An Overview on Opinion Mining Techniques and Sentiment Analysis

Penubaka Balaji¹, D.Haritha² and O.Nagaraju³
¹Research Scholar, ²Professor & Head-BES Department of Computer Science & Engineering, K.L.University, Guntur, AP, India. penubakabalaji.cse@gmail.com, haritha_donavalli@kluniversity.in
³Head of Computer Science, Government College, Macherla, AP, India. profonr@gmail.com

February 1, 2018

Abstract

Now a days, customers opinions are plays the major role in the E-commerce applications such as Flipkart, Amazon, eBay etc. Based on customer feedback on the product or seller in the form reviews or comments are the difficulty process by potential buyers to choose a products through online. In the proposed system, the various sentiment analysis techniques to provide a solution in two main areas. 1) Extract customer opinions on specific product or seller. 2) Analyze the sentiments towards that specific product or seller. In this paper, we analyzed several opinion mining techniques and sentiment analysis and their correctness in the categories of opinions or sentiments.

Key Words : Opinion mining, Sentiment analysis (SA), Reviews, Comments.
1 Introduction

Opinion Mining or Sentiment Analysis is the evaluation model to learning of public opinions, attitudes and feelings toward any item, product or seller. The object can characterize persons, objects or topics [1]. Opinion Mining is one of the greatest dynamic research area in Natural Language Processing. Sentiment Analysis defines a procedure of mining, classifying, analyzing and describing the feelings or sentiments in the form of word-based data using Machine Learning, Natural Language Processing or Statistics. The two terminologies sentiment analysis or opinion mining are more substitutable. Opinion Mining mine the textual data and evaluates public’s attitude around an object whereas sentiment analysis classifies the sentiment articulated in a script then examines it. Sentiment analysis can be measured a taxonomy process as illustrated in Fig. 1. There are three main categories in sentiment analysis: document-level SA, sentence-level SA, and aspect-level SA.

**Document-level SA:** its main objective is, to categorize an attitude text as articulating a positive or negative attitude or sentiment. It deliberates the complete text a basic data unit.

**Sentence-level SA:** its main objective is, to categorize sentiment articulated in individual sentence. The initial stage is to classify either the sentence is subjective nor objective. If the sentence is subjective, Sentence-level sentiment analysis will decide whether the sentence articulates a positive or negative feelings.

**Aspect-level SA:** its main objective is, to categorize the sentiment through feature to the exact features of objects. The primary stage is to classify the objects and their features.

The opinion holders can give dissimilar opinions for dissimilar features of the same object like this sentence "The camera of this phone is not good, but the voice clarity is excellent".

The entity sets are most important issue in the sentiment analysis field. The main sources of data are the product reviews, news articles or political debates etc. These are plays the major role in the selection of business decision making in such that they analyze the results of customers opinions about their items. The review sources are main review websites. The sentiment analysis in not only used in items or product reviews but can be used in stock markets, news articles or political debates. In the social network
sites such as amazon, flipkart, etc. are recorded customer opinions or reviews for their products or sellers. These are very useful information for new buyer to choose products or sellers in their sites and other sites too.

2 Opinion/sentiment components and its types

2.1 Opinion/Sentiment components:

There are three main components in the opinion/sentiment [12].

i) Opinion holder: Person who gives a comment.
   
   *Ex. The camera quality of this phone is excellent.*

ii) Opinion object: Object on which comment expressed.

   *Ex. The opinion object is "the camera quality of this phone is excellent".*

iii) Opinion orientation: Find the comment either positive or negative or neutral

   *Ex. The camera quality of this phone is excellent.*
2.2 Opinion/Sentiment Types:

There are two main types

2.2.1 Regular type: A regular opinion is often referred simply as an opinion in the literature and it has two subtypes.

2.2.1.1. Direct Opinion: A direct opinion denotes to an attitude articulated straight on an object or an object aspect. For example, "The battery life of this mobile phone is good"

2.2.1.2. Indirect Opinion: It denotes to an opinion that is articulated indirectly on an object or an object aspect. For example, "After taking this syrup, my body pains relieved"

2.2.2 Comparative type: A comparative opinion states a relation of similarities or differences between two or more entities. For example, the sentences, "Boost tastes better than Horlicks" and "Boost tastes the best" express two comparative opinions.

3 Literature review

Sentiment Classification techniques can be categorized into two main approaches [1] such as machine learning approach, lexicon based approach.

Machine Learning Approach relates to the well-known machine learning algorithms and uses semantic features by using supervised and un-supervised learning mechanisms.

The Lexicon-based Approach trusts on a sentiment wordlist, a pool of well-known and precompiled feeling or sentiments expressions. It is categorized into dictionary-based approach and corpus-based approach which use statistical or semantic methods to find sentiment polarity.

The hybrid Approach associations of together methodologies and is very collective with sentiment wordlists playing a important role in the mainstream of approaches. The several methods and the best widespread algorithms of sentiment or opinion classification are illustrated in Fig. 2.1 as shown below.
There are many more applications and improvements on opinion mining and sentiment analysis algorithms. This survey gives several sentiment or opinion mining technique and their accuracy given by authors (Table 1).

---

<table>
<thead>
<tr>
<th>S.No</th>
<th>Authors</th>
<th>Title</th>
<th>Opinion Mining Techniques used</th>
<th>Accuracy in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pravekumar Singh, Mohd Shahid Hussain</td>
<td>Methodological Study of Opinion Mining and Sentiment Analysis</td>
<td>Naive Bayers, Multilayer Perception</td>
<td>76</td>
</tr>
</tbody>
</table>

*Source: Am Shams Engineering Journal (2014) 5:1093–1113*
<table>
<thead>
<tr>
<th></th>
<th>Author(s)</th>
<th>Technique Description</th>
<th>Approach Type</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Gayathri Deepthi, K. Sashi Rekha</td>
<td>Opinion Mining and Classification of User Reviews in Social Media</td>
<td>Naive Bayes</td>
<td>84</td>
</tr>
<tr>
<td>3</td>
<td>Richa Sharma, Shweta Nigam, Rekha Jain</td>
<td>Mining of Product Reviews at Aspect Level</td>
<td>Dictionary based Approach</td>
<td>67</td>
</tr>
<tr>
<td>4</td>
<td>Vaishali Mehta, Prof. Ritesh K Shah</td>
<td>Approaches of Opinion Mining and Performance Analysis: A Survey</td>
<td>Naive Bayes Genetic Algorithm Naive Bayes+Genetic Algorithm</td>
<td>91 91 93</td>
</tr>
<tr>
<td>5</td>
<td>Gautami Tripathi, Nagama S</td>
<td>Feature Selection and Classification approach for Sentiment Analysis</td>
<td>Support Vector Machine</td>
<td>85</td>
</tr>
<tr>
<td>6</td>
<td>Gagandeep Singh, Kamaljeet Kaur Mangat</td>
<td>Performance Analysis of Supervised learning Methodologies for Sentiment Analysis of Tweets</td>
<td>Naive Bayes Maximum Entropy Support Vector Machine K Nearest Neighbour</td>
<td>71 75 81 85</td>
</tr>
<tr>
<td>7</td>
<td>Dola Saha, Prajna Paramita Ray</td>
<td>Sentiment Analysis on Tweet Dataset using Data Mining Techniques</td>
<td>K-means</td>
<td>77</td>
</tr>
<tr>
<td>8</td>
<td>Preety, Sunny Dahiya</td>
<td>Sentiment Analysis using Support Vector Machine And Naive Bayes Algorithm</td>
<td>Naive Bayes Support Vector Machine Naive Bayes+K Means</td>
<td>80 84 89</td>
</tr>
<tr>
<td>9</td>
<td>K. Uma Maheswari,</td>
<td>Opinion Mining using Hybrid Methods</td>
<td>Support Vector Machine</td>
<td>68</td>
</tr>
<tr>
<td>10</td>
<td>S.P. Raja Mohana, G. Aishwarya Lakshmi</td>
<td>Support Vector Machine+Particle Swarm Optimization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Chetanshi Bhadane, Hardi Dalal, Heenal Doshi</td>
<td>Sentiment Analysis: Measuring Opinions</td>
<td>Support Vector Machine</td>
<td>78</td>
</tr>
</tbody>
</table>

Table 1 Performance index of Opinion mining techniques given by different authors
4 Conclusion & future work

Opinion mining or sentiment analysis is an important role of data mining applications to mine the pearl knowledge from large volume of customer feedback, comments or reviews of any item, product or topic. A lot of work has been discussed and conducted to extract sentiments such as document, sentence, and aspect feature level opinion analysis. The data sources from social websites, microblogs, news articles and forums are mostly used in opinion analysis now a days. These data sources are used in expressing people’s feelings or feedback about specific item or topic. This paper offered many sentiment or opinion mining techniques, levels and their types and finally these are applied by the authors and the accuracy produced.

By using LDA, topic modelling, extrinsic and intrinsic feature selection algorithms, we can compute more accuracy than previous one.

References


