International Journal of Advanced Science and Technology
Vol. 29, No. 7s, (2020), pp. 5094 - 5102

Machine Learning Approach Based on Sentiment Analysis for Effective Managerial Decision on Product Quality Management and Customer Satisfaction

[1] Saumya Chaturvedi, [2]Vimal Mishra, [3]Nitin Mishra

[1]*Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India,*
E-mail: saumyanmishra5@gmail.com

[2]*Director IERT Allahabad , India,*
E-mail: vimalmishra@iert.ac.in

[3]*Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India,*
E-mail: drnitinmishra10@gmail.com.

Abstract

This paper narrates the application of Machine Learning Approach based on Sentiment Analysis for product quality management and enhanced business process decision. It is based on customer opinion analysis which is built over customer review posted on the internet in natural language. We have used R studio to implement machine learning algorithms for sentiment analysis. We have contributed two algorithms for natural language processing. We have also proposed a framework for sentiment analysis. Machine Learning, web data extraction and natural language processing are used for the understanding of customer review analysis. In this paper, specific internet resources such as imdb.com, amazon.com, and yelp.com are used for customer review sentiment classification. The attributes are text sentences extracted from reviews of products, movies, and restaurants. The review analysis is utilized in making a managerial decision on product quality in accordance with customer satisfaction. The satisfaction is measured over two scales such as sentiment towards the product and aspect based sentiment. We have evaluated the efficiency of machine learning methods for sentiment classification and analysis. An intelligent decision system is proposed for quality management based on customer sentiment indicating satisfaction. The proposed approach is seen as a prototype of the decision system which will help in taking an effective managerial decision on product quality management according to customer feedback.

Keywords: *Sentiment Analysis,Machine Learning,customer satisfaction research, Natural Language processing, Business Intelligence and Analytics, decision support system, Aspect-based sentiment analysis.*

1 Introduction

In order to deliver a quality product a company should always make effective decisions. There are typically six steps associated with effective decision process no matter whether a decision is manual or automated. Managers accost decisions requirement in the form of either a problem or an opportunity. For making efficient managerial decision a manager required information available to make the decisions and depth knowledge of information analysis.An organization should device a process of automatic collection of Information and its additional proceeding and analysis. Decision making should be founded on the understanding and assumptions obtain during the analysis of gathered data.

Figure 1 Six steps of decision process and Quality Assurance vs Quality Control

The Total Quality Management (TQM) is a management philosophy committed to a focus on continuous improvements of product and services with the involvement of the entire workforce.Quality Assurance is presently attained through a process approach based on the model of a quality management system (QMS)[1](see figure 2). The QMS model describes the interaction of the company and its customer during the process of product manufacturing and consumption.The objective of this paper is to develop decision support approach for Quality Management guided by sentiment analysis of customer reviews on internet by using machine learning techniques. It will help in taking better managerial decisions such as, which aspect should keep in the product? what areas need changes?

2 Related work

Text mining for analysis of customers review posted online is not new. There are many studies concerning Tools and techniques for data collection sentiment analysis and information extraction. A lexicon-based method predicts classification of reviews with 90 percent accuracy . Lu et .al[2] and Jo and Oh [3] focused in the work on the problems of automatic product aspect identification.For automatic analysis of reviews some social monitoring systems and frameworks have been developed recently. Liu et .al [4] presented Framework named as opinion observer for analyzing and comparing consumer opinions of competing products. Ganu et .al [5] focused on an analysis of free-text reviews by means of classification of reviews at the sentence level, with respect to both the topic and the sentiment

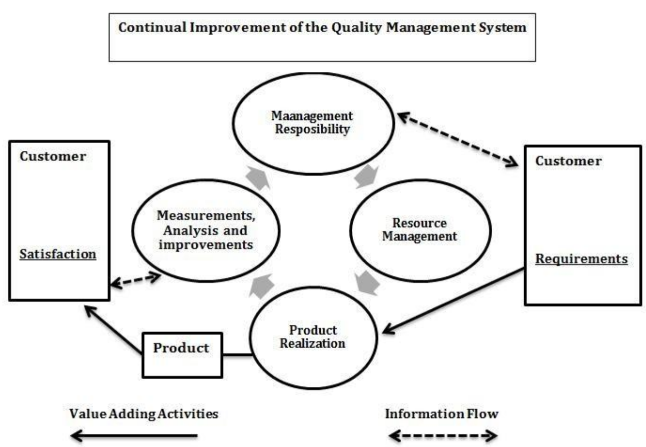


Figure 2 Process based Quality Management System model

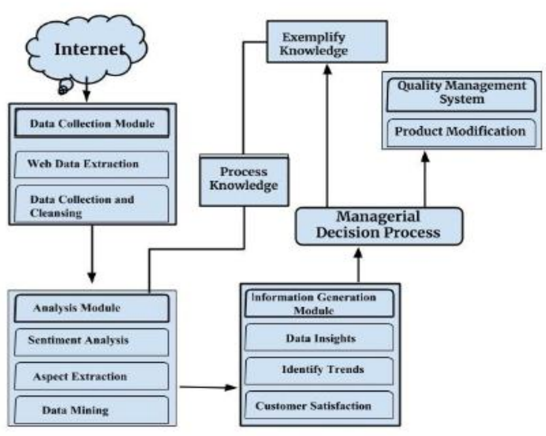
expressed in the sentences.Bjørkelund et .al [6] described how the results of sentiment analysis of textual reviews can be visualized using Google Maps, providing possibilities for users to easily detect good hotels and good areas to stay in. In this paper we focus on sentiment identification, its classification and proposed a model which can utilize this knowledge to take decision for the product quality.

3 Methodology and Framework

The suggested methodology for effective managerial decisions on product quality management is accomplished by collecting the data from web using intelligent decision system based on natural language processing. The architecture of proposed methodology is presented in figure 3. The projected framework in figure 4 offers sentiment analysis on customer review data in a better way by using S/A convertor before applying the trained machine learning model to unseen review data. The framework gives better results in terms of prediction accuracy on unseen data even when applied to a variety of data from different domains.

3.1 Training and Deployment Module

Labeled data is very important to apply supervised flavor of machine learning.Initially, the reviews are collected which are termed as Raw Corpora Data. The sentiment label extraction is performed using natural language processing on this raw corpus data. The Training data is feed to the training module for learning. The classifier learns according to the learning algorithm with the help of training data. Finally, it generates the trained model. The testing data portion from labeled data is kept and utilized to test the trained model that is built. This is termed as the evaluation of the model.In the deployment module we have added SA convertor for classification of a variety of data for which the model has not been trained. In general, when the classifier is trained for a word “good” and if “excellent” appears in the text the model is not able to classify. It deteriorates the performance of the machine learning model on real and new unseen data. SA convertor alters the unseen patterns to their



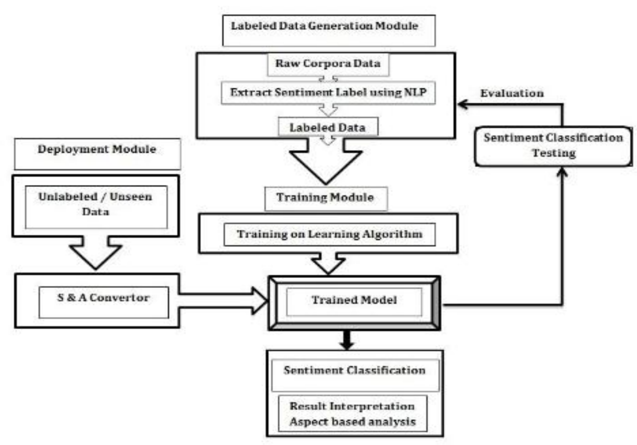


Figure 3 Architecture of Intelligent Decision System

Figure 4 Framework for Sentiment Analysis known synonyms and if no synonym is available then antonym is utilized with negation and sentiment classification is done.

4 Applied Machine Learning Techniques

In this section the implemented machine learning techniques for customer review analysis and sentiment analysis for support of decision making over quality of products are described. We have mainly examined Support Vector Machine (SVM) and Naive Bayes machine learning algorithm with its variants according to framework proposed. For experimentation purpose we have taken sentences from different websites imdb.com, amazon.com and yelp.com. for each website we have taken 500 positive and 500 negative sentences. Those sentences were selected randomly from larger data sets of reviews. The data set contains sentences labelled with positive and negative sentiment. The attributes of data are text sentences, extracted from reviews of products , movies and restaurants.

4.1 Sentiment Analysis (SA)

Sentiment analysis is the most needed step after data collection. The Sentiment analysis process broadly have two approaches. First, Lexicon- based approach which works on subjective lexicons. Second, Machine Learning approach. In Lexicon based approach score is linked with each word indicating the nature of text such as positive, negative and neutral. The aggregation of scores of all subjective words is estimated. The highest score will decide the polarity of the text [7].Lexicon -based approach is subdivided into two further approaches (a) Dictionary based approach in which a set of opinion words are manually collected to form a list.(b) Corpus based approach which work over collection written text generally on some specific area. In this method list of seed is prepared and expanded with the help of corpus text [8]. so this method works on domain oriented text. In our work we have used supervised machine learning approach.In this paper we have focused on naive Bayesian classifier and Support Vector Machines. The realization of these techniques is based on as described by Pang and Lee in their work [9].

4.2 Aspect Based SA and Customer Satisfaction Measures

Sentiment analysis of reviews provide a crisp evaluation of customer satisfaction with product. But, it does not clearly inform about what customers like about the product and what they do not like about the product.Aspect based sentiment analysis is more difficult task. Frequency-based aspect extraction is performed as defined by Scaffidi et al. in their work [10]. The customer satisfaction for product or service can be calculated by equation number (1) and customer satisfaction for product's or service's aspect group can be calculated by equation number (2).

$$CS = \frac{R^{pos}}{R^{pos} + R^{neg}} .100 \qquad (1)$$

$$CS_{asp} = \frac{R^{pos}_{asp}}{R^{pos}_{asp} + R^{neg}_{asp}} .100 \qquad (2)$$

asp + *Rasp*

5 Experiments and Results

For measuring efficacy of the designed intelligent decision system framework, we have taken review data from three different domains which contain reviews of products,movi6

and restaurants. For training and building effective sentiment classifier, we have evaluated the classification accuracy with the help of a confusion matrix shown in the table 1.

Table 1 Confusion Matrix summarizing

	Event	**No-**
Eve	*true*	*false*
No-	*false*	*true*

We have compared machine learning methods for sentiment classification with tagging technique and conversion technique.After implementation of framework for designing of model and predicting the sentiment classification for 10 no of runs we have observed result accuracy as depicted in table 2.The comparison of accuracy of methods for sentiment analysis is shown in figure 5. We can observe that the maximum accuracy is attained by Naive Bayes with tagging and sA convertor. We have used results obtained by Pang and Lee for movie-review domain as a baseline. Even with the variation in training data we found Naive Bayes classifier with tagging and sa convertor most efficient machine learning technique.

Table 2 Comparison of methods for sentiment classification

No of runs	S V	N a	SVM (with Tagging and	Naïve Bayes (with Tagging and
1	7	9	93.1	98.2
2	8	8	95.2	99.9
3	8	8	85.6	94.5
4	8	8	89.2	89.2
5	8	8	92.3	90.6
6	9	9	89.1	90.6
7	8	8	81.1	85.5
8	8	8	86.3	88.2
9	8	8	92.6	85.1
10	8	8	90	93.1

Using the aspect extraction algorithm we have identified certain aspects of object "restaurant". The extracted aspects and their sentiment terms with their polarity or orientation is depicted in figure 6. So we can observe the aspects of a restaurant extracted in our experiments are "ambience", "staff", "food" and "service".

6 Conclusion and future work

Deprived quality of products and services always contributes to a lessening of customer satisfaction.Therefore, it is very important to maintain high-quality standards of products and services. For maintaining these standards managerial decisions can be made with the help of opinion or review analysis of feedback expressed by customers in natural

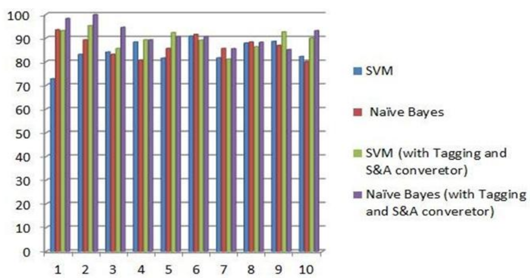


Figure 5 Comparison of accuracy of methods for sentiment analysis

Aspect	Matched Sentiment Term	Polarity
Ambience	Loving	Positive
	Noisy	Negative
Staff	Helpful	Positive
	Great	Positive
	Rude	Negative
Food	Good	Positive
	Delicious	Positive
	Disappointed	Negative
	Tasty	Positive
Service	Poor	Negative
	Slow	Negative
	Awesome	Positive

Figure 6 Extracted Aspects and sentiment term with polarity for object "Restaurant"

language.The aspect-based sentiment analysis provides clear details of sentiments of the customer in regards to features of products and services. This leads to a better understanding of good and bad features of product or service.The suggested conception of the framework for aspect-based sentiment analysis allows extracting good and bad features along with the customer satisfaction level. This automation can effectively alter managerial decisions and efficiently reduce the labor cost of manual work. Future research can be done in the direction of feedback given in speech and audio formats. This could be done by implementing methods of machine translation with sentiment analysis.

References

[1] FJ Domínguez-Mayo, M José Escalona, Manuel Mejías, Margaret Ross, and Geoff Staples. A quality management based on the quality model life cycle. *Computer Standards & Interfaces*, 34(4):396–412, 2012.

[2] Bin Lu, Myle Ott, Claire Cardie, and Benjamin K Tsou. Multi-aspect sentiment analysis with topic models. In *2011 11th IEEE International Conference on Data Mining Workshops*, pages 81–88. IEEE, 2011.

[3] Yohan Jo and Alice H Oh. Aspect and sentiment unification model for online review analysis. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 815–824. ACM, 2011.

[4] Bing Liu, Minqing Hu, and Junsheng Cheng. Opinion observer: analyzing and comparing opinions on the web. In *Proceedings of the 14th international conference on World Wide Web*, pages 342–351. ACM, 2005.

[5] Gayatree Ganu, Amlie Marian, and Nomie Elhadad. Ursa-user review structure analysis: understanding online reviewing trends. *DCS technical report2010*, 2010.

[6] Eivind Bjørkelund, Thomas H Burnett, and Kjetil Nørvåg. A study of opinion mining and visualization of hotel reviews. In *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services*, pages 229–238. ACM, 2012.

[7] Hanhoon Kang, Seong Joon Yoo, and Dongil Han. Senti-lexicon and improved naïve bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications*, 39(5):6000–6010, 2012.

[8] Fazel Keshtkar and Diana Inkpen. A bootstrapping method for extracting paraphrases of emotion expressions from texts. *Computational Intelligence*, 29(3):417–435, 2013.

[9] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pages 79–86. Association for Computational Linguistics, 2002.

[10] Christopher Scaffidi, Kevin Bierhoff, Eric Chang, Mikhael Felker, Herman Ng, and Chun Jin. Red opal: product-feature scoring from reviews. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 182–191. ACM, 2007.

[11] Saumya Chaturvedi, Vimal Mishra, and Nitin Mishra. Sentiment analysis using machine learning for business intelligence. pages 2162–2166, 09 2017.

[12] Katrin Marquardt, Marieta Olaru, and Ioana Ceausu. Study on the development of quality measurements models for steering business services in relation to customer satisfaction. *Amfiteatru Economic*, 19(44):95, 2017.

[13] Dietmar Gräbner, Markus Zanker, Gunther Fliedl, Matthias Fuchs, et al. *Classification of customer reviews based on sentiment analysis*. Citeseer, 2012.

[14] Walter Kasper and Mihaela Vela. Sentiment analysis for hotel reviews. In *Computational linguistics-applications conference*, volume 231527, pages 45–52, 2011.

[15] Sasha Blair-Goldensohn, Kerry Hannan, Ryan McDonald, Tyler Neylon, George A Reis, and Jeff Reynar. Building a sentiment summarizer for local service reviews. In *WWW workshop on NLP in the information explosion era*, volume 14, pages 339–348, 2008.

[16] Edward C Malthouse, Michael Haenlein, Bernd Skiera, Egbert Wege, and Michael Zhang. Managing customer relationships in the social media era: Introducing the social crm house. *Journal of Interactive Marketing*, 27(4):270–280, 2013.

[17] Mathias Bank. Aim-a social media monitoring system for quality engineering. 2013.

[18] Hongning Wang, Yue Lu, and Chengxiang Zhai. Latent aspect rating analysis on review text data: a rating regression approach. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 783–792. ACM, 2010.

[19] Adnan Duric and Fei Song. Feature selection for sentiment analysis based on content and syntax models. *Decision support systems*, 53(4):704–711, 2012.

[20] Jimmy Lin and Alek Kolcz. Large-scale machine learning at twitter. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, pages 793–804. ACM, 2012.

[21] Marco Pennacchiotti and Ana-Maria Popescu. A machine learning approach to twitter user classification. In *Fifth International AAAI Conference on Weblogs and Social Media*, 2011.

[22] Duyu Tang, Furu Wei, Bing Qin, Ming Zhou, and Ting Liu. Building large-scale twitter- specific sentiment lexicon: A representation learning approach. In *Proceedings of coling 2014, the 25th international conference on computational linguistics: Technical papers*, pages 172–182, 2014.

[23] Chid Apte. The role of machine learning in business optimization. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 1–2, 2010.

[24] Seth Earley. Analytics, machine learning, and the internet of things. *IT Professional*, 17(1):10–13, 2015.

[25] James A Rodger, Ganesh Bhatt, Pankaj Chaudhary, Germaine Kline, and William McCloy. The impact of business expertise on information system data and analytics resilience (isdar) for disaster recovery and business continuity: An exploratory study. *Intelligent Information Management*, 7(04):223, 2015.

[26] Ee-Peng Lim, Hsinchun Chen, and Guoqing Chen. Business intelligence and analytics: Research directions. *ACM Transactions on Management Information Systems (TMIS)*, 3(4):17, 2013.

[27] Andreas Metzger, Philipp Leitner, Dragan Ivanović, Eric Schmieders, Rod Franklin, Manuel Carro, Schahram Dustdar, and Klaus Pohl. Comparing and combining predictive business process monitoring techniques. *IEEE transactions on systems, man, and cybernetics: systems*, 45(2):276–290, 2015.

[28] Jeonghee Yi, Tetsuya Nasukawa, Razvan Bunescu, and Wayne Niblack. Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In *Third IEEE international conference on data mining*, pages 427–434. IEEE, 2003.

[29] Freddy Chong Tat Chua, Hady W Lauw, and Ee-Peng Lim. Predicting item adoption using social correlation. In *Proceedings of the 2011 SIAM International Conference on Data Mining*, pages 367–378. SIAM, 2011.

[30] Hsinchun Chen, Roger HL Chiang, and Veda C Storey. Business intelligence and analytics: From big data to big impact. *MIS quarterly*, 36(4), 2012.

[31] Rajiv Kohli and Varun Grover. Business value of it: An essay on expanding research directions to keep up with the times. *Journal of the association for information systems*, 9(1):1, 2008.

[32] Ryan KL Ko, Stephen SG Lee, and Eng Wah Lee. Business process management (bpm) standards: a survey. *Business Process Management Journal*, 15(5):744–791, 2009.

[33] Haluk Demirkan and Dursun Delen. Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems*, 55(1):412–421, 2013.

[34] Amy Shuen. *Web 2.0: A Strategy Guide: Business thinking and strategies behind successful Web 2.0 implementations*. O'Reilly Media, 2018.

[35] Peter SH Leeflang, Peter C Verhoef, Peter Dahlström, and Tjark Freundt. Challenges and solutions for marketing in a digital era. *European management journal*, 32(1):1–12, 2014.

[36] Benjamin T Hazen, Christopher A Boone, Jeremy D Ezell, and L Allison Jones- Farmer. Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154:72–80, 2014.

[37] Tammo HA Bijmolt, Peter SH Leeflang, Frank Block, Maik Eisenbeiss, Bruce GS Hardie, Aurélie Lemmens, and Peter Saffert. Analytics for customer engagement. *Journal of Service Research*, 13(3):341–356, 2010.

[38] Barbara Wixom and Hugh Watson. The bi-based organization. *International Journal of Business Intelligence Research (IJBIR)*, 1(1):13–28, 2010.

[39] Rajeev Sharma, Sunil Mithas, and Atreyi Kankanhalli. Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations, 2014.

[40] James R Evans and Carl H Lindner. Business analytics: the next frontier for decision sciences. *Decision Line*, 43(2):4–6, 2012.

[41] Oyku Isik, Mary C Jones, and Anna Sidorova. Business intelligence (bi) success and the role of bi capabilities. *Intelligent systems in accounting, finance and management*, 18(4):161–176, 2011.

[42] Sentiment labelled sentences data set. http://archive.ics.uci.edu/ml/datasets/Sentiment+Labelled+Sentences. [Online; accessed 24- August-2018].