

Sentiment Analysis by Visual Inspection of User Data from Social Sites - A Review on Opinion Mining

Vijayshri R. Ingale¹, Rajesh Nandkumar Phursule²

¹Department of Computer Engineering, JSPM's Imperial College of Engineering & Research, Pune, Maharashtra, India

²Professor, Department of Computer Engineering, JSPM's Imperial College of Engineering & Research, Pune, Maharashtra, India

Abstract: *Positive online reviews have a significant impact on customer's decision-making process. On the other hand, online customer complaints, if not handled properly, could easily cause customers to lose loyalty for related products/services and create negative word-of-mouth. Thus, online customer feedback of products/service is useful for customer behavior analysis and is important for businesses. Customers can give their feedback in various websites and its difficult for product vendor/service provider or in case of our paper, hotel owner/manager to collect all reviews and analyze them. As a result, there is a growing need to extract and analyze customer opinions from large collections of online customer reviews. Recently, much effort has gone into automatic opinion mining, making it possible to obtain online customer opinions from various websites. However, visually examining and analyzing such mining results have not been well addressed in the past. Effective visual analysis of online customer opinions is needed, as it has a significant impact on building a successful business. In this paper, we present Opinion-Seer, an interactive visualization system that could visually analyze a large collection of online hotel customer reviews. The visual metaphor provides users with an integrated view of multiple correlations, allowing them to find useful opinion patterns quickly. Furthermore, it enables a fast side-by-side visual comparison of opinions of different customer groups, which is useful for finding out whether the opinions are influenced by a specific demographic factor.*

Keywords: Sentiment analysis, visual inspection, opinion mining

1. Introduction

In the proposed work extraction of reviews of hotel customers from various websites provided/selected by hotel manager/owner. Review comments are usually classified into three categories of positive, negative, and neutral. However, a positive review on an object does not always indicate that the opinion holder has positive opinions on all aspects or features of the examined object. To further obtain such detailed aspects, feature-level opinion mining has been proposed and extensively studied on product reviews to find opinions expressed on individual product features by visual inspections.

2. Requirement of Sentiment Analysis

Sentiment analysis aims to assist users to automatically detect relevant opinions within a large volume of review collection and create a coherent overview of these opinions. Review comments are usually classified into three categories of opinions: positive, negative, and neutral. Many approaches have been proposed to mine the overall opinion information at the document level or sentence level. However, a positive review on an object does not always indicate that the opinion holder has positive opinions on all aspects or features of the examined object. To further obtain such detailed aspects, feature-level opinion mining has been proposed and extensively studied on product reviews to find opinions expressed on individual product features.

3. Classification of Sentiment Analysis

In general, a sentiment analysis can be classified into three levels including the document level (review), the sentence level (semantic phrases) and feature (aspect)-based level. Compared to simple text summarizers, structured summarization of opinions has been formed according to feature based sentiment analysis, in which useful and relevant information will be available. We can extract features of the text, analyze the sentiment, integrate and summarize opinions by the developed ontology. Using the framework of the proposed ontology, the output results of the structured summarization will be presented in this paper

4. Proposed Methodology

Proposed work is a general analysis tool to analyze and detect hidden patterns in raw text data, and provide a user-friendly visual presentation to end users such as hotel managers. For hotel managers, the system allows them to identify useful and meaningful relationships quickly among vast amounts of textual data uploaded by customers on the e-channel, so that an effective decision can be better formulated to give timely and appropriate responses to the customers. Moreover, instead of inventing an unfamiliar visual representation, we augment familiar visual metaphors to convey the results from complex opinion analysis. Considering the analytical task and data characteristics of opinion mining, we combine the simplicity and familiarity of radial visualization, scatter plots.

The major contributions of work are as follows:

- Combine an opinion mining technique with subjective logic to model uncertainty in opinions and fuse the opinions
- Design a new visual representation with an integrated multidimensional view which can naturally encode the uncertainty information.

5. Visualization Technique for Data Extraction

There has been recent growing interest in visualizing opinions extracted from customer reviews posted online. These methods can be classified into two categories: document-level and feature-level opinion visualization. Document-level visualization focuses on visualizing opinion data at the document level. For example, Morinaga et al. suggested a 2D scatter plot called positioning map to show the group of positive or negative sentences. Gamon et al. derived a number of topics and estimated the average sentiment value for each topic. A Tree Map-style user interface called Pulse was designed to visualize the topics and their sentiment values.

6. Feature Level Opinion Visualization

Although the document-level opinion visualization provides a high level opinion overview of customer reviews, but not enough details are presented for users to understand customer opinions on certain product/service features (e.g., room, service, and price). With the development of feature-based opinion mining, visualization researchers have developed feature-level opinion visualization. Liu et al. proposed a method to extract feature-level opinions from customer reviews, and augmented traditional bar charts to facilitate visual comparison of extracted feature-level opinions. Oelke et al. introduced several visualization techniques including visual summary reports, cluster analysis, and circular correlation map to facilitate visual analysis of customer feedback data at the feature level. In addition, while existing methods do not consider the uncertainty of opinion extraction, our visualization approach explicitly accounts for uncertainty to reveal faithfully the underlying data.

7. Gaps in the previous work

From literature review it comes to know that, in existing techniques, there are lot of drawbacks. These drawbacks are summarized as;

- Previous works consider document level opinion mining which provides a high level opinion overview of customer reviews, but not enough details are presented for users to understand customer opinions on certain product/service features (e.g., room, service, and price).
- Issue of uncertainty remains unsolved.
- Previous literatures didn't implement visualization of customer feedback review comment and one who implemented are limited to basic.

8. Scope of the Project Proposed

As discussed above, existing works have some drawbacks summarized above. In our proposed work, we try to solve and give answer to these drawbacks in an effective manner.

The opinion-mining model in Opinion Seer is built on the method feature-level opinion mining, but is focused on visualizing the opinion mining results, which accounts for uncertainty to effectively model and analyze customer opinions. Moreover, we provide users with visual interaction tools to examine the results from multiple perspectives. To obtain detailed aspects, feature-level opinion mining has been proposed and extensively studied on product reviews to find opinions expressed on individual product features.

To address the need and to effectively communicate opinion-mining results and facilitate the analytical reasoning process, we designed and developed Opinion Seer. In this work, we propose a new feature-based opinion mining technique to faithfully model the uncertainty in the reviewed text. In addition, subjective logic is used to handle and organize multiple opinions with degrees of uncertainty. Moreover, instead of inventing an unfamiliar visual representation, we augment familiar visual metaphors to convey the results from complex opinion analysis. Considering the analytical task and data characteristics of opinion mining, we combine the simplicity and familiarity of radial visualization, scatter plots, and tag clouds while addressing their shortcomings, such as the lack of relationship analysis among multiple facets.

9. Extract features based review analysis

Different methods used to extract features in the review can be divided into five categories: 1) frequent nouns and noun phrases 2) based on relations between the feature and the opinion 3) supervised learning methods 4) topic modelling techniques and 5) hybrid methods. Most initial researches into extracting features from the document were based on nouns and relations between a feature and sentiment expressions.

After extracting reviews from various websites, rearrange them. Since, hotel industry is a seasonal business, reviews are also comes are seasonal. Hence, it's necessary to analyze the reviews by seasons. Therefore, first off let us classify the reviews. Now, feature extraction method is implemented. After extracting the features and determining the sentiment of reviews, obtained results are combined so as to produce a summary of opinions about various features. Hence, similar features in synonymous groups should be merged together and their correspondent sentiments should be averaged. Now, values from various reviews of respected features are combined together and averaged to form a single value. The extracted features must belong to any one of service provided by hotel. Hence next step is to group features and their value related to services provided by hotel. Now, we have values indicating performance of different service sectors such that departments of hotel.

Example:

Suppose our application extracts following reviews from different websites;

- Management staff was good but restaurant service was not impressive.
- The way food serve was good but it was not tasty.
- Lodging facility was good but restaurant service was not impressive.
- Lodging facility was bad; also taste of restaurant food was bad.
- Food serve in good manner and management staff was also good.

Table 1. shows how application must extract features and their corresponding value in following manner

Table 1 : Extract features and their values

Feature	Value (Review no)	Percentage / Satisfaction Level
Management Staff	Good (1) + Good (5)	$2 (+) / 2 = 100\%$ satisfaction
Restaurant Service	Not impressive (1) + Not impressive (3)	$2 (-) / 2 = \text{No satisfaction}$
Food Serve	Good (2) + Good (5)	$2 (+) / 2 = 100\%$ satisfaction
Food Taste	Not Tasty (2) + Bad (4)	$2 (-) / 2 = \text{No Satisfaction}$
Lodging Facility	Good(3)+Bad(4)	$1 (+) \text{Union} 1 (-) / 2 = 50\%$ satisfaction

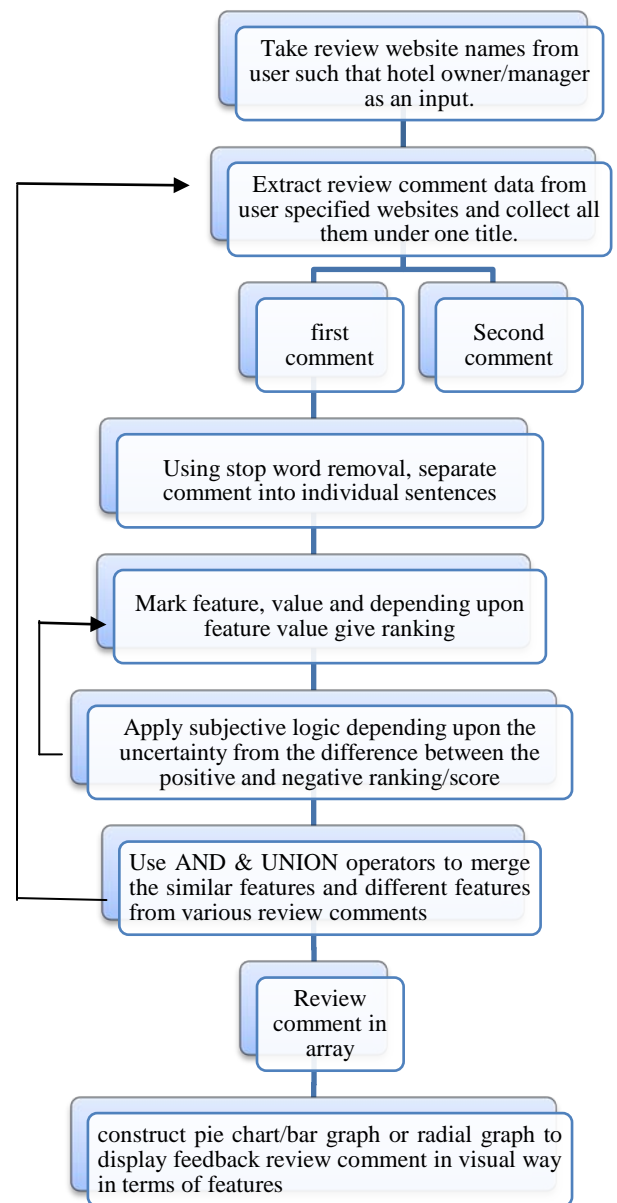
Now we have features and its value. Also, satisfaction level was formulated/calculated from various values obtained from reviews. But still it's not in good easily understandable format. Hence, visualizing is our last step to display result to hotel owner/manager in graphical format.

10. Implementation of /Methodology

Opinion Seer has two possible uses. Hospitality researchers can use it as a general analysis tool to analyze and detect hidden patterns in raw text data, and provide a user-friendly visual presentation for end users such as hotel managers. For hotel managers, the system allows them to identify useful and meaningful relationships quickly among vast amounts of textual data uploaded by customers on the e-channel, so that an effective decision can be better formulated to give timely and appropriate responses to the customers.

11. Basic Algorithm

Flow chart of Logical sequence



12. Conclusion

Future trends in data mining and visualization are becoming more evident with each new introduction of newer data mining solutions and data visualizations tools. Most of such advancements are based on Artificial Intelligence, such tools aims at intelligence like human. The aim is to replace the human aspect in decision-making. Data conditioning is a hopeful key to the amounts of data that is on the rise, which is of need to be mined. For visualization tools, interactivity is the new gesture, allowing users to touch, rotate, and select how to view data sets on its stir. Future work includes more challenges facing the development and wide spread use of data mining and visualization techniques. The current rising programmed data conditioning tools that provide a more effective dataset to be processed by the data mining.

References

- [1] Emitza Guzman, Padma Bhuvanagiri, Bernd Bruegge, "FAVe: Visualizing User Feedback for Software," 2014 Second IEEE Working Conference on Software Visualization, 978-0-7695-5305-4/14
- [2] Yingcai Wu, Furu Wei, Shixia Liu, Norman Au,
- [3] Weiwei Cui, Hong Zhou, and Huamin Qu, Member IEEE, "OpinionSeer: Interactive Visualization of Hotel Customer Feedback", accepted 1 August 2010 IEEE Conference on Software Visualisation
- [4] Abdul Wahab, Muhammad Faheem Khan, Aurangzeb Khan, Aziz Ullah Khan, "IMPORTANT FEATURES SELECTION DURING SENTIMENT ANALYSIS " Sci.Int(Lahore),26(2),961-966,2014.
- [5] Vijay B. Raut, Prof. D.D. Londhe, "Survey on Opinion Mining and Summarization of User Reviews on Web" International Journal of Computer Science and Information Technologies, Vol. 5 (2), 2014, 1026-1030
- [6] Mily Lal, Kavita Asnani, "ASPECT EXTRACTION & SEGMENTATION IN OPINION MINING", International Journal Of Engineering And Computer Science, Volume 3 Issue 5 May, 2014 Page No. 5873-5878
- [7] Yee W. Lo, Vidyasagar Potdar, "A review of opinion mining and sentiment classification framework in social network", 2009 3rd IEEE International Conference on Digital Ecosystems and Technologies, 978-1-4244-2346-0/09
- [8] G.Angulakshmi, Dr.R.ManickaChezian, "An Analysis on Opinion Mining: Techniques and Tools", International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 7, July 2014, 2319-5940
- [9] Gayathri Deepthi. V, K. Sashi Rekha, "Opinion Mining and Classification of User Reviews in Social Media ", International Journal of Advance Research in Computer Science and Management Studies, Volume 2, Issue 4, April 2014.
- [10] Walaa Medhat, Ahmed Hassan, Hoda Korashy, "Sentiment analysis algorithms and applications: A survey", 2014 Production and hosting by Elsevier B.V. on behalf of Ain Shams University.
- [11] Neha S. Joshi, Suhasini A. Itkat, "A Survey on Feature Level Sentiment Analysis", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4), 2014, 5422-5425
- [12] Abdullah Dar, Anurag Jain, "Opinion Mining through on to Tree", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 6, June 2014.
- [13] Poonam S. Patil, Rajesh. N. Phursule, "Survey Paper on Big Data Processing and Hadoop Components", International Journal of Science and Research, Volume 3 Issue 10, October 2014, Paper ID: OCT14251
- [14] Nikhil Nagnath Boriwale, Satish .R.Todmal, "A Survey of Forensic Analysis on Document Clustering", International Journal of Science and Research, Volume 3 Issue 11, November 2014, Paper ID: OCT141295
- [15] Chaitrali T. Chavan, Prof. Vinod S. Wadne, "Categorization and Detection of Adaptive Novel Class of Feature Evolving Data Streams" International Journal of Advanced Engineering Research and Science (IJAERS), Vol-1, Issue-4, Sept.- 2014.
- [16] Katre Ketaki Ganesh, Prof. Devendra P. Gadekar, "Survey paper on string transformation ", International Journal of Computer Engineering and Applications, Volume VIII, Issue III
- [17] Deepali Virmani, Vikrant Malhotra, Ridhi Tyagi, "Sentiment Analysis Using Collaborated Opinion Mining", International Journal of Soft Computing and Engineering (IJSCE), Volume-4, Issue-ICCIN-2K14, March 2014.