Improving Conclusion of Product Quality Using Opinion Mining Considering Customer Evaluation

Ankit Nagesh Limkar

M. Tech Student, Department of Computer Engineering, Swaminarayan College of Engineering & Technology,

Kalol, India

ABSTRACT: Customer gives review in terms of opinions and it plays a vital role in current scenario. Today's customers consider other individuals' opinions while purchasing a new product, who already have purchased & used them before. After purchase of a new product, customers post their opinions in terms of reviews and ratings, for many products through review sites, blogs, and social networking sites. Corporate sector and large scale business organizations are always keen to find and observe customer trends and demands to make certain decision of their products, support and services. Decision is very crucial for E- Commerce, Online shopping and tourism to analyse bulky social data present on the web. So, it is very much important to create methods that automatically classify them and help us to take decision instantly. Sentiment Analysis, reviews, views, ratings, emotions and opinions are also considered in analysis of Opinion Mining from text, big data and speech by using various methods. In this thesis, we are going to see how mining algorithm can be used to analyse the reviews posted by the online customers considering their reviews, ratings etc. Our main goal is to create a system for quicker analysis of opinions which foster judgment and decision making capabilities of different consumer products, support and services.

KEYWORDS: Opinion Mining, Mining Algorithm, Online Reviews, Sentiment Classification, SentiWordNet, Min-Max Normalization

I. INTRODUCTION

Opinion mining dissects individual opinions, sentiments, assessments, mentality, and feelings from written text. It has pulled in a number of analysts from distinctive areas of exploration including NLP, information mining, machine learning, phonetics, and even social science.
A bigger portion of content handling systems works with authentic data. The enormous volumes of opinionated content hold by the web. Web users express particular emotions and opinions on nearly anything at review sites, blogs, and forums and so on. This important data is freely accessible for internet clients. The substantial gathering of opinions on the web makes it extremely tough to get helpful data effectively. Perusing all reviews and emotions to settle on an educated choice is a much time taking task. Perusing distinctive and potentially even conflicting opinions composed by diverse commentators may make organizations, users and customers more confused. These needs have made another line of examination on mining user and customer opinions, which is called opinion mining.

II. OBJECTIVE

The objective of this paper is to design a mechanism to get an opinion about the consumer product. In case of mining opinion of any product or service the decision making part is crucial. In e-commerce, online shopping and online tourism, it is very crucial to analyse the good amount of social data present on the web. In decision making one of the most essential things is sentiment classification considering analysis of reviews, views and emotions. A frequent item set mining algorithm will be used for mining reviews from online reviews those are posted by customers. The main theme is to create a system for analysing opinions which implies judgement of different consumer products.

III. IMPLEMENTATION METHOD

Following are the steps for proposed approach for performing opinion mining and what techniques and algorithms we are going to use to perform Opinion Mining and Sentiment Analysis for getting useful information from online customer reviews.

A. *Extract nouns, adjectives, verbs and adverbs:*

To perform this step, we are using dictionary approach. After this step we get all the Nouns, Adjectives, Verbs and Adverbs that are present in the customer reviews file. We use these words for getting the frequent words in the next step.

B. *Identify frequent words by using frequent item set mining algorithm:*

In this step, we get frequent words by using the frequent item set mining algorithm other than Apiori algorithm.

C. *Sentiment Analysis on the frequent words using SentiWordNet:*

In this step we perform Sentiment Analysis on the frequent words that we got from frequent item set mining algorithm by using SentiWordNet. It provides a value for each and every word.
Sentiment Analysis deals with the usage of automated techniques for anticipating the introduction of subjective substance on text reviews or comments, with usage in various fields that includes recommendation system and advertising, user intelligence and opinion retrieval. SentiWordNet is an opinion vocabulary and can be considered as extended from the Wordnet database where each one term is connected with numerical scores demonstrating positive and negative sentiment data. This examination shows the consequences of applying the SentiWordNet lexical asset to the issue of automated sentiment arrangement of customer film reviews or comments.

D. *Min-max normalization:*

Min-Max normalization is the technique of taking data calculated in its own units and converting it to a value between 0 and. We use normalization because star ratings values lies between 1 to 5 and word polarity of SentiWordNet values lies between -1 and +1.
Suppose we have some n rows with five variables, A, B, C, D and E, in the data. We use variable B as an example for understanding the normalization concept in the calculations below. All the other variables in the rows are normalized in the similar way.
The normalized value of Bi for variable B in the ith row is calculated as:
normalised $(Bi) = Bi – Bmin / Bmax− Bmin$
where,
Bmin = the least value for variable E
Bmax = the highest value for variable E
If Bmax = Bmin then Normalized (Bi) is set to 0.5

E. *Provide Visualization:*

Finally visualization is provided as a conclusion of following the above steps.

IV.   CONCLUSION

The need of customer reviews and feedbacks has become extremely important. Opinion Mining is an area to consolidate the scattered data of opinions from social media and ecommerce as well as review sites. The vital phase in opinion mining is identifying frequent patterns by using frequent item set mining algorithms. So far, Apriori algorithm has been used widely for this phase. The main aim of this research work is to use other efficient algorithm to enhance and improve efficiency such that it provides speedier convergence rate and compare the results produced afterwards to prove that the algorithm satisfies the objective of this research.

REFERENCES

1. Lu´ıs Sarmento, Paula Carvalho, M´ario J Silva, and Eug´enio de Oliveira. Automatic creation of a reference corpus for political opinion mining in user-generated content. In Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion, pages 29–36. ACM, 2009.
2. John Krumm, Nigel Davies, and Chandra Narayanaswami. User-generated content. IEEE Pervasive Computing, 7(4):10–11, 2008.
3. Bo Pang and Lillian Lee. Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2):1–135, 2008.
4. Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 168–177. ACM, 2004.
5. Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of the conference on human language technology and empirical methods in natural language processing, pages 347–354. Association for Computational Linguistics, 2005.
6. Alexander Pak and Patrick Paroubek. Twitter as a corpus for sentiment analysis and opinion mining. In LREC,2010.
7. Andrea Esuli. Automatic generation of lexical resources for opinion mining: models, algorithms and applications. In ACM SIGIR Forum, volume 42, pages 105–106. ACM, 2008.
8. Richa Sharma, Shweta Nigam, and Rekha Jain. Supervised opinion mining techniques: A survey.
9. Akshat Bakliwal. Fine-grained opinion mining from different genre of social media content. 2013.
10. Samaneh Moghaddam and Martin Ester. Aspect-based opinion mining from online reviews. In Tutorial at SIGIR Conference, 2012.
11. Bo Pang and Lillian Lee. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In Proceedings of the 42nd annual meeting on Association for Computational Linguistics, page 271. Association for Computational Linguistics, 2004.
12. Ellen Riloff, Janyce Wiebe, and William Phillips. Exploiting subjectivity classification to improve information extraction. In Proceedings of the National Conference On Artificial Intelligence, volume 20, page 1106. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2005.
13. Stephan Raaijmakers and Wessel Kraaij. A shallow approach to subjectivity classification. In ICWSM, 2008.
14. Janyce M Wiebe, Rebecca F Bruce, and Thomas P O'Hara. Development and use of a gold-standard data set for subjectivity classifications. In Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics, pages 246–253. Association for Computational Linguistics, 1999.
15. Anindya Ghose and Panagiotis G Ipeirotis. Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. Knowledge and Data Engineering, IEEE Transactions on, 23(10):1498–1512, 2011.
16. Nitin Jindal and Bing Liu. Opinion spam and analysis. In Proceedings of the 2008 International Conference on Web Search and Data Mining, pages 219–230. ACM, 2008.
17. Nitin Jindal and Bing Liu. Review spam detection. In Proceedings of the 16th international conference on World Wide Web, pages 1189–1190. ACM, 2007.
18. Wei Zhang, Clement Yu, and Weiyi Meng. Opinion retrieval from blogs. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management, pages 831–840. ACM, 2007.
19. Lun-Wei Ku, Yu-Ting Liang, and Hsin-Hsi Chen. Opinion extraction, summarization and tracking in news and blog corpora. In AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, volume 100107, 2006.
20. Hoa Trang Dang and Karolina Owczarzak. Overview of the tac 2008 opinion question answering and summa- rization tasks. In Proc. of the First Text Analysis Conference, 2008.
21. Murthy Ganapathibhotla and Bing Liu. Mining opinions in comparative sentences. In Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1, pages 241–248. Association for Computational Linguistics, 2008.
22. Charu C Aggarwal and ChengXiang Zhai. Mining text data. Springer, 2012.
23. Xiaowen Ding, Bing Liu, and Philip S Yu. A holistic lexicon-based approach to opinion mining. In Proceedings of the 2008 International Conference on Web Search and Data Mining, pages 231–240. ACM, 2008.
24. Raffaele Perego, Salvatore Orlando, and P Palmerini. Enhancing the apriori algorithm for frequent set counting.In Data Warehousing and Knowledge Discovery, pages 71–82. Springer, 2001.
25. Christian Borgelt and Rudolf Kruse. Induction of association rules: Apriori implementation. In Compstat, pages 395–400. Springer, 2002.
26. Bruno Ohana and Brendan Tierney. Sentiment classification of reviews using sentiwordnet. In 9th. IT & T Conference, page 13, 2009.
27. Andrea Esuli and Fabrizio Sebastiani. Sentiwordnet: A publicly available lexical resource for opinion mining. In Proceedings of LREC, volume 6, pages 417–422, 2006.

28. Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In LREC, volume 10, pages 2200–2204, 2010.
29. Anil Jain, Karthik Nandakumar, and Arun Ross. Score normalization in multimodal biometric systems. Pattern recognition, 38(12):2270–2285, 2005.
30. Soo-Min Kim and Eduard Hovy. Determining the sentiment of opinions.   In Proceedings   of the 20th international conference on Computational Linguistics, page 1367. Association for Computational Linguistics, 2004.

BIOGRAPHY

Ankit Limkar completed his Bachelors in Computer Engineering in 2008 and currently pursuing Masters in Computer Engineering from Gujarat Technological University. He has 2 years of industry experience and 5 years of academic experience.