ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

Analyzing Public Sentiment Variations on Twitter and Facebook

[I]Ashwini Patil, [II]Sanket Thakare
[I]Assistant Professor, Thakur College of Engineering and Technology, Mumbai, India
[II]Software Developer, Ebuild Softwares and Solutions, Mumbai, India.

Abstract
Nowadays, Sentiment analysis and opinion mining have become one of the utmost developing fields of research and fascinated attention of researchers, analysts and industrialists. Millions of users, per day, exchange their views, ideas, expressions, feelings and opinions on social networking sites like Twitter and Facebook. The organisations can use this vast textual information for analyzing people's opinion about a specific target. Opinion of people about a specific target may change with the time. Most of the existing research work mainly focused on classifying and predicting public sentiments on Twitter. In this work, we have analyzed the public sentiment variations in a specific time period about a specific target on Twitter and Facebook both. One can do further analysis to know the useful insights causing public sentiment variations. This kind of analysis is helpful in various fields for taking proper decisions and reverting public opinion.

General Terms
Sentiment Analysis, Opinion Mining

Keywords
Supervised machine learning, Public Sentiment Variation

I. Introduction

With the intensive development of Web applications such as blogs, forums, and social networking sites, there is growth in user generated content in the form of posts, reviews, ratings, recommendations and feedbacks about the products. This vast textual information can be about anything including politicians, products, people, events, etc. This huge textual content available on web is needed by companies, web practitioners, researchers, politicians, service providers and analysts to mine and analyze this for different uses. With this vast user generated content, there is need of automated techniques for analyzing and mining, as manual analysis is very difficult for such a huge content. Sentiment analysis or opinion mining is the automated and computational study of people's opinion, expression, feelings, thoughts, and emotions towards a particular target. The term opinion mining deals with extracting and analyzing people's opinion about any entity, while the term sentiment analysis aims to automate the task of finding sentiment and classifying it into sentiment polarity. Nowadays, every individual wants to know other's opinion before purchasing any product. The organizations and businesses want to find their consumers feedback about their product or services. The organizations nowadays, are not dependent on surveys, opinion polls and focus groups because of the vast amount of data available publicly for analysis.

Per day, millions of users are sharing their opinions/feelings on social media like Twitter and Facebook. But the opinions of people are bounded for the particular time period. With time people's opinion may change about a specific target. This can be called as 'Public Sentiment Variation'. Much of the previous work, in this research field mainly focuses on predicting and classifying public sentiment on microblogging services. So, anyone can do the valuable analysis to find the variations in public sentiments according to the time. This kind of analysis can be useful for taking corrective actions, for example, if public sentiment towards a particular product changes suddenly, the company can make the proper decisions to revert the public sentiment. This type of analysis can be useful for further analysis to mine the possible reasons behind such variations. This type of work is done in [12].

Basically, there are two approaches for sentiment analysis 1) Machine learning based and 2) Lexicon based. The machine learning based approach uses various supervised and unsupervised learning algorithms like Naïve Bayes, Maximum Entropy, and Support Vector Machines for sentiment classification. The Lexicon based techniques use a lexicon dictionary with sentiment words related to specific domain for sentiment classification. Some researchers combined the machine learning and lexicon based techniques [7, 10]. According to literature survey, both these techniques have their own pros and cons, but still the supervised machine learning techniques give more accuracy of sentiment classification than lexicon based techniques [13].

In this work, we have analyzed public sentiment variations on Twitter and Facebook about the specific targets like "IPhone" and "Kindle". To track public sentiment variations, we have classified the tweets and posts regarding these two targets into positive, negative and neutral. For this we used Stanford NLP's Sentiment Analysis Tool, which is based on Maximum Entropy Classifier [14].

II. Related Work

Many efforts have been taken to automate the task of sentiment analysis in last few years. Different Machine learning and lexicon based techniques have been proposed for sentiment classification of online text data.

A. Machine Learning Based Approaches

The machine learning techniques became popular for sentiment analysis from the work done by Pang and Lee [1]. Three Supervised Classifiers: Naive Bayes (NB), Maximum Entropy (ME) and Support Vector Machines (SVM) were experimented on movie reviews collected from imdb.com. They compared the efficiency of these three classifiers with different feature selection method such as uni-gram, n-gram, combining uni-gram and bi-gram and by combining uni-gram and POS tagging. They reached a conclusion that if the feature set is small then it is better to consider feature presence than feature frequency. While Naive Bayes performed well on small feature set, SVM performed well on large feature space. When there was a large feature space, Maximum Entropy performed better than Naive Bayes. SVM showed the best performance with an average precision of 82.9%. Their work was extended to improve the precision of NB [2]. They proposed subjectivity identification in the text as a pre-processing step and best accuracy up to 87.2% was achieved on the movie reviews dataset. The experiments with SVM, NB and ME on movie and car brand reviews by Boiy, E., Hens, P., Deschacht, K., and Moens, have also confirmed high performance of machine learning techniques with accuracy up to 90.25% [3]. N-gram modelling approach with Naive Bayes was studied for subjectivity and polarity analysis at document level on legal weblogs and movie reviews, by Cornad and Schilder [4].

B. Lexicon Based Approach

An unsupervised learning algorithm for classifying reviews was proposed by Turney, in 2002 [5]. In that approach, first, Adjectives and adverbs were extracted and then semantic orientation of extracted phrases was calculated using Pointwise Mutual Information (PMI). Finally reviews were classified as per average semantic orientation of the phrases. This approach was able to achieve 66% accuracy for the movie review domain. The experiments are conducted on different datasets like movie, bank and automobile. Positive and negative sentiment based summaries for product features from reviews (Amazon, CNET) were proposed by Hu and Liu (2004). Opinionated sentences were identified using dictionary of adjectives (sentiment lexicon). They used 30 adjectives as seed list. They achieved 80% accuracy of the polarity classification in terms of precision [6]. In an attempt to assign sentiment scores to each distinct entity in the text and then assigning an overall subjectivity score to the text, Godbole et al. (2007) proposed a sentiment lexicon based semantic approach [8]. Pang and Lee (2008), in their study gives the overview of the different approaches used in sentiment analysis. Their study concludes that in some unsupervised learning approaches, a sentiment lexicon is generated and later used to determine the text unit's degree of positivity or subjectivity. Creating the sentiment lexicon through unsupervised polarity or subjectivity labelling of words or phrases is crucial [9].

III. Proposed Methodology

In this work, for analyzing public sentiment variations, we have performed following three steps. First we extracted the tweets and Facebook posts regarding specific targets (e. g. "IPhone", "Kindle" etc.) from the collected tweets and posts. The extracted tweets and posts are pre-processed to remove the noise in the data. The task of preprocessing is essential to make the sentiment analysis more efficient. Second, we classified these tweets/posts into three classes like positive, negative and neutral by using Stanford NLP Classifier [14]. Finally, on the basis of classification of each tweet and post, the sentiment variations regarding the specific targets have been tracked with the help of descriptive statistics. These steps are described in the following subsections.

1. Extracting and Processing Sentiment Words

We extracted all the tweets and posts related to the specific targets. To extract the related tweets/posts, from the whole dataset, we filtered the tweets/posts which contain the keywords of the target. Generally noise in the tweets and posts affects the results of sentiment analysis. The preprocessing techniques remove the noise and make them efficient for sentiment classification. For preprocessing we have applied following techniques.

1. Stop word removal: Extremely common words are not considered in order to speed up results. These filtered words are known as 'Stop Words'. We maintained a list of commonly used stop words and it is used to remove all the stop words from tweets/posts.
2. Stemming: Stemming is done for reducing inflected (or sometimes derived) words to their stem, base or root form. For example the word "sadness", indicate the negative sentiment. To extract the correct sentiment from these type of words. They should be stemmed to its root i.e. the word "sad". A popular Porter stemmer algorithm is used for stemming process [15].
3. Conversion of slang words: User Tweets and posts often contain slang words. The slang words like omg, lol play important role in sentiment analysis. We converted commonly used slang words into their standard forms and then used them to analyze the sentiments.
4. Removal of URL: Most of the tweets and posts contain URLs. These URLs should be removed to make the task of sentiment analysis more effective.

2. Classifying Sentiments

To classify the tweets and posts as positive, negative and neutral, we used sentiment analysis tool of Stanford NLP framework [14]. This tool uses Maximum Entropy Classifier which classifies the tweets/posts into positive, negative and neutral classes. This Stanford NLP tool has proven its quality and accuracy, as various sentiment processing applications has adopted it as a popular sentiment analysis tool in machine learning. Thus we can estimate the probability that an opinion may contain positive, negative or neutral sentiments.

3. Analyzing Sentiment Variation

After classifying all the extracted tweets and posts about a target, we can analyze the sentiment variations using the number of positive and negative tweets/posts. But the number of positive and negative tweets is not so useful, as this number may change consistently. In this work, we considered percentage of positive, negative and neutral tweets/posts to analyze the sentiment variations, e.g. the percentage of negative/positive tweets/posts is increasing more than 50%. The statistics obtained after classification is the useful indicator for sentiment variation over the period of time.

IV. Performance Analysis

The performance of sentiment analysis techniques can be evaluated by using different performance metrics like overall accuracy, precision, recall and F1 score. The values for these metrics can be obtained by using the following confusion matrix [7].

Table 2 : Confusion matrix

	Predicted positives	Predicted negatives
Actual positive examples	Total True Positive examples (TP)	Total False Negative examples (FN)
Actual Negative examples	Total False Positive examples (FP)	Total True Negative examples (TN)

The performance of sentiment classification is evaluated by the Overall Accuracy, which is given by

Overall Accuracy= TP+TN/TP+FP+TN+FN …… (1)

Another popular evaluation metrics are:
PRECISION is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. It is usually expressed as a percentage.
Precision= TP/TP+FN ……… (2)
RECALL is the ratio of the number of relevant records retrieved to the total number of relevant records in the database. It is usually expressed as a percentage.
Recall = TP/TP+FP ……….. (3)

1. Datasets

Twitter Dataset: We obtained Twitter dataset from http://help.sentiment140.com/for-students. This dataset consists of around 1.5 million tweets. We have tested our system on the subset of the dataset, which covers around 3000 tweets on IPhone.
Facebook Dataset: We collected Facebook dataset from http://cucis.ece.northwestern.edu/projects/Social. It covers 1000 user posts on Kindle. This dataset is used in the paper [11].

V. Results and Discussions

We tested the work on around 3000 Twitter tweets and 1000 Facebook posts. Twitter Dataset and Facebook Dataset are processed in two steps: Pre-processing and Sentiment Classification. In Pre-processing stage, all the stop words are removed and remaining words are brought to their base forms using stemming. In sentiment classification the entire tweet's and post's sentiment label is decided. The tweets and posts have been manually assigned the dates and time. We considered a time window of 10 days for tracking variations in the public sentiments. We analyzed the percentage of positive, negative and neutral tweets/posts for each time window. The figure 1 and 2 show the public sentiment variations for "IPhone" on Twitter and "Kindle" on Facebook respectively.

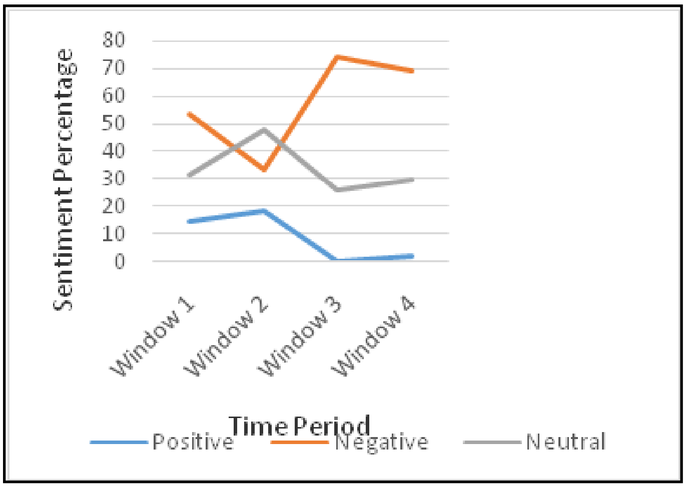

Fig 1 : Public Sentiment Variation for "IPhone" on Twitter

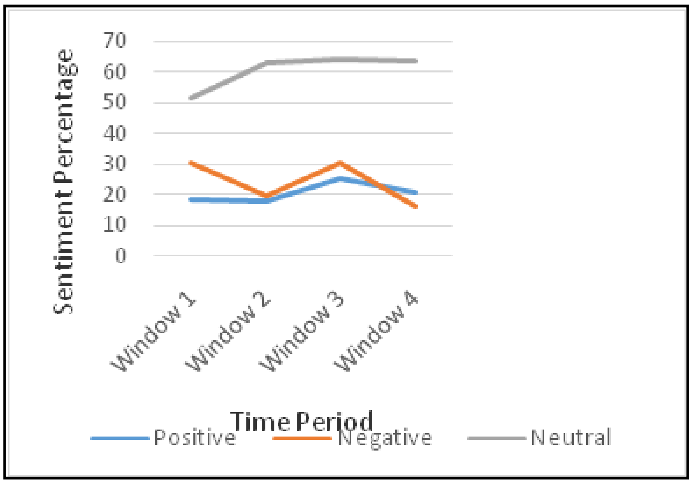

Fig. 2 : Public Sentiment Variation for "Kindle" on Facebook

To evaluate accuracy of the proposed work, we manually labelled the Twitter tweets and Facebook posts as positive, negative and neutral. The accuracy of the system is measured against the ground truth of the sentiment labels for each tweet/post. The Overall accuracy is measured in terms of precision and recall. The precision and Recall values are evaluated by using equations 2 and 3, on different time windows. The number of tweets/posts varied for each time window. We compared the performance of our approach on Twitter and Facebook. The following graph shows the precision and recall curves for Twitter and Facebook datasets. The system has achieved 53.96% accuracy on Twitter dataset and 63.16% accuracy on Facebook dataset.

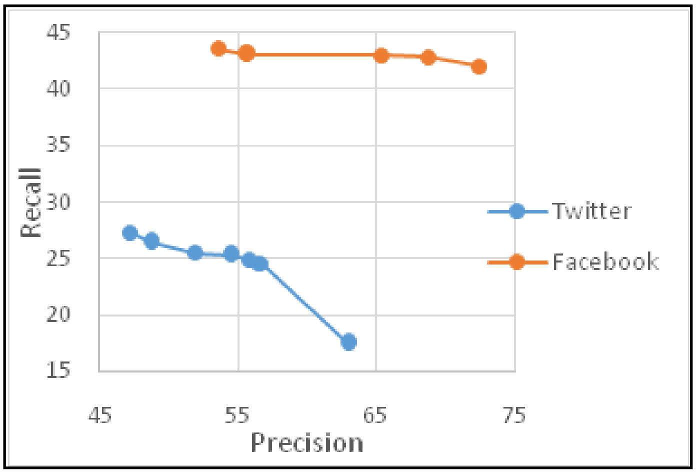

Fig. 3 : Precision and Recall Curves for Our Approach on Twitter and Facebook Datasets

VI. Conclusion

This paper focuses on analyzing the public sentiment variations on Twitter and Facebook. To track the public sentiment, we used Stanford NLP's sentiment analysis tool, which classifies the tweets/posts into positive, negative and neutral classes. On the basis of classification, the public sentiment variations have been tracked with the help of some descriptive statistics. We compared the performance of our approach on two different platforms Twitter and Facebook, with two targets, "IPhone" and "Kindle". Our experimental results show that our approach is able to achieve 53.96% accuracy on Twitter and 63.16% accuracy on Facebook. Our Twitter dataset contains around 3000 tweets, while Facebook dataset contains 1000 posts. We conclude that our approach works better with small size dataset, irrespective of character limitation on Twitter, which is 140 characters. For Facebook there is no such limitation on characters for a single post. Previous research studies focused on Twitter for analysing public sentiment variations. In this paper we tried to analyse public sentiment variations on Facebook also.

VII. Future Scope

After analyzing the public sentiment variations about a specific target, we can do further analysis to mine the possible reasons behind these variations. The public sentiment variation is useful indicator for knowing the possible reasons behind the variations in particular time period.

References

[1] *B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," Proceedings of the ACL-02 conference on Empirical methods in natural language processing, vol.10, 2002, pp. 79-86.*

[2] *Pang B., Lee, L., "A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts",Proc. Of 42nd annual meeting of the association for Computational linguistics(ACL), 2004, 271-278*

[3] *Boiy, E., Hens, P., Deschacht, K., and Moens, M.F. "Automatic sentiment analysis in on-line text" Proceedings Of 11th International Conference on Electronic Publishing(Vienna, Ausrtria) 2007*

[4] *Cornard, J.G., and Schilder, F. "Opinion mining in legal blogs" Proc. Of 11th International Conference on Artificial Intelligence and Law (ICAiL'07), ACM,2007, 231-236*

[5] *P. Turney, "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews", Proceedings of the Association for Computational Linguistics (ACL), 2002, pp. 417–424.*

[6] *M. Hu and B. Liu, "Mining and summarizing customer reviews," Proceedings of the tenth ACM international conference on Knowledge discovery and data mining, Seattle, 2004, pp. 168-177.*

[7] *Anuj Sharma, ShubhamoyDey "Artificial Neural Network Based Approach for Sentiment Analysis of Opinionated Text", ACM transaction, 2012*

[8] *Godbole,N., Srinivasaiah, M.Skiena S. "Large scale sentiment analysis for news and blogs", Proceedings of the International Conference on Weblogs and Social media (ICWSm'07), 2007*

[9] *B. Pang and L. Lee, "Opinion mining and sentiment analysis," Foundations and Trends in Information Retrieval 2(1-2), 2008, pp. 1–115.*

[10] *A. Mudinas, D. Zhang, M. Levene, "Combining lexicon and learning based approaches for conceptlevel sentiment analysis", Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining, ACM, New York,NY, USA, Article 5, pp. 1-8, 2012*

[11] *Kunpeng Zhang, Yu Cheng, YushengXie, Ankit Agrawal, Diana Palsetia, Kathy Lee, and AlokChoudhary, "SES: Sentiment Elicitation System for Social Media Data", ICDM-SENTIRE 2011*

[12] *Shulong Tan, Yang Li, Huan Sun, Ziyu Guan, XifengYan, Jiajun Bu, , Chun Chen, , and XiaofeiHe,IEEE Transactions on "Interpreting the Public Sentiment Variations on Twitter," VOL. 26, NO. 5, MAY 2014.*

[13] *Ashwini Patil, Shiwani Gupta, "A Review on Sentiment Analysis Approaches" Multicon-W 2015 (ICWCCV 2015), vol 1, pp.243-247, ISBN: 978-0-9884925-7-8, February 2015*

[14] *http://nlp.stanford.edu/software*

[15] *M.F. Porter, "An algorithm for suffix stripping", Program, Vol. 40, Iss: 3, 2006, pp.211 - 218*

Author's Profile

Ashwini Patil, *Received the Master's degree(M.E.) in Computer Engineering in 2016 from Thakur College of Engineering, Mumbai. She is currently working as Assistant Professor in department of Computer Engineering, TCET, Mumbai. Her area of specialization is Data Mining and Sentiment Analysis.*
Email-Id: patil.ashwini.03@gmail.com

SanketThakare, *Received the Master's degree(M.E.) in Computer Engineering in 2016 from Thakur College of Engineering, Mumbai. He is currently working as a Software Developer in Ebuild Software Solutions, Mumbai.*
Email-Id: sanketmthakare@gmail.com