

# Review of Dual Sentiment Analysis

Bhagyashri Ramesh Jadhav<sup>1</sup>, Manjushri Mahajan<sup>2</sup>

<sup>1,2</sup>Computer Network, G.H.R.CEM, Wagholi, Pune, Maharashtra.

**Abstract:** *With the vast growth of the social media content on the Internet in the past few years, people now express their opinion on almost anything in discussion. Finding the opinion sites and monitoring them on the web is difficult task. Thus there is a need for automatic opinion Sentiment Analysis is the area of study that analyzes customer feedback, opinions, sentiments, evaluations, attitudes, and from written language. It is one of the most useful research areas in Natural Language Processing and is also widely studied in data mining. The importance of sentiment analysis coincides with the growth of social media such as reviews, discussion forums, blogs, and social networks.*

**Keywords:** sentiments, sentiment analysis, Natural Language Processing

## 1. Introduction

“What people think” has always been an important piece of information during the decision making process. In the past, when an individual needed to make a decision he typically asked for opinions from friends and family. When an organization wanted to find opinions of the general public about its products and services, it conducted surveys.

With the explosive growth of the social media content on the internet in the past few years, the world has been transformed. E Commerce sites, online communities, forums, discussion groups, web logs, product rating sites, chat rooms are some of the sources on which people can now express their views on almost anything in discussion

**Sentiment analysis** refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials. Sentiment analysis is widely applied to reviews and social media for a variety of applications, ranging from marketing to customer service. Analysers used for polarity identification. Analysers are of two types manual (domain oriented) and automatic (generalized oriented) we used domain oriented in are methodology. In manual analyser predefined data set exit which similar/ related term have to feed and result occurs and other hand automatic analyser consist huge data set and also capable to handle multiple language at a time. Sentiment analysis is used to classify polarity and the sentiment analyser is used to define polarity opinion expressed is (+) tive, (-) tive or (=) neutral[1].

In this paper we do a survey of papers on Dual Sentiment Analysis and detail the techniques used.

## 2. Basic Concepts

### a) Sentiments

Sentiment is a sincere and refined sensibility, a tendency to be influenced by emotion rather than reason or fact: to appeal to sentiment. Sentimentality implies affected, excessive, sometimes mawkish sentiment: weak sentimentality. The study of emotions in text can be conducted from two points of view. Firstly, one can investigate how emotions influence a writer of a text in choosing certain words and/or other linguistic elements.

Secondly, one can investigate how a reader interprets the emotion in a text, and what linguistic clues are used to infer the emotion of the writer [4]

### b) Sentiment Analysis

Sentiment Analysis is process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc. is positive, negative, or neutral.

“Sentiment Analysis is the task of identifying positive and negative opinions, emotions, and evaluations”. Sentiment Analysis has many names. It's often referred to as subjectivity analysis, Opinion mining, and appraisal extraction, with some connections to affective computing (computer recognition and expression of emotion) [5].

### c) Dual Sentiment Analysis

A model called dual sentiment analysis (DSA), to address this problem for sentiment classification. We first propose a novel data expansion technique by creating a sentiment-reversed review for each training and test review. On this basis, we propose a dual training algorithm to make use of original and reversed training reviews in pairs for learning a sentiment classifier, and a dual prediction algorithm to classify the test reviews by considering two sides of one review [1]

## 3. Source of Data

The major source for obtaining data is websites, software, fact, figures etc. Huge data can be found about company, consumers, producers, retailers, legal documents, data warehouse etc. Opinion mining is used to analyze this data. Details are given below.

### a) World Wide Web

In web there are huge amount of data is present this data is the big source for researchers and users both in the world of web blogs, comment box, form feeding etc techniques is used for data extracting and to transfer data emails, face book, twitter is used.

**b) Sites**

In the current era many recognized groups are doing the work of analyzing data and maintaining the sites. Industries are hiring people to performing the respective work. Finally on the bases of obtained data such as price, quantity, ranking the product result occurs.

**c) Web Based Interface**

Interface is a medium between user and the web this consist text messaging, digital audio/video, e-mail, links etc. Interface plays an essential role between user and web because without interface it is not possible to connect or communicate. [3]

**4. Sentiment Analysis Techniques**

Sentiment analysis includes two techniques ,these are supervised and unsupervised

**a) Supervised**

Supervised approach is called Machine learning. Machine learning approach uses training data to build a predictive model. Predictive models such as decision trees, logistic regressions or neural networks are used to make prediction on documents that are outside the training set.

This approach has an advantage as it is based on learning patterns that are useful for making automated and efficient predictions. Also the algorithms are capable of discovering unimagined and complex patterns that would be beyond what a human could anticipate. However it has drawbacks as huge training data is necessary to build the model and validating it is time consuming and challenging. A rating is required to be provided for every document, and if there are attributes of documents it is needed to provide a rating for each of these as well. Another complication arises if two different reviewers assign two different sentiment ratings to the same document, then this can introduce unexpected errors in building and measuring the performance of model.

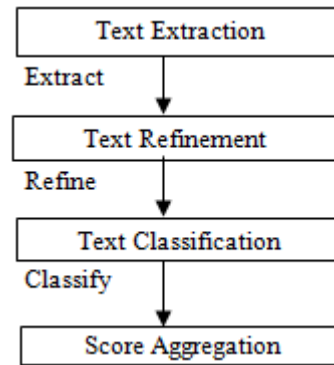
**b) Unsupervised**

Unsupervised approach is called Natural language processing. Natural language processing (NLP) is a field of artificial intelligence that deals with automatically extracting meaning from natural language text. It uses entities and syntactic patterns in the text to understand its meaning. It also uses a combination of language dictionaries, linguistic constructs like parts of speech, and noun phrases along with a range of operators.

The major advantage of rule-based methods is that it provides freedom for the rule developers to use their domain knowledge to devise rules for analysis purpose. Rule-based methods are completely unsupervised and they do not require any training data. This is a main advantage in real-life applications where training data is scarce. Additionally it provides the facility to refine the rules over time based on the feedback from analysts or subject-matter experts to adjust the models. The major drawback of NLP approach is that they require a lot of human involvement in developing the rules and it completely rely on the domain knowledge of rule developers.

**5. Main Steps of Sentiment Analysis**

- Text Extraction – This step involves extracting words from text that influence the outcome of the result.
- Text Refinement – This step involves refining text in form of relevant phrases, words etc.
- Text Classification – This step includes classification of text into its class(positive/negative)

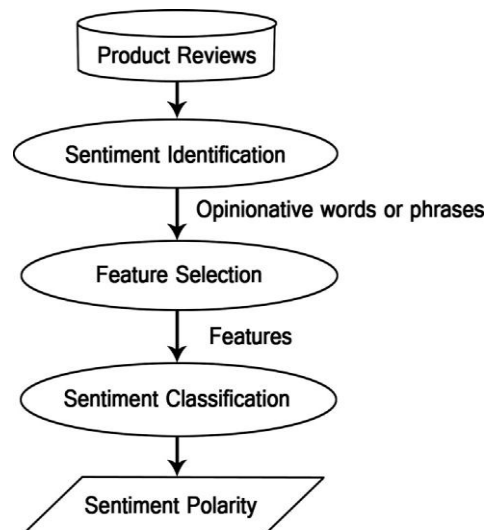


**Figure 1:** Steps of Sentiment Analysis

- Score Aggregation – This step collects total scores from classifier and then aggregates it further to produce the total sentiment score [6].

**6. Sentiment Analysis According To Product Review**

Sentiment Analysis can be considered a classification process as illustrated in Fig.2. There are three main classification levels in SA: document-level, sentence-level, and aspect-level SA. Document-level SA aims to classify an opinion document as expressing a positive or negative opinion or sentiment



**Figure 2:** Sentiment analysis process on product reviews

It considers the whole document a basic information unit(talking about one topic). Sentence-level SA aims to classify sentiment expressed in each sentence. The first step is to identify whether the sentence is subjective or objective. If the sentence is subjective, Sentence-level SA will

determine whether the sentence expresses positive or negative opinions [7].

## 7. Sentiment Analysis at different level classification

### A Document Level classification

Identify if the document (e.g. product reviews, blogs, forum posts) expresses opinions and whether the opinions are positive, negative, or neutral. (Technique: Heuristics)

### B Sentence Level classification

Identify if a sentence is opinionated and whether the opinion is positive, negative, or neutral. (Technique: Part of Speech Tagging)

### C Attribute Level classification

Extract the object attributes (e.g. image quality, zoom size) that are the subject of an opinion and the opinion orientations.

## 8. Application of Sentiment Analysis

Some of the applications of sentiment analysis include online Advertising ,hotspot detection in forums etc. Online advertising has become one of the major revenue sources of today's Internet ecosystem. Sentiment analysis find its recent application in Dissatisfaction oriented online advertising [9] and Blogger-Centric Contextual Advertising [8], which refers to the assignment of personal ads to any blog page, chosen in according to bloggers' interests When faced with tremendous amounts of online information from various online forums, information seekers usually find it very difficult to yield accurate information that is useful to them. This has motivated the research on identification of online forum hotspots, where useful information is quickly exposed to those seekers. Nan Li et al., [10] used sentiment analysis approach to provide a comprehensive and timely description of the interacting structural natural groupings of various forums, which will dynamically enable efficient detection of hotspot forums. In order to identify potential risks, it is important for companies to collect and analyze information about their competitors' products and plans. Sentiment analysis find a major role in competitive intelligence [11] to extract and visualize comparative relations between products from customer reviews, with the interdependencies among relations taken into consideration, to help enterprises discover potential risks and further design new products and marketing strategies. Opinion summarization summarizes opinions of articles by telling sentiment polarities, degree and the correlated events. With opinion summarization, a customer can easily see how the existing customers feel about a product, and the product manufacturer can get the reason why different stands people like it or what they complain about. Ku et al., in [12] investigated both news and web blog articles. Algorithms for opinion extraction at word, sentence and document level are proposed. The issue of relevant sentence selection is discussed, and then topical and opinionated information are summarized.

Opinion summarizations are visualized by representative sentences. Finally, an opinionated curve showing supportive and non-supportive degree along the timeline is illustrated by an opinion tracking system. Other applications includes online message sentiment filtering-mail sentiment classification, web blog authors attitude analysis etc. Review Seer is a tool that automates the work done by aggregation sites. Naive Bays classifier is used with positive and negative review sets for assigning a score to the extracted feature terms. The classifier did not perform well for web pages crawled from the result of a search engine. It displays attributes and score of the attribute along with review sentences. Web Fountain uses beginning definite Base Noun

Phrase (bBNP) heuristic for extracting product features. To assign sentiments to the features, reviews are parsed and traversed with two linguistic resources namely the sentiment lexicon and the sentiment pattern database. The sentiment lexicon defines the polarity of terms and sentiment pattern database defines sentiment extraction patterns for a sentence predicates [13].

## 9. Conclusion

Sentiment analysis is essential for anyone who is going to make a decision. Sentiment analysis is helpful in different field for calculating, identifying and expressing sentiment. It is helpful for everyone when they want to buy a product and they can decide which product is best. Sentiment analysis is very important for Enterprises and helps them to know what customers think about their products. Therefore companies can take decisions about their products based on customer's feedback Thus companies can modify their products features and introduce new products according to customers' opinions in a better and faster way.

## References

- [1] Biber, D; Finegan, E. Styles of stance in english: Lexical and grammatical marketing of evidentiality and affect. Text 9, 1989, pp. 93-124.
- [2] 4. Erik Boiy; Pieter Hens; Koen Deschacht; Marie-Francine Moens" Automatic Sentiment Analysis in On-line Text", Proceedings ELPUB2007 Conference on Electronic Publishing – Vienna, Austria – June 2007
- [3] Abdullah Dar\*, Anurag Jain "Survey paper on Sentiment Analysis: In General Terms" Nov 2014
- [4] Pang, B. and Lee, L. (2008). Opinion mining and sentiment analysis. *Foundation and Trends in Information Retrieval*, 2(1-2):1–135.
- [5] Sasha Blair- Goldensohn, Kerry Hannan, Ryan McDonald, Tyler
- [6] Neylon, George A. Reis, Jeff Reynar. *Building a Sentiment*
- [7] Summarizer for local service Reviews, 2008.
- [8] Wilson T, Wiebe J, Hoffman P. Recognizing contextual polarity
- [9] in phrase-level sentiment analysis. In: Proceedings of HLT/EMNLP; 2005.
- [10] Fan Teng-Kai, Chang Chia-Hui, Blogger-Centric Contextual Advertising. In: Expert Syst. Appl., 2011.

- [11] Guang Qiu, Xiaofei He, Feng Zhang, Yuan Shi, Jiajun Bu, Chun
- [12] Chen: DASA: Dissatisfaction-oriented Advertising based on Sentiment Analysis .Expert Syst. Appl. 37(9): 6182-6191 (2010)
- [13] Nan Li a, Desheng Dash Wu b,c, Using text mining and sentiment analysis for online forums hotspot detection and forecast, Decision Support Systems 48 (2010) 354–368
- [14] Kaiquan Xu , Stephen Shaoyi Liao , Jiexun Li, Yuxia Song, “Mining comparative opinions from customer reviews for Competitive
- [15] Intelligence”, Decision Support Systems 50 (2011) 743–754.
- [16] Ku, L.-W., Liang, Y.-T., & Chen, H, “Opinion extraction,
- [17] summarization and tracking in news and blog corpora”. In AAAICAAW“ 06.
- [18] Yi and Niblack, “Sentiment Mining in Web Fountain”,Proceedings of 21st international Conference on Data Engineering, pp. 1073-1083, Washington DC,2005
- [19] Rui Xia, Feng Xu, Chengqing Zong, Qianmu Li, Yong Qi, and Tao Li” Dual Sentiment Analysis: Considering Two Sides of One Review” IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 27, NO. 8, AUGUST 2015 1041-4347 \_ 2015 IEEE..