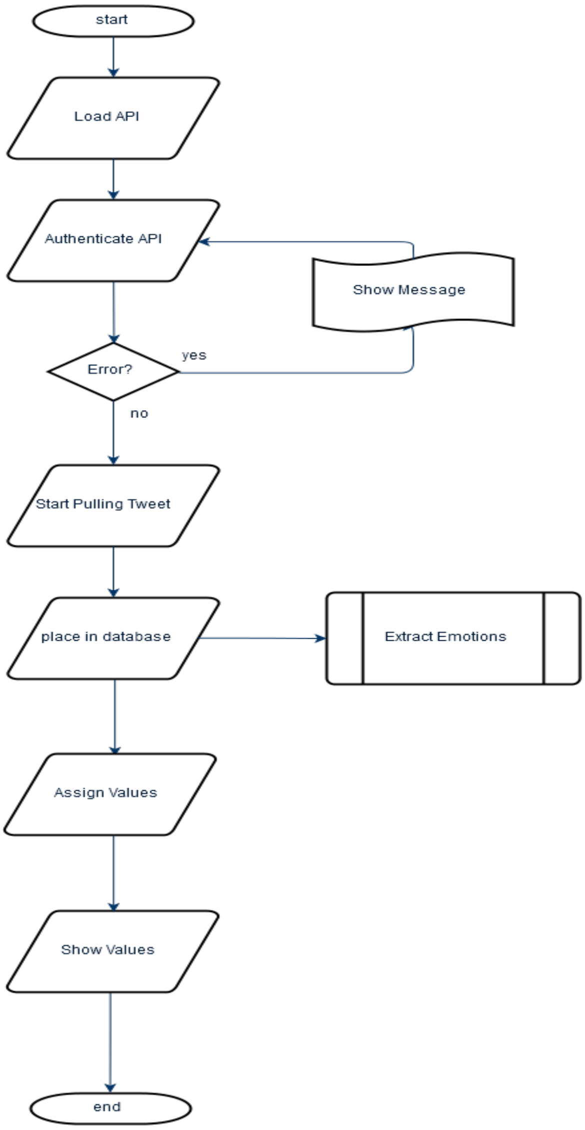


Figure 5.1: Activity Diagram of Implementation

The data flow throughout the system is shown below in the Data Flow Diagram Where we can see from user end User pass its Credential to the system as an authentication process to the system. The system connects with Twitter Website to pull authorization details.

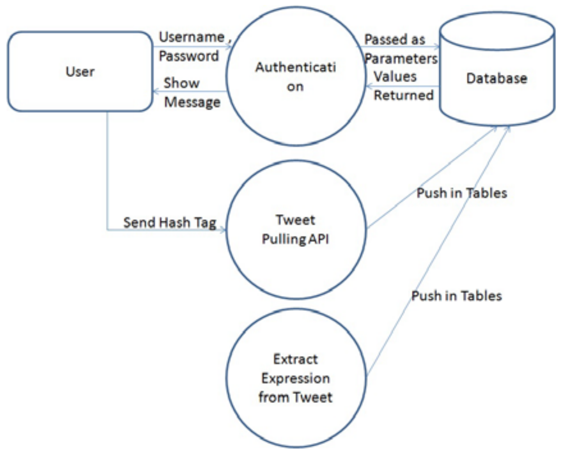


Figure 5.2 Data Flow Diagram

6. The Result

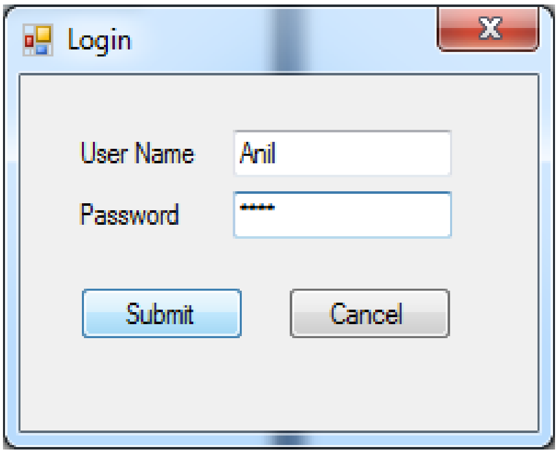


Fig 6.1: Login Screen

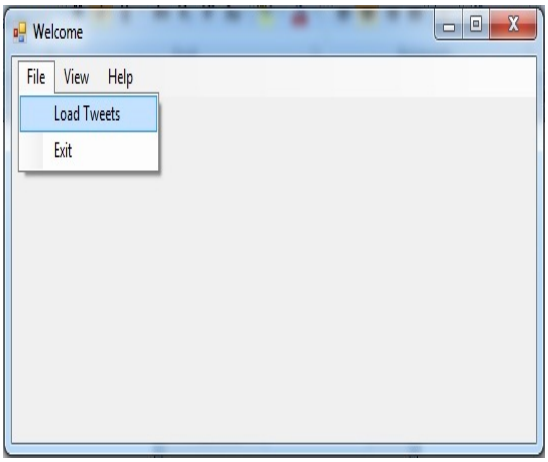


Fig 6.2: Main Screen

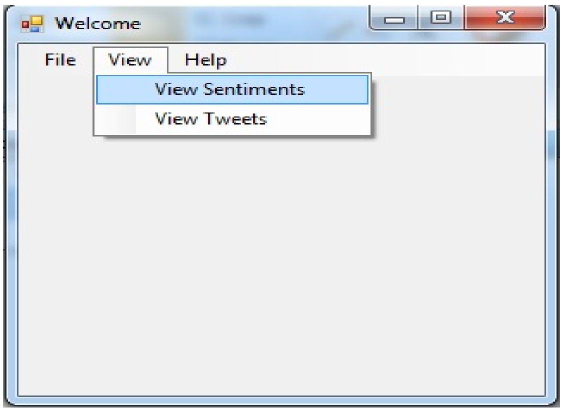


Fig 6.3 View Sentiments Menu

Fig 6.4 Load Tweets

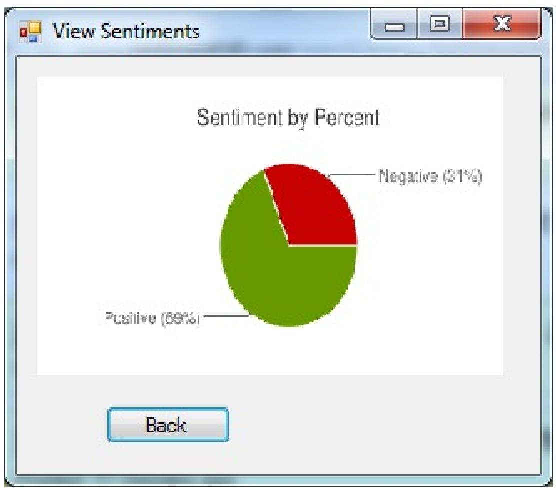


Fig 6.5 View Tweets Sentiments

7. Conclusion

The Expression Extraction Algorithm system extracts

the expression in the large number of tweets. It first collects all the tweets and extracts expressions from it. From these extracted expressions the values is assigned. These values provide a Meta data relevant to a brands current image/ campaign.
This is the analysis that has been done on closed features data sets of tweets, showing how sentiments flow in the general user space.

8. References

[1] Anil Ahir, Santosh Tamboli," AN APPROACH TOWARDS ANALYSIS OF MICROBLOGGING WITH TWITTER FOR EVALUATING THE BRANDING TRENDS", International Journal of Science and Research", IJSR ISSN 2319-7064, Volume 5, Issue 6, JUNE 2015.

[2] Sandip Mali, Ashish Baleroa," Sentiment of Sentence in Tweets: A Review", IOSR Journal of Computer Engineering, e-SSN: 2278-0661, Volume 17, Issue 6.

[3] https://simple.wikipedia.org/wiki/Twitter

[4] Apoorv Agarwal, Fadi Biadsy, and Kathleen Mckeown. 2009. Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams. Proceed-ings of the 12th Conference of the European Chapter of the ACL (EACL 2009), pages 24–32, March.

[5] Luciano Barbosa and Junlan Feng. 2010. Robust senti-ment detection on twitter from biased and noisy data.
Proceedings of the 23rd International Conference on Computational Linguistics: Posters, pages 36–44.

[6] Adam Bermingham and Alan Smeaton. 2010. Classify-ing sentiment in microblogs: is brevity an advantage is brevity an advantage? ACM, pages 1833–1836.

[7] C. Fellbaum. 1998. Wordnet, an electronic lexical database. MIT Press.

[8] Michael Gamon. 2004. Sentiment classification on cus-tomer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. Proceedings of the 20th international conference on Computational Lin-guistics.

[9] Alec Go, Richa Bhayani, and Lei Huang. 2009. Twit-ter sentiment classification using distant supervision. Technical report, Stanford.

[10] David Haussler. 1999. Convolution kernels on discrete structures. Technical report, University of California at Santa Cruz.

[11] M Hu and B Liu. 2004. Mining and summarizing cus-tomer reviews. KDD.

[12] S M Kim and E Hovy. 2004. Determining the sentiment of opinions. Coling.

[13] Dan Klein and Christopher D. Manning. 2003. Accurate unlexicalized parsing. Proceedings of the 41st Meet-ing of the Association for Computational Linguistics, pages 423–430.

[14] Alessandro Moschitti. 2006. Efficient convolution ker-nels for dependency and constituent syntactic trees. In

Proceedings of the 17th European Conference on Ma-chine Learning.

[15] Alexander Pak and Patrick Paroubek. 2010. Twitter as a corpus for sentiment analysis and opinion mining.
Proceedings of LREC.

[16] Pang and L. Lee. 2004. A sentimental education: Sen-timent analysis using subjectivity analysis using sub-jectivity summarization based on minimum cuts. ACL.

[17] Turney. 2002. Thumbs up or thumbs down? seman-tic orientation applied to unsupervised classification of reviews. ACL.

[18] C M Whissel. 1989. The dictionary of Affect in Lan-guage. Emotion: theory research and experience, Acad press London.

[19] Wilson, J. Wiebe, and P. Hoffman. 2005. Recognizing contextual polarity in phrase level sentiment analysis.
ACL.