Fourth International Conference on Recent Trends in Computer Science & Engineering.

Chennai, Tamil Nadu, India

Sentiment Analysis:A Comparative Study On Different Approaches

Devika M D[a]*, Sunitha C[a], Amal Ganesh[a]

Department of CSE, Vidya Academy of Science and Technology, Thrissur 680501, India

Abstract

Sentiment analysis (SA) is an intellectual process of extricating user's feelings and emotions. It is one of the pursued field of Natural Language Processing (NLP). The evolution of Internet based applications has steered massive amount of personalized reviews for various related information on the Web. These reviews exist in different forms like social Medias, blogs, Wiki or forum websites. Both travelers and customers find the information in these reviews to be beneficial for their understanding and planning processes. The boom of search engines like Yahoo and Google has flooded users with copious amount of relevant reviews about specific destinations, which is still beyond human comprehension. Sentiment Analysis poses as a powerful tool for users to extract the needful information, as well as to aggregate the collective sentiments of the reviews. Several methods have come to the limelight in recent years for accomplishing this task. In this paper we compare the various techniques used for Sentiment Analysis by analyzing various methodologies.

Keywords: Sentiments*; Lexicons; Polarity*

1. **Introduction**

Sentiment analysis is a kind of text classification that catalogs texts based on the sentiment orientation of opinions they contain. It thus plays an important part of Natural Language Processing. NLP is a field of computer science and artificial intelligence that mainly deals with human-computer language interaction. This field is particularly of use to merchants, stock traders, and in election works.

Sentiment analysis is the process of detecting the contextual polarity of the text. It determines whether given text is positive, negative or neutral. It is otherwise called as opinion mining too, since it derives the opinion or attitude of the speaker. For this analysis, the opinions are collected from the users, which can be employed for further

* Corresponding author. Tel.: 0091-9846947813
E-mail address: md.devika@gmail.com

improvements. The social networks act as a medium where the users can post many opinions a day and these blogs are used for classification. A lot of research work is being held in the field of sentiment analysis due to its significance in the marketing level competition and the changing needs of the people. Sentiment analysis requires the usage of a training set for its performance, and its quality plays a great role in the accurate evaluation of the text. The semantic analysis of the sentence also increases the meaning and accuracy of the result. POS tagging will be helpful to users for understanding whether the review or comment corresponds to the relevant subject searched for.

2. Levels of Analysis

In general, sentiment analysis has been investigated mainly at three levels [1]. In document level the main task is to classify whether a whole opinion document expresses a positive or negative sentiment. This level of analysis assumes that each document expresses opinions on a single entity. In sentence level the main task is to check whether each sentence expressed a positive, negative, or neutral opinion. This level of analysis is closely related to subjectivity classification, which distinguishes objective sentences that express factual information from subjective sentences that express subjective views and opinion. Document level and the sentence level analyses do not discover what exactly people liked and did not like. Aspect level performs finer-grained analysis. Instead of looking at language constructs (documents, paragraphs, sentences, clauses or phrases), aspect level directly looks at the opinion itself.

3. Sentiment Analysis Methods

Sentiment analysis played a great role in the area of researches done by many, there are many methods to carry out sentiment analysis. Still many researches are going on to find out better alternatives due to its importance in this scenario. Some of the methods are discussed in this paper.

3.1. Machine learning Approach

Machine learning strategies work by training an algorithm with a training data set before applying it to the actual data set. Machine learning techniques first trains the algorithm with some particular inputs with known outputs so that later it can work with new unknown data [2]. Some of the most renowned works based on machine learning are as follows:

3.1.1 Support Vector Machine

It is a non-probabilistic classifier in which a large amount of training set is required. It is done by classifying points using a (d-1)-dimensional hyper plane. SVM finds a hyper plane with largest possible margin [3]. Support Vector Machines make use of the concept of decision planes that define decision boundaries. A decision plane is one that separates between a set of objects having different class membership. An illustration is given in Fig. 1a. In this the objects belong to either class red or green, and the separating line defines the boundary. Here the original objects are (left side of Fig. 1b) mapped or rearranged using a mathematical function known as kernel and this is known as mapping or transformation. After transformation, the mapped objects are linearly separable and as a result the complex structures having curves to separate the objects can be avoided.