Scientific Journal of Impact Factor (SJIF): 3.134

ISSN(Online): 2348-4470
ISSN(Print) : 2348-6406

International Journal of Advance Engineering and Research Development

Volume 1, Issue 11, November -2014

SURVEY PAPER ON ANALYZING SENTIMENT OF PUBLIC ON TWITTER

Chandrakant .M.Jadhav[1],Ashwini P.Patil[2], Gurunath G.Machhale[3]

[1]*Department of Computer Science & Engineering, BIGCE College Solapur,chandaronly@gmail.com*
[1]*Department of Computer Science & Engineering, BIGCE College Solapur, ash15april@gmail.com*
[3]*Department of Information Technology, SCOE College Pune , gurunath.machhale@gmail.com*

Abstract *—An important part of information gathering behavior is to find out what other people think. With the growing availability of opinion-rich resources. Opinion resources are online review sites and personal blogs. This are new opportunities and challenges arise as people can do actively use information technologies to understand the opinions of others. Today micro blogging has become a popular communication tool among Internet users. Everyday millions of user share their opinions on different aspects of life. For sentiment analysis and opinion mining micro blogging web-sites provides raw data. In our paper, we focus on twitter posts or remarks given by users. We collects data for sentiment analysis, then we determine the positive, negative and neutral sentiment for a product given by the user.*

Keywords*-Opinion mining, Sentiment analysis, Opinion extraction, Sentiment polarity, Sentiment classification.*

I. INTRODUCTION

There has been a lot of researcher work on opinion mining and sentiment analysis. Opinions are central to all human activities. Sentiments, evaluations, emotions and attitudes. this are the opinion mining related concepts. Sentiment analysis is also called as opinion mining. Opinion mining is the field of analyzes people opinions, sentiments and emotions towards entities such as product, services, events, topics and related attributes. Sentiment analysis or opinion mining contains many tasks such as opinion extraction, sentimentmining, reviewing and subjectivity analysis.

The term sentiment analysis or opinion mining is commonly used in industry area, but in academia these term are frequently employed. Opinion is a person thinking or perspective about an object. Mining is refers to extraction of knowledge, information from raw data available on internet [1].The increasing use of internet by people for searching information about various products, latest launched product etc. People are also write their opinion and comments on social media .Survey of such opinion also affect the remarks related to a product so that they can improve the quality of product based upon the remarks of people.

Opinion mining and sentiment analysis are used in business to analyze their product remarks. Remarks helps to take effective measures for positive remarks. Sentiment analysis has been divided into three levels.

A. Document Level: At this level to classify a whole document that will express opinion is positive or negative sentiment .For example, product review expresses an overall positive or negative opinion about product .this is commonly referred as document level sentiment classification. In this level analysis assumes that each document review express opinions on a single product. This level is not applicable to documents which consider multiple entities.

B. Sentence Level: Each sentence express a positive ,negative or neutral opinion. In this level of analysis determine subjective sentences that express subjective views and opinions.

C. Entity and Aspect Level: Document level and sentence level analyses do not show exactly what people liked and did not like. Aspect level analysis is based on the idea that opinion consists of a positive or negative sentiment. The importance of opinions helps to understand the sentiment analysis. In this level of analysis is to consider sentiment on entities or aspects. Based on this result, a structural summary of opinions about entities can be produced and their aspects are discovered.

There are two types of opinions.

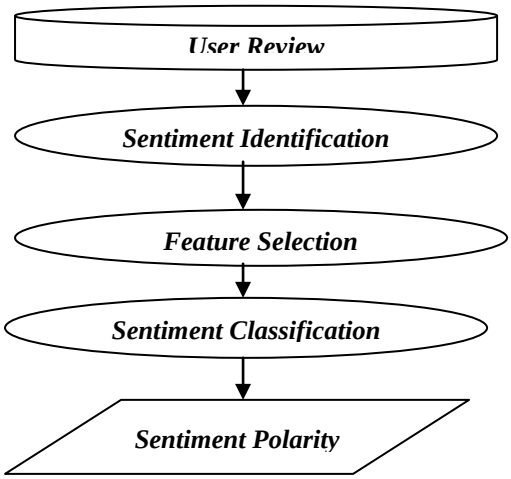
A. Regular Opinions : A regular opinion express only sentiment on an single entity or an aspects of the single entity.

B. Comparative Opinion: A comparative opinion considers multiple entities .It compares multiple entities based on their aspects.

II. LITERATURE SURVEY

Many algorithm have been proposed to implement opinion mining and sentiment analysis. Various tools are available for opinion extraction, sentiment analysis. Researchers developed models for measuring polarity of words and sentences [2]. Some researchers stated that opinion mining and sentiment analysis have slightly different notions [3]. Opinion Mining extracts people's opinion about an entity and analyzes opinion about an entity. Sentiment Analysis identifies the sentiment expressed in a text. The target of Sentiment Analysis is to find opinions, identify the sentiments and then classify their polarity they have expressed.

Sentiment analysis classification process as illustrated in Figure 1. Document-level, sentence-level, and aspect-level sentiment analysis as shown in Figure 1. A Document-level analysis classify an opinion document. It considers the whole document a basic information unit or raw data. Sentence-level analysis classify sentiment expressed in each sentence. The first step in sentiment analysis is to identify whether the sentence is subjective or objective. If the sentence is subjective then Sentence-level analysis will determine whether the sentence expresses positive or negative opinions about entity [4] .Aspect-level identify the entities and their aspects. The opinion holders can give different opinions for different aspects of the same entity or product.



"Figure 1. Sentiment analysis process on product reviews"

The data sets are used in sentiment analysis. The main sources of raw data are taken from the product reviews. These reviews are analyzed. These analyzed result are important to the business holders . Business holders can take business decisions according to the analysis results of users 'opinions about their products. The reviews sources are mainly review sites that provide users opinion . In political debates for example, we could find out people's opinions on a certain political parties. The election results can also be predicted from posts. The social network sites sites are a very good source of information because people share and discuss their opinions about a certain topic freely. They are also used as data sources in the sentiment analysis process.

We use a dataset formed of collected short messages from Twitter. Twitter contains a large number of very short messages are written by the users .The contents of the messages are different from personal thoughts to public statements. Table 1. shows examples of typical posts taken from Twitter. As the audience, user of micro blogging platforms and its services grows every day. Data from these sources can be used in opinion mining and sentiment analysis tasks. For example, manufacturing companies may be interested in the following questions.

A. What do people think about our product quality, its services, company?
B. How positive or negative are people about our product?
C. What would people prefer our product to be like?

Social organizations may ask people's opinion on current debates. All this information can be obtained from microblogging services. Users post everyday what they like, what they dislike, and their opinions on many aspects of their life. Microblogging can be used for sentiment analysis purposes. Microblogging platforms are used by different people to express their opinion about different topics. It is a valuable source which provide people's opinions. Twitter contains number of text posts, short messages and it grows every day. Twitter's audience are not only regular users but celebrities are also write their post on twitter, company representatives write their opinion, politicians, and even country presidents also share their messages. Therefore, it is possible to collect text posts of users from different social and interests groups.

Twitter's audiences are from many countries. It is possible to collect data in different languages. We collect text posts from Twitter .Split that posts into following three sets of texts.
A. Texts containing positive emotions, such as happiness, amusement or joy
B. Texts containing negative emotions, such as sadness, anger or disappointment
C. Objective texts that only state a fact or do not express any emotions .

The most important indicators of sentiments are sentiment word.. sentiment words express positive or negative sentiments. For example, good, wonderful, and amazing are positive sentiment words, and bad, poor, and terrible are negative sentiment words. Sentiment words are source to sentiment analysis for obvious reasons. A list of such words is called as sentiment lexicon. A positive or negative sentiment word may have opposite orientations in different application

"Table 1. Examples of twitter posts with expressed users' opinions"

funkeybrewster: @redeyechicago I think Obama's visit might've sealed the victory for Chicago. Hopefully the games mean good things for the city.
vcurve: I like how Google celebrates little things like this: Google.co.jp honors Confucius Birthday— Japan Probe
mattfellows: Hey world. I hate faulty hardware on remote systems where politics prevents you from moving software to less faulty systems.
brroooklyn: I love the sound my iPod makes when I shake to shuffle it. Boo bee boo

III. PROPOSED CONCEPTS

This paper proposes an algorithm for sentiment analysis. The basic concept identifying sentiment words first. These words express user opinion and to find where in the document opinion is present. the opinions extracted .Extracted opinion are analyzed to check polarity of the opinion. This is the bottom-up approach which is used in most of the algorithms [2]. The proposed algorithm analyses the review given by user word by word. Analyzing the opinion requires analyzing of words in a sentence. For such analysis, algorithms proposed have used thesaurus and word net [5][6]. In this paper, sentiment words are stored into database. Sentiment words are used for analysis of opinion. Every sentiment word in the database has been given a value. When a sentiment word is detected in a sentence the value saved in the database is used for evaluating the opinion value.

3.1. Proposed Algorithm

This paper proposes an algorithm for identifying the polarity of remarks. A case has been considered where in a set of students give their feedback about a particular teacher. The algorithm is applied on every remark to identify the polarity of each remark. The algorithm will generates a numeric value for the opinion. If the opinion value are high the opinion are considered positive. Lower opinion value represents negative remarks. The algorithm analyses the remarks word by word [2]. Sentiment words are identified .Each sentiment word contain a combined value which is given to each sentence. The database along with the sentiment word saves an associated value for the opinion word. The value assigned to each sentiment word is based on how much strong or weak sentiment is being used. The value ranges from zero to ten. If a sentiment word emotes strongly positive opinion higher is its value in the database. A sentiment word that represents strong negative opinion lower is its value in the database. When a sentence is analyzed, for each sentiment word found in the sentence, its opinion value is fetched from the database. Then the sentiment value of that sentence is estimated. If there is negation in a sentence the value of opinion score is decreased/ increased by a certain amount [2].

Step 1. For each word in sentences.
Step 2. Check word is negation or sentiment word.Every sentiment word is stored into database referred as dataset.
Step 3. Every sentiment word is assign a value .
If a value is less than 5
It represents negative opinion.
If a value is greater than 5
It represents positive opinion.
Low and high value may be decided using thesaurus.
Step 4. If negation is present in document before sentiment is increased or decreased by 2 depending upon whether the sentiment value is high or low .
Step 5. An average is calculated for all the opinion scores calculated for each remark given by students.The range of value is 0 to 10.
If value is less than 2
Very low score
If value is greater than 2 but it is leass than 4.5
Low score
If the value is greater than 4.5 but less than 5.5
Moderate score
If the value greater than 5.5 but less than 8
High score
If value is greater than 8
Very high.

IV. CONCLUSION

This paper proposes an algorithm for sentiment analysis. The algorithm check each word and compare with sentiment word which are stored in database. These words express user opinion. Extracted opinion are analyzed to check polarity of the opinion. The proposed algorithm analyses the review given by user word by word. Analyzing the opinion

requires analyzing of words in a sentence. Sentiment words are used for analysis of opinion. Every sentiment word in the database has been given a value. When a sentiment word is detected in a sentence the value saved in the database is used for evaluating the opinion value.

In this paper, we analyzed public sentiment and the possible reason behind the variation. Sentiment classifier determines positive, negative and neutral sentiments of document.

REFERENCES

[1] Baker, R., &Yacef, K. (2009).“The State of Educational Data mining in 2009: A Remark Future of Bessel Visions.”Journal of Educational Data Mining.

[2] Lun-Wei Ku, Yu-Ting Liang and Hsin-Hsi Chen, “Opinion Extraction, Summarization and Tracking in and Blog Corpora”, 2006 American Association for Artificial Intelligence.

[3] TsytsarauMikalai, Palpanas Themis. “Survey on mining subjective data on the web”. Data Min KnowlDiscov 2012; 24:478–514.

Wilson T, Wiebe J, Hoffman P. “Recognizing contextual polarity in phrase-level sentiment analysis”. In: Proceedings of HLT/EMNLP; 2005.

[5] M. Hu and B. Liu, “Mining and summarizing customer remarks,” in Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), pp. 168–177, 2004.

[6] R. McDonald, K. Hannan, T. Neylon, M. Wells, and J. Reynar, “Structured models for fine-to-coarse sentiment analysis,” in Proceedings of the Association for Computational Linguistics (ACL), pp. 432–439, Prague, Czech Republic: Association for Computational Linguistics, June 2007.

