

## A Review on Opinion Mining and Sentiment Analysis Performance at Different Levels

Mrs. K. Chitra, Dr. A. Tamilarasi

Assistant Professor, Department of Computer Applications Kongu Engineering College, Perundurai  
Professor and Head Department of Computer Applications Kongu Engineering College, Perundurai

---

**Abstract:** Generally, the opinions are gathered from the public about an entity like any product, event, movie etc. is very essential because of customer satisfaction for a specific entity. Opinions are statements that represent people's perception or sentiment, which is used to finding the customer feelings or thoughts about a purchasing of a particular product or topic with the use of natural language processing. Consider, if anyone (new user) can purchase any product then they can get an idea from the existing user's opinion or statement whether the product is good, bad or neutral. Based on existing user views the new user can conclude whether the product is purchase or not. It involves building a system to collect and examine opinions about the product made in many online purchasing sites. Opinion mining is a part of web content mining. Web content mining is a part of Data mining. The important role of Sentiment analysis is to train computer to be able to understand, recognize and generate emotions. The objective of this paper is to provide knowledge about how the customer opinions are performed at different levels.

---

### I. Introduction

Sentiment analysis, also known as opinion mining, is the analysis of the feelings that is people's opinions, sentiments, attitude, emotions, evaluations, appraisals towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes using natural language processing tools. It deals with opinion oriented natural language processing. Such opinion-oriented studies include among others, genre distinctions, emotion and mood recognition, ranking, relevance computations, perspectives in text, text source identification and opinion oriented summarization. Sentiment analysis has turned out as an exciting new trend in social media with a gamut of practical applications that range from applications in business (marketing intelligence; product and service bench marking and improvement), applications as sub component technology (recommender systems; summarization; question answering) to applications in politics. It has great potential to be used in business strategies and has helped organizations get a real-time feedback loop about their marketing strategy or advertisements from the reaction of the public through tweets, posts and blogs. For a new product launch it can give them instant feedback about the reception of the new product. It can have what their brand image is, whether they are liked or not.

As the field of sentiment analysis is relatively new, the terminology used to describe this field of research is many. The terms opinion mining, subjectivity analysis, review mining and appraisal extraction are used interchangeably with sentiment analysis. Subjectivity analysis or subjectivity classification is focused on the task of whether the sentence or document is expressing opinions or sentiments of the author or just merely stating facts.

Recently a lot of interest has been generated in the field of sentiment analysis, with researchers recognizing the scientific trials and potential applications supported by the processing of subjective language. Sentiment direction, where users can search for positive or negative tweets on a particular topic.

This paper gives an overview of sentiment analysis, different ways to evaluate opinion and its basic terminology, how can assess the whether the opinion is positive or negative through its workflow. Lastly, it discusses the performance of sentiment analysis at various levels and algorithms to get the better knowledge in that field.

### II. Ways Of Evaluation In Opinion Mining

- Direct opinion, gives positive or negative opinion about the object directly. For example, "This film was not likable".
- Comparison means to compare the object with some other similar objects. For example, "The film Y is better than that X" expresses a comparison.

An opinion (or regular opinion) is simply a positive or negative sentiment, attitude, emotion or appraisal about an entity or an aspect of the entity from an opinion holder. Positive, negative and neutral are called opinion orientations (also called sentiment orientations, semantic orientations, or polarities).

### III. Opinion Mining Terminologies

This represents basic terminologies currently used in the area of opinion mining.

**Fact:** A fact is something that has really occurred or is actually the case.

**Opinion:** An opinion is a belief about matters commonly considered subjective, and is the result of emotion or interpretation of facts.

**Subjective/opinionated text:** A text is subjective or opinionated if it expresses personal feelings or beliefs, e.g. opinions.

**Objective text:** An objective text expresses some information about the world.

**Item:** An item is a concrete or abstract object such as product, service, person, event, organization. An item can be represented as a hierarchy of components, sub-components, etc.

**Review:** A review is a subjective text containing a sequence of words describing opinions of reviewer regarding a specific item. Review text may contain complete sentences, short comments, or both.

**Short comments or pros/cons:** The reviewer can describe pros and cons of the item.

**Object:** An entity which can be a product, person, event, organization, or topic. The object can have attributes, features or components associated with it. Further on the components can have subcomponents and attributes .

**Feature:** An attribute (or a part) of the object with respect to which evaluation is made.

**Opinion orientation or polarity:** The orientation of an opinion on a feature f indicates whether the opinion is positive, negative or neutral. Most work has been done on binary classification i.e. into positive or negative. But opinions can vary in intensity from very strong to weak [7]. For example a positive sentiment can range from content to happy to ecstatic. Thus, strength of opinion can be scaled and depending on the application the number of levels can be decided.

**Opinion holder:** The holder of an opinion is the person or organization that expresses the opinion

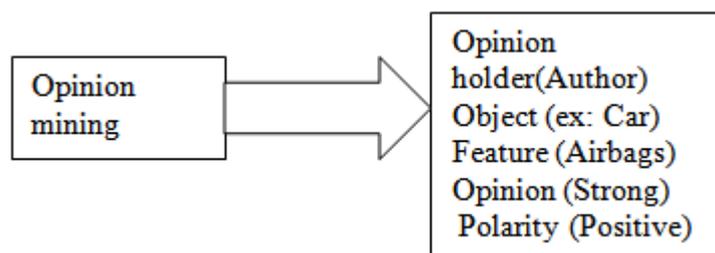


Fig.1 Terminology of Sentiment Analysis

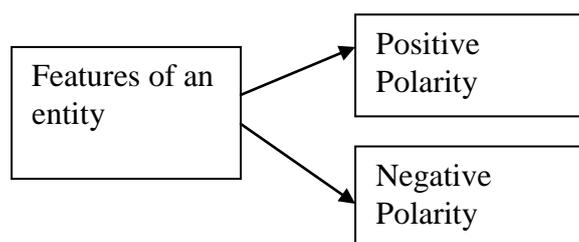


Fig.2 Polarity Prediction based on features of entity

The above diagram (Fig1 & 2) represents based on the Features of an entity we may have predict the polarity belongs to positive or negative because the opinion is based on the features of an entity. If the Opinion is weak then the polarity is negative otherwise positive one.

### IV. Opinion Mining Work Flow

- ✓ Web users
- ✓ Comments Reviews
- ✓ Input Document
- ✓ Preprocessing
- ✓ Sentiment Classification(Positive or Negative)
- ✓ Rating
- ✓ Opinion impact

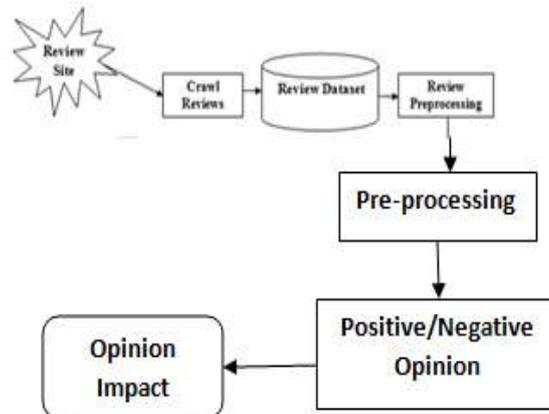


Fig.3: Opinion Mining Work Flow

## V. Sentiment Analysis Performance

Sentiment Analysis tasks are mainly divided into the polarity of a given text at the document, sentence and feature level / attribute level / aspect level / phrase level to find whether it give positive opinion, negative opinion or neutral. This is also referred as 'Sentiment Polarity Prediction' [1]. The Sentiment Analysis performance is carried out into three levels,

- ✓ The document level
- ✓ The Sentence Level
- ✓ The Feature Level

### 5.1 Document Level Sentiment Analysis

Here , the whole document considered as a single entity and classify the sentiment as positive or negative or neutral. Document-level sentiment classification assumes that the opinionated document expresses opinions on a single target and the opinions belong to a single person, product, or service. It is clear that this assumption is true for customer reviews of products documents, which usually focus on one product, and single reviewer writes it.

There are two types of classification techniques that have been used in document-level sentiment classification such as supervised method and unsupervised method.

#### 5.1.1 Supervised Method

Sentiment classification task can be spontaneous as a supervised learning problem with two classes, positive and negative. Product reviews are used as training and testing data. Another method such as online reviews have rating scores assigned by their reviewers, e.g., 1-5 stars, the positive and negative classes are determined using the ratings. A review with 4 or 5 stars is considered a positive review, and a review with 1 to 2 stars is considered a negative review.

A training set is used by an automatic classifier to learn the differentiating characteristics of documents, and a test set is used to check how well the classifier performs. The most commonly used features in sentiment classification are introduced below

##### 5.1.1.1. Terms and their frequency

The features can be considered as individual words termed as unigram and their n-grams with associated frequency counts. It can be calculated as weightage form.

##### 5.1.1.2. Part of speech (POS)

POS information is a very important indicator of sentiment expression.

##### 5.1.1.3. Sentiment words and phrases

Sentiment words and phrases that is express in positive and negative sentiments. For example good, fantastic, amazing words are positive sentiment and bad, boring, slow, worst and poor are negative sentiment. For example rubbish (noun), hate and like (verb) can indicate opinion in some documents.

#### **5.1.1.4. Negations:**

Negation words are very important to evaluate the polarity of a sentence because they can transform the sentiment orientation in a sentence. For instance, the sentence “I don’t like this phone” has negative orientation.

#### **5.1.1.5. Syntactic dependency:**

It is represented as a parse tree and it contains word dependency based features.

### **5.1.2 Unsupervised Method**

The Unsupervised learning method has a set of inputs, like clustering, labels are not known during training. Classification is performed using some fixed viti patterns which are used to express opinions. The part-of-speech (POS) tags are used to compose viti patterns. The unsupervised learning can be done by extracting the opinion words inside a document .The point-wise mutual information can be made use of to find the semantics of the extracted words.

### **5.2 Sentence Level Sentiment analysis**

In the sentence level sentiment analysis, the polarity of each sentence is calculated. The same document level classification methods can be applied to the sentence level classification problem. Objective and subjective sentences must be found out. The subjective sentences contain opinion words which help in determining the sentiment about the entity. After which the polarity classification is done into positive and negative classes. Knowing that a sentence is positive or negative is of lesser use than knowing the polarity of a particular feature of a product. The advantage of sentence level analysis lies in the subjectivity/ objectivity classification.

Sentence level sentiment Analysis has two tasks:

#### **5.2.1 Subjectivity Classification**

A sentence can be either subjective sentence or objective sentence. Objective sentence contains the facts. It has no judgement or opinion about the object or entity while subjective sentence has opinions. The advantage of sentence level analysis lies in the subjectivity/ objectivity classification .

#### **5.2.2 Sentiment Classification**

Sentence can be classified as positive, negative or neutral depending upon the opinion words present in it. A number of researches focus on finding how to classify the text effectively. The same document level classification methods can be applied to the sentence level classification problem. A number of different methods are discussed and compared under supervised machine learning approach. Finally, they observe that first and last lines of a review are often indicative of review polarity. They performed an in depth analysis on different sentence types and the semantic score of subjective sentence is extracted from SentiwordNet lexical resource.

### **5.3 Feature Level Sentiment Analysis:**

This level of sentiment classification is a much more pinpointed method to opinion mining. This type of classification considers the opinions on features of particular objects. Features of the product are defined as attributes, components and other aspects of the product, Analysis of such features are recognizing sentiment of the document is called as Feature based Sentiment Analysis. The task of Feature Level sentiment classification is to extract the features of the commented object and after that conclude the opinion of the object. Positive or negative and then group the feature synonyms and make the summary report.

## **VI. Sentiment Lexicon Construction**

Sentiment lexicon words are identified by opinion words are also known as opinion-bearing words. Sentiment words are always divided into two categories such as positive or negative. Positive opinion words are used to express some desired states while negative opinion words are used to express some undesired states. There are three methods to construct a sentiment lexicon: manually construction, corpus-based methods and dictionary-based methods.

### **1) Manually Construction:**

The manual construction of sentiment lexicon is a very hard and time-consuming task and always cannot be used alone but it can be combined with other methods to improve the accuracy of these methods.

## **2) Corpus –Based Method**

Corpus-based methods can produce opinion words with relatively high accuracy. Most of these corpus based methods need very large labeled training data. This approach has a major advantage that the dictionary-based approach does not have. It can help find domain specific opinion words and their orientations.

## **3) Dictionary -Based Method**

In this methods is to collect an initial seed set of sentimental words and their orientation manually, and then searching in a dictionary to find their synonyms and antonyms to expand this set. The new seed set are used iteratively to generate new sentiment words.

## **VII. Aspect- Based Sentiment Analysis**

The major steps involved to produce this detailed summary of product reviews are

- (1) Aspect extraction
- (2) Aspect sentiment orientation detection

### **1) Aspect extraction**

Aspect extraction indicate as feature extraction is one of the key tasks in aspect-based sentiment analysis. This method proposed is based on information extraction approach that identifies frequently occurring noun phrases. This approach is generally useful in finding aspects which are strongly associated with a single noun. The disadvantage of this approach is that it cannot detect the aspect terms which are of low frequency and noun phrases.

### **2) Aspect sentiment orientation detection**

It resolves the sentiment orientation expressed on each aspect in a sentence is the second task in aspect based sentiment analysis. It must determine whether the sentiment orientation on each aspect is positive, negative or neutral. This task can be divided into the following sub tasks:

1. Extracting opinion words or phrases.
2. Identifying the polarity of each opinion words or phrases
3. Handling opinion shifters (such as no, not, don't) and opinion intensifiers (such as very, extremely)
4. Handling but clauses.
5. Aggregating opinions (if there is more than one opinion word or phrase in a sentence).

## **VIII. Conclusion**

This paper offers a review of research work done in several aspects of Sentiment Analysis. It describes earlier review, text mining preprocessing, and sentiment analysis performance at various levels. The Extraction of opinion words / sentences is very challenging task. That is difficult to differentiate objective sentence or subjective information's. There are many techniques, algorithms are used to perform this task, and many of the studies are remain unsolved. More future research works could be committed to these difficulties and challenges. Main contribution of this review paper is to extracting the survey of people's opinions, emotions and sentiments in the field of opinion mining.. This review paper will be useful for researchers and beginners in the field of Opinion mining

## **References**

- [1]. K.G. Nandakumar, Dr.T.Christopher "Opinion Mining: A Survey", International Journal of Computer Applications, Volume 113, No.2, March 2015.
- [2]. Sindhu C1, Dr. S. ChandraKala, "A Survey on opinion mining and sentiment polarity classification". International Journal of Emerging Technology and Advanced Engineering. , Volume 3, Special Issue 1, January 2013).
- [3]. Ziqiong Zhang, Qiang Ye, Zili Zhang, Yijun Li, "Sentiment Classification of Internet Restaurant Reviews written in Cantonese", Expert Systems with Applications, 2011.
- [4]. B.Liu. 2010 "Sentiment Analysis and Subjectivity", Second Edition, The Handbook of Natural Lanugage Processing.
- [5]. Bing Liu, Sentiment Analysis, Mining opinions, Sentiments, and Emotions, Book, June (2015).
- [6]. Muhammaad Zubair, Aurangzeb Khan, Shakeel Ahmad, FazalMasudKundi and Asghar,2014.A Review of Feature Extraction in Sentiment Analysis. ISSN 2090-4304 Journal of Basic and Applied Scientific Research.
- [7]. W. Medhat et al., Sentiment analysis algorithms and applications: a survey, Ain Shams Eng. J. (2014).
- [8]. X. Ding. S. M. Street, B. Liu, and P. S. Yu, A Holistic Lexicon based approach to opinion mining (2008) pp. 231-239.
- [9]. Haseena Rahmath P," Opinion Mining and Sentiment Analysis challenges and Applications", IJAIEM Volume 3, Issue 5, May 2014.
- [10]. K Ravi and V Ravi, A survey on opinion mining and sentiment analysis: Tasks, approaches and applications, in Knowledge Based System (2015).
- [11]. S. Veeramani1, S. Karuppusamy,"A Survey on Sentiment Analysis Technique in Web Opinion Mining". International Journal of Science and Research (IJSR).