**International Journal of Management, IT & Engineering**
Vol. 8 Issue 9, September 2018,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

# PREDICTION OF ASPECT CATEGORY ON CUSTOMER REVIEWS USING ASSOCIATION RULE MINING

**Miss.Gayatri D.Khot***

**Mr. Hemant. A. Tirmare****

**Abstract**

Now a days online consumer review is a most powerful tool for decision making. This term serves as electronic word of mouth which become increasingly popular. Millions of people are now buying products and services via online. Web services are provided this feature to users openly. The web can provide an extensive source of consumer reviews. The user can read all the reviews and evaluate fair view of product or service. This can apply only to a limited number of reviews presented on the web. The web contain more than hundreds of reviews then problem arrived and time consuming also. A text processing framework is desirable which summarize all the reviews. This framework would find out general aspect category addressed in all review sentences. The method presented in this framework which applies association rule mining on co-occurrence frequency data to find out these aspect categories. From this result, generate polarity score for each aspect category. This polarity score helps to evaluate fair decision making for the customer as well as the company. The graph representation is also provided by the system for quickly evaluate the decision for products or services provided by the web.

***Keywords:***

consumer reviews;

aspect category;

co-occurrence

frequency data;

polarity score;

## 1. Introduction

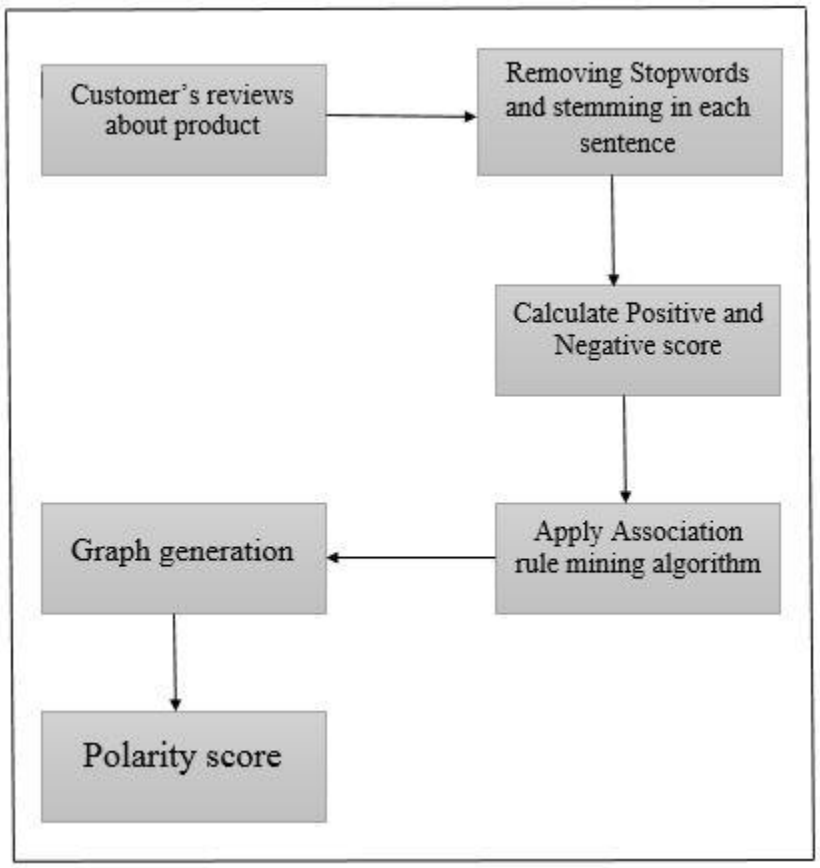
In the present association, clients wish to buy the item or products by on the web. Informal exchange is a conventional method which gives an exhortation and suggestions about any online item or products by orally. Family as well friends are principle key elements for this conventional strategy. So these works were not made in a point of view for expansive dataset. So there is requirement for framework which gives a yield such that clearify the perspective of item into clients mind. So this will build the profitability and also execution of item or products. This system serves to client and additionally organizations.

Lately there has been an in the long run development in enthusiasm from brands, organizations and scientists in Sentiment Analysis and its application to business applications. Associations have two kinds of client feeling or criticism information that they measure, store, and break down: organized and unstructured information. Organized information will be data that is obviously characterized and simple to give an account of. It is the sort of information that is by and large found in an overview and can be composed in a spreadsheet: name, area, age, and rating (3 out of 5 stars, for instance, or a 10 for "most fulfilled" versus a 1 for "minimum satisfied").Unstructured information as it exists today seems to be, essentially, content, in spite of the fact that it can likewise incorporate other media, for example, sound, photographs, or recordings. Unstructured information can be caught in an email, the "extra remarks" area of an overview, voice accounts of client associations, a post on a client survey site, in web-based social networking, call focus notes, talk transcripts, and many different spots. Examining this information accurately is basic, since it uncovers everything from purchasing patterns to item imperfections and gives a huge business advantage. Associations frequently battle to do this examination, be that as it may, in light of the fact that unstructured information is essentially harder to order and cover than organized information. It can be difficult to parse because of syntactic mistakes or slang, it regularly contains different inconsequential thoughts, and it can speak to different levels of notion identified with every thought.

System Architecture:

In this system, take input as customer’s reviews, process on that to removes stopwords and stemming words i.e. less frequency words. Output given to aspect category detection module to

find out the aspect category for each statements.The sentimental analysis module after this generate polarity score.




**System Architecture**

There are following modules in our proposed system.

a. Data Accumulation

b. Preprocessing of data

c. Transformation

d. Aspect Category

e. Evaluation

**Module1:Data Accumulation**

getInputFile() method is used to take the file from storage. This method is available in java. This module take input as different statements on which system able to process and analyze.

**Module2:Preprocessing of data**

This module removes stopwords and streeming words from customer reviews statements. Stopwords are like a, an, am, and, are, as, at, be, been, both, did, do, so, some, was etc. Streeming words like ed, ing, ation , lly etc. All stopwords and streem words are removed from reviews so delection of this words are not affected on system as well as processing time is also minimized.

**Module3:Transformation**

In the transformation process, the score for each sentence is calculated in thepresented documents.This score helps to detect the each positive or negative scoring.

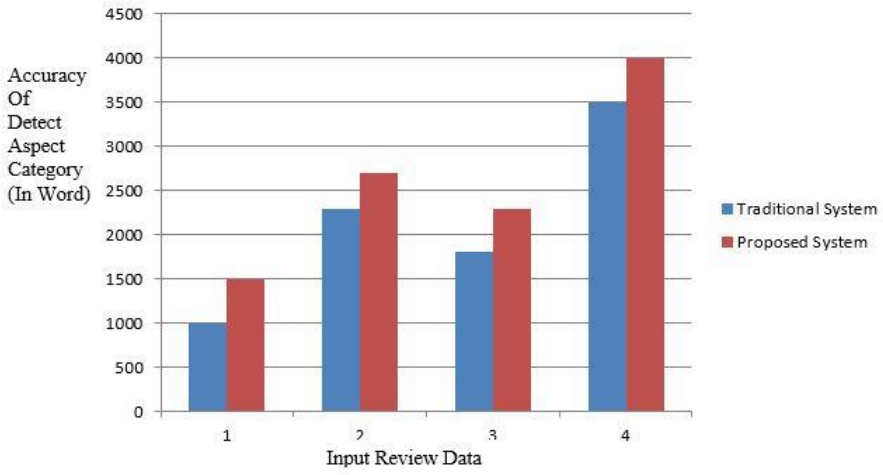
**Module4:Aspect Category**

Apriori association rule mining algorithm is used to find this relationship. This algorithm generates config value for each relation exist for particular input categories.

**Module5:Evaluation**

It gives the review result in graph.The representation of scoring is important on every framework. This helps to improve accuracy of system.There are three graph generates first is polarity score of each reviews, second graph is aspect category detection score and third is positive or negative scoring graph.

## 3. Results and Analysis

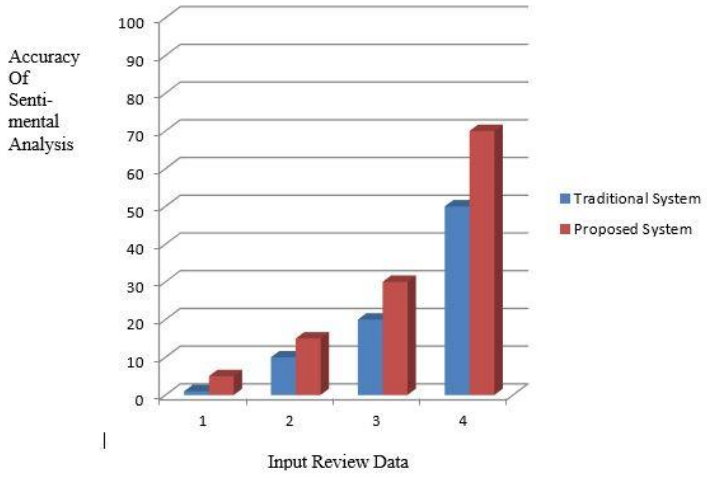


Comparison of Traditional System and Proposed System for Aspect Category

| Review Data | Aspect categories in traditional System | Aspect categories in proposed System |
|---|---|---|
| 1 | 1000 | 1500 |
| 2 | 2300 | 2700 |
| 3 | 1800 | 2300 |
| 4 | 3500 | 4000 |

Aspect category comparison

Above graph representation of traditional verses proposed system shows how the aspect category is detected in two different frameworks.Prediction of aspect category is one of biggest task in sentiment analysis.As per traditional system they only shows the statements i.e. reviews which is written by customers is processed.System directly apply the process and get the output, it is time consuming process.Because directly predict the aspect not efficient method. In proposed system, first framework anlyze the all statements and removes then further procedure begins.



Comparison of Traditional System and Proposed System for Sentimental Analysis

| Review Data | Positive/Negative scoring of traditional System | Positive/Negative scoring of proposed System |
|---|---|---|
| 1 | 1 | 5 |
| 2 | 10 | 15 |
| 3 | 20 | 30 |
| 4 | 50 | 70 |

Sentimental analysis comparison

Detetcting a positive and negative scoring of each reviews is the task for framework.From this resuly customer as well as companies improves their decision making power. In traditional system, less amount of scoring given for one document file.With the help of this system get overall percentage of all reviews simultaneously.

**4. Conclusion**

The issue of previous system is only for limited number of reviews presented on it. The web contain more than hundreds of reviews then problem arrived and time consuming also. The proposed system designs a text processing framework. This framework find out general aspect category addressed in all review sentences. Each sentence is checked by the framework and system gives an appropriate category for particular review. The method presented in this framework which applies association rule mining on data to find out these aspect categories. The system is for both structure sentences and unstructured sentences. For this type of sentences, positive and negative scoring of each sentences is needed. So system is able to find out whether sentence is positive or negative. From this result, generate polarity score for each sentences. This polarity score helps to evaluate fair decision making for customer as well as company.

**References**

[1] Kim Schouten, Onne van der Weijde, Flavius Frasincar, and Rommert Dekker,"Supervised and Unsupervised Aspect Category Detection for Sentiment Analysis With Co-Occurrence Data" IEEE TRANSACTIONS ON CYBERNETICS, Volume: PP, Issue: 99,April 2017.

[2] Deepa Ananda, Deepan Naorem," Semi-supervised Aspect Based Sentiment Analysis for Movies using Review Filtering" 7th International conference on Intelligent Human Computer Interaction, IHCI 2015.

[3] Walaa Medhat , Ahmed Hassan, Hoda Korashy," Sentiment analysis algorithms and applications: A survey", Ain Shams Engineering Journal (2014) 5, 1093–1113.

[4] Aitor Garc__a-Pablos, Montse Cuadros, German Rigau," W2VLDA: Almost Unsupervised System for Aspect Based Sentiment Analysis", July 19, 2017.

[5] Xinjie Zhou, Xiaojun Wan and Jianguo Xiao,"Representation Learning for Aspect Category Detection in Online Reviews", Twenty-Ninth AAAI Conference on Artificial Intelligence.

[6] Doaa Mohey El-Din Mohamed Hussein," A survey on sentiment analysis challenges", Journal of King Saud University – Engineering Sciences (2016),

[7] Md Shad Akhtar, Asif Ekbal, and Pushpak Bhattacharyya," Aspect Based Sentiment Analysis: Category Detection and Sentiment Classification for Hindi", Indian Institute of Technology Patna India-801103.

[8] Alessia D'Andrea,Fernando Ferri,Patrizia Grifoni,Tiziana Guzzo,"Approaches, Tools and Applications for Sentiment Analysis Implementation", International Journal of Computer Applications (0975 – 8887) Volume 125 – No.3, September 2015.

[9] Samuel Brody, Noemie Elhadad," An Unsupervised Aspect-Sentiment Model for Online Reviews", The 2010 Annual Conference of the North American Chapter of the ACL, pages 804–812.

[10] Chetashri Bhadanea,Hardi Dalalb, Heenal Doshic," Sentiment analysis: Measuring opinions", International Conference on Advanced Computing Technologies and Applications (ICACTA-2015).

[11] Emma Haddi, Xiaohui Liu, Yong Shi," The Role of Text Pre-processing in Sentiment Analysis", Information Technology and Quantitative Management (ITQM2013).

[12] Thellaamudhan C, Suresh R, Raghavi P," A Comprehensive Survey on Aspect Based Sentiment Analysis",International Journal of Advanced Research in Computer Science and Software Engineering,Volume 6, Issue 4, April 2016.

[13] Sujata L. Sonawane, Pallavi V. Kulkarni, Extracting Sentiments from Reviews: A Lexicon-Based Approach",IEEE2017

[14] Jyoti S Deshmukh, Amiya Kumar Tripathy," Text Classification using Semi-supervised Approach for Multi Domain", 2017 International Conference on Nascent Technologies in the Engineering Field (ICNTE-2017).

[15] Qing Sun , Jianwei Niu , Zhong Yao , Dongmin Qiu,"Research on Semantic Orientation Classification of Chinese Online Product Reviews Based on Multi-Aspect Sentiment Analysis", Big Data Computing Applications and Technologies (BDCAT), 2016 IEEE/ACM 3rd International Conference.

[16] Mondher Bouazizi, Tomoaki Ohtsuki,"A Pattern-Based Approach for Multi-Class Sentiment Analysis in Twitter",IEEE Access (Volume: PP, Issue: 99 ).

[17] Nipuna Upeka Pannala, Chamira Priyamanthi Nawarathna , J.T.K.Jayakody,Lakmal Rupasinghe , Kesavan Krishnadeva," Supervised Learning Based Approach to Aspect Based Sentiment Analysis", 2016 IEEE International Conference on Computer and Information Technology.

[18] Y. Tang, Y.-Q. Zhang, N. V. Chawla, and S. Krasser, "SVMs modeling for highly imbalanced classification," IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 39, no. 1, pp. 281–288, Feb. 2009.

[19] T. Brychcin, M. Konkol, and J. Steinberger, "UWB: Machine learning approach to aspect-based sentiment analysis," in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin, Ireland, 2014, pp. 817–822.

[20] A. Garcia-Pablos, M. Cuadros, S. Gaines, and G. Rigau, "V3: Unsupervised generation of domain aspect terms for aspect based sentiment analysis," in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin, Ireland, 2014, pp. 833–837.

[21] R. Feldman, "Techniques and applications for sentiment analysis,"Commun. ACM, vol. 56, no. 4, pp. 82–89, 2013.

[22] S. Sen and D. Lerman, "Why are you telling me this? An examination into negative consumer reviews on the Web," J. Interact. Marketing, vol. 21, no. 4, pp. 76–94, 2007.

[23] D. Smith,S.Menon, and K.Sivakumar, "Online peer and editorial recommendations, trust, and choice in virtual markets," J. Interact.Marketing, vol. 19, no. 3, pp. 15–37, 2005.

[24] A. Katifori, C. Vassilakis, and A. Dix, "Ontologies and the brain: Using spreading activation through ontologies to support personal interaction," Cognitive Syst. Res., vol. 11, no. 1, pp. 25–41, 2010.

[25] SemEval-2014 dataset which contain both training sets and test sets [Online].Available:https://github.com/pedrobalage/SemevalAspectBasedSentimentAnalysis/tree/master/semeval_data

[26] Linguist. Syst. Demonstrations, 2014, pp. 55–60. [Online]. Available: http://www.aclweb.org/anthology/P/P14/P14-5010.

[27] X. Zheng, Z. Lin, X. Wang, K.-J. Lin, and M. Song, "Incorporating appraisal expression patterns into topic modeling for aspect and sentiment word identification," Knowl. Based Syst., vol. 61, no. 1, pp. 29–47,2014.

[28] G. Castellucci, S. Filice, D. Croce, and R. Basili, "UNITOR: Aspect based sentiment analysis with structured learning," in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin, Ireland, 2014,pp. 761–767.

[29] K. Schouten and F. Frasincar, "Survey on aspect-level sentiment analysis," IEEE Trans. Knowl. Data Eng., vol. 28, no. 3, pp. 813–830,Mar. 2016.

[30] Wikipedia. "Fundamental human needs". Internet: https://en.Wikipedia .org/wiki/Fundamental_human_needs [Apr. 29, 2016].

[31] 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.

[32] B. Bickart and R. M. Schindler, "Internet Forums as Influential Sources of Consumer Information," Journal of Interactive Marketing, vol. 15, no. 3, pp. 31–40, 2001.