Study on Product Opinion Analysis for Customer Satisfaction on E-Commerce Websites

Ms. Vijaya Sagvekar [a, 1], Dr. Prashant Sharma [b]

[a] *Research Scholar, Department of Computer Engineering, PAHER, Udaipur, India*

[b] *Associate Professor & HOD Computer Science & Engineering Department PAHER, Udaipur (Rajasthan),* India

Abstract---- The E-commerce websites have been emerged in a high range of marketing benefits for the users to publish or share the experience of the received product by posting review that contain useful comments, opinions and feedback on the product. These days, a large number of clients acquire freedoms to look at comparative items in online sites and pick their top choices in computerized retailers, like Amazon.com and Taobao.com. Client audits in online media and electronic trade Websites contain important electronic word data of items. Sentiment Analysis is broadly applied as voice of clients for applications that target showcasing and client care. Sentiment extractors in their most essential structure classify messages as either having a good or negative or once in a while neutral supposition. A typical application of sentiment investigation is the programmed assurance of whether an online review contains a positive or negative review. Subsequently, in this paper, with the use of the strategies on sentiment analysis, obstinate sentences alluding to a particular element are first recognized from item online audits. We have proposed deep learning strategy as a classification model for discovering the condition of review. The outcomes showed suggested site for the client dependent on the early audits, past reviews and answer given to inquiry audit for the client. Additionally, it is seen that the proposed strategy can ready to answer every one of the reviews with a superior closeness like a human reaction to the client.

Keywords--- E-commerce, opinion analysis, sentiment analysis, deep learning

1. Introduction

The marketers and manufacturers have been focused on the market performance from long time. The development of product marketing strategy for managing the product quality is helpful in making a better decision shows improvement in the performance [1]. The manufacturers and advertisers continuously gather the product information that useful in analyzing the performance in the market. In the traditional methods, the main data sources for analysis of main data sources were collected from the manufacturers" internal data, offline customer reviews by the surveys, review forms should be in handwritten format etc [2].

---

[1] Ms. Vijaya Sagvekar, Research Scholar, Department of Computer Engineering, PAHER, Udaipur
Email: vijayasagvekar@gmail.com.

As the technology became advanced, the e-commerce websites enabled to publish the opinion of users on product and the users got a public platform to share their useful comments and opinions towards the product they purchased. According to the survey, 97 % of consumers will be influenced for purchasing the product after going through the reviews posted by other customers. So, a most of clients will see online reviews prior to settling purchase decision [3,4].

91 % of people regularly or occasionally read online reviews before purchasing a product and also the early reviews on product before purchasing has high impact on succeeding the product sales [5]. Even though, the early reviewers contribute small proportion of reviews, it is easy to determine the failure or success of new services and new products. Based on the early reviews from the early reviewers helps to adjust marketing strategies in product improvement on designs and helps in succeeding the new product [6]. Based on the early reviews, the companies will recognize the early reviews thereby improve the product designs, marketing strategies that lead to a success on newly launched products. Thus, early analysts become the accentuation to screen and pull in at the early advancement phase of an organization in the world [7-9]. The early reviews have pulled in showcasing experts broadly investigate the customer buy goals. For example one of the largest e-commerce companies is Amazon in the world, where it provides early reviewer program opportunities for making early reviews that helps the company to acquire those early reviews on products that have few reviews or no reviews. With the help of Amazon shoppers program, will provide information about buying a product by making smarter decision. Based on the above discussions, we can see that for product marketing the most important is early reviewers of product [10-14].The present research will take initiative for studying the behavior characteristics for the early reviewers and posting it in e-commerce platform such as Amazon, Yelp. The main aim of the research work is to analyze the consumer satisfaction on products bought from E-commerce sites and improved the performance. The overall characteristics of early reviews have to be analyzed from the early reviewers needed to be compared to majority and laggard reviewers. An early reviewer tends to post more helpful reviews and helps to gives a higher average rating score to products. The rating behaviours are characterized that helps to find the scores received from others that helps to determine the correlation of the reviews based on the product popularity [15]. The discoveries with the character factors are connected with the hypothesis as follows: higher normal rating scores can be considered as the ideal attitude towards the items, and higher support votes of early surveys given by others can be seen as an intermediary proportion of the opinion leadership.

2. Literature Review

Many researches have been developed for predicting the ranking based on the online reviews. Some of the studies are as follows. In the literature survey, a survey of recent techniques is highlighted with its advantage and limitations.

Table 1. Literature Review

Authors	Method	Advantage	Disadvantage
Ahani, A., et al [16]	Self-Organizing Map and Higher Order Singular Value decomposition clustering algorithm	The hybrid algorithm was used for assisting that overcame the data related complications for online reviews and presented spa hotel market segmentation for predicting the travel choice using machine learning algorithms.	The available customer data from TripAdvisor included only the general preferences of spa hotel customers degraded the performance rate

Rita, P et al [17]	Four-dimensions of e-service quality model	The four dimensions of e-service quality were considered such as the impact on customer trust, satisfaction, customer behaviour and building existed literature based on e-service quality during online shopping.	The quality of online stores, in general, was not based on the product segments sold in online stores and the measurements were not applicable for assessing product segments.
Lucini, F.R., et al [18]	Texting Mining approach Latent Dirichlet Allocation (LDA)	The developed model presented a novel framework for customer satisfaction measuring in the airline industry using Latent Dirichlet Allocation (LDA) that detected the popular topic using the natural language processing and machine learning process	The data restricted the diversity of opinions as the proficient in English were in more and likely provided airline experience information.
Bai, T. et al [19]	Margin based Embedding Ranking Model (MERM)	The characterized and predicted the early reviewers for E-commerce sites to present effective product marketing. The developed model used Margin based Embedding Ranking Model (MERM) that predicted the early reviewers in a cold-start setting	The developed model used Margin based Embedding Ranking Model (MERM) that predicted the early reviewers in a cold-start setting.
Zhao, Y.,et al [20]	Technical attributes and sentiment polarity	The review samples were taken from trip advisor that predicted overall customer satisfaction using technical attributes in online and textual reviews of customers.	The textual reviews are influenced by languages and are different cultures needed extension for examining different language reviews.
Liu, Y. et al [21]	Product Competitive and Quality Management and Marketing Strategy	The developed a product competitive advantage analysis for providing an essential basis for quality management and marketing strategy development on social media. The novel method provided an essential basis for managing the marketing strategy and quality using user generated content	The customer comments relied which comes after target product availability of the customers and was not applicable during the design and development of the product.
Jian Jin et al [22]	Product feature extraction and sentiment analysis	The developed model performed opinionated representative for specific product based on the features especially for the competitive products. The sentimental analysis was performed for opinionated sentences that refer specific features for online reviews.	During choosing comparative sentences in the review that were of many for different products resulted lower information comparativeness values.
Sun, Q., et al [23]	Sentiment analysis eWOM	The developed model extracted large volume of online customer reviews and performed sentiment analysis for eWOM products. The developed model used semi-supervised fuzzy product ontology mining algorithm for extraction of features with negative or positive labels.	The developed model required improvement in positive and negative opinion words extraction and also polarity computation.
Thompson, J.J., et al [24]	Lexicon-Based Sentiment Extraction Analysis	The developed model performed sentimental analysis for chat messaging by the player who was involved in video game StarCraft 2. The developed model performed sentimental analysis was applicable for toxicity detection and also identified the players and their messages that are threat for the player.	The human raters disagree not only the sentiment but also disagreed the toxicity that suggested that the task is not straight forward and thus the performance at high rate was impossible

| Kumar, S., et al [25] | Sentimental Analysis For The Product Review For EEG Response | The developed model performed multimodal framework for the estimation of product rating on customer product and their brands. The reviews obtained from the global viewers were processed using Natural Language Processing (NLP) technique that computed the score for global rating. | The emotional state needed to be considered for improving the performance. |

3. Motivation for Study

There are numerous quantities of studies that examine the monetary results or the drivers of online reviews and, simultaneously, propose suggestions for the plan of review frameworks, for example, giving reviewers a predefined review format. Albeit a significant number of review framework configuration highlights have been proposed throughout the long term, truth be told, not very many have really been dissected. The new online plans of action and conditions have arisen including two-sided stage organizations (e.g., Uber). These empower two-sided reviews and require adjusted plan highlights to, for example, alleviate correspondence in two-sided review frameworks. Finally, most plan highlights of review frameworks have been investigated for fixed gadgets like PCs. Notwithstanding, online reviews are progressively delivered and burned through cell phones, which require explicit plan highlights. These are the unsolved problems in the ranking analysis for online reviews and many research work based on this problem are still examining.

4. Motivation for Study

There are numerous quantities of studies that examine the monetary results or the drivers of online reviews and, simultaneously, propose suggestions for the plan of review frameworks, for example, giving reviewers a predefined review format. Albeit a significant number of review framework configuration highlights have been proposed throughout the long term, truth be told, not very many have really been dissected. The new online plans of action and conditions have arisen including two-sided stage organizations (e.g., Uber). These empower two-sided reviews and require adjusted plan highlights to, for example, alleviate correspondence in two-sided review frameworks. Finally, most plan highlights of review frameworks have been investigated for fixed gadgets like PCs. Notwithstanding, online reviews are progressively delivered and burned through cell phones, which require explicit plan highlights. These are the unsolved problems in the ranking analysis for online reviews and many research work based on this problem are still examining.

5. Proposed Methodology

In our paper, the proposed a application for predict online items in different E-commerce sites like Amazon, Paytm, Flipkart, ShopClues and Snapdeal in reviews of clients. Firstly, we processed the datasets of different items of particular organization from E-commerce sites. Each product has its own features for categorized depends on a particular feature. This features such as significantly positively or negatively influences customer ratings, readability, subjectivity, length–significantly, and sentiment polarity. The entire surveys of client that refreshed itemsbased are generated for collecting FCM model [27].
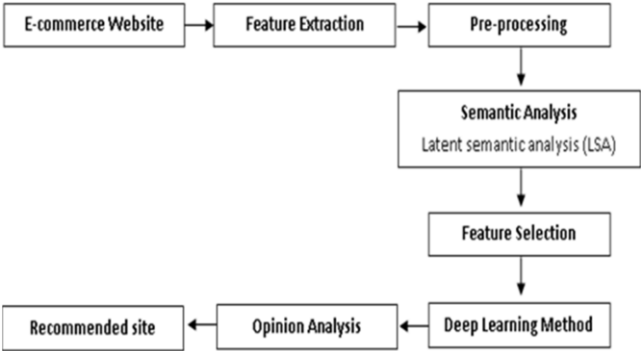
Figure 1. Flow graph of Proposed Method

Dataset

The proposed research uses a row dataset obtained from online E-shopping sites (ShopClues, Paytm, Flipkart, Amazon and Snapdeal). Each and every site is important in introducing a chance to buy the product by great population. The popularity of these sites is identified by various administrations that help to improve the distribution of various products belonging to various categories. Among the E-commerce database, Amazon dataset is used for the present research that has around142 million product reviews taken for the duration May 1996 to July 2014. Each of the review from the site consisted of textual comment that was posted by the product user that accompanied to publish time stamp accurate in the study. Usually the reviews are associated scale that is having up to five star of rating that is associated with textual description. A review an audit is related with a rating score in a five-star scale. Every item is related with a category label and a textual description.

Feature Extraction The reviews from the dataset are extracted using feature extraction technique that reviews of collected item from the E-commerce sites improved the quality of review during analysis. The feature extraction includes the items enhancement for performing audit rating of each item. The attributes here refers to the cost, quality, positive survey. The feature extraction mines the opinion of customer reviews, summarize the reviews, store and produce step preparation.

Pre-processing Transforming text into something an algorithm can process is a confounded process. In this part, the means associated with text processing are as per the following. (i) Tokenization: The superfluous tokens are not difficult to filter, where a document is changed into passages or sentences into words. In our work, the online review is tokenizing into words. (ii) Removal of Unnecessary labels and Punctuation: The following stage is to eliminate Punctuation, as the Punctuation doesn't do any additional data while treating text data. (iii) Removing stop words: Frequent words such as "the", "is", etc. that do not have explicit semantic. (iv) Lemmatization: Another way to deal with eliminate affectation by deciding the grammatical form and using point by point data set of the language.

Semantic Analysis The semantic analysis is performed for the preprocessed data where, the natural language content reads all the words captures the content by performing real meaning of any text. The text elements are assigned and are assigned based on the logical and grammatical role. The semantic analysis analyzes the surrounding context in the text to accurately disambiguate the exact meaning of words. The relationship

between the concepts in the text is also developed to identify the most relevant elements in text and understand the topic discussed. In the semantic analysis, Latent Semantic Analysis (LSA) was used in NLP that analyses the relationship among the set of documents and the terms they contain by creating a set of concepts related to the documents and terms.

Feature Selection From the extracted concepts, feature selection process is performed for selecting the subset terms that were occurred in training and these selected subsets were treated as features that performed text classification. Firstly, the training is performed and is applied for the classifier decrease the size of the vocabulary effectively. Secondly, feature selection improves the classification accuracy thereby eliminates the noise features in figure1.

Deep Learning the deep learning methods provide an opportunity that faces challenges in NLP problems such as sequence-to-sequence prediction. The developed model performed deep learning methods for learning the features based on the NL which is required by the model specifies the required features and were extracted. In natural language processing, the performance of deep learning is depending on genuine outcomes and that the enhancements give off an impression of being proceeding and maybe speeding up.

Opinion Analysis The optimal ranking opinion assesses the significance of each element relatively with respect to the sentiment score that utilized for rank highlights measuring. Based on the opinion analysis, the commentators rated the sites that are having vital data distinguished the untruthful opinions.

Recommended Site Based on these recommendations, a perspective positioning calculation were performed for ranking the vital angles resulted a viewpoint recurrence that impacted opinions for each perspective for general sentiments.

6. Results and Discussion

For accurate recommendation of products a new opinion analysis system is developed, by analyzing the reviews that users are posted for the products. The main aim of this research is to create a accurate keyword extraction technique and clustering approach for recommending the products using amazon customer review dataset interms of positive and negative form. In this paper, with GWO algorithm, a keyword extraction method (LDA) is used for selecting the proper key words. The acquired same keywords are clustered using PFCM algorithm. The developed recommendation system has main advantage is that, system has ability to find the fake products, keep track of clients satisfaction etc.

Table 2. Comparison between proposed systems and different classifiers.

	Precision	Recall	F1-Measure	AUC	Accuracy
Random Forest	69.003	73.25	71.132	48.263	73.187
Decision Tree	75.766	75.583	74.618	54.628	75.528
Proposed	76.106	76.781	74.167	57.404	77.236

The proposed system conveyed a powerful execution through quantitative analysis and comparative analysis. From the test analysis, the proposed system accomplished around 77.236% 77.236% of classification accuracy, but the existing techniques achieved

limited accuracy in amazon customer review dataset. In future work, an accurate system is implemented for further improvement in the classification accuracy for recommendation of product.

7. Conclusion

Now a days the developer community is more and more focusing on the user experience of browsing, because number of users browsing the internet are exponential increased. In this paper, proposed the approach to predict the most effective web-based shopping sites. By client survey determined whether poor or great is the product from business websites. It is more important for any business to have the knowledge about the reviews of customer regarding its any particular items.

In this paper, to analyze the characteristics and categorize early reviewers on an e-shopping sites and their effect on product popularity will use the Amazon datasets. This proposed study review posting process and develop a deep learning model for the prediction of reviewers. We also mine the summary of surveys, customer reviews. Based on the opinion analysis, a recommended perspective positioning computation to rank the fundamental points by contemplating both the viewpoint recurrence and the impact of opinions gives the recommended site.

References

[1] Smith, A.N., Fischer, E. and Yongjian, C., 2012. How does brandrelated user-generated content differ across YouTube, Facebook, and Twitter?. Journal of interactive marketing, 26(2), pp.102-113.

[2] Dong, J., Li, H. and Zhang, X., 2014, November. Classification of customer satisfaction attributes: an application of online hotel review analysis. In Conference on e-Business, e-Services and e-Society (pp. 238-250). Springer, Berlin, Heidelberg.

[3] De Giovanni, P. and Zaccour, G., 2014. A two-period game of a closed-loop supply chain. European Journal of Operational Research, 232(1), pp.22-40.

[4] Farhadloo, M., Patterson, R.A. and Rolland, E., 2016. Modeling customer satisfaction from unstructured data using a Bayesian approach. Decision Support Systems, 90, pp.1-11.

[5] Elish, M.O., 2014. A comparative study of fault density prediction in aspect-oriented systems using MLP, RBF, KNN, RT, DENFIS and SVR models. Artificial Intelligence Review, 42(4), pp.695-703.

[6] Clavel, C. and Callejas, Z., 2015. Sentiment analysis: from opinion mining to human-agent interaction. IEEE Transactions on affective computing, 7(1), pp.74-93.

[7] Nahm, Y.E., 2013. A novel approach to prioritize customer requirements in QFD based on customer satisfaction function for customer-oriented product design. Journal of Mechanical Science and Technology, 27(12), pp.3765-3777. Yang, L., Geng, X. and Liao, H., 2016. A web sentiment analysis method on fuzzy clustering for mobile social media users. EURASIP Journal on Wireless Communications and Networking, 2016(1), p.128.

[8] Yang, J.B., Wang, Y.M., Xu, D.L., Chin, K.S. and Chatton, L., 2012. Belief rule-based methodology for mapping consumer preferences and setting product targets. Expert Systems with Applications, 39(5), pp.4749-4759.

[9] Decker, R. and Trusov, M., 2010. Estimating aggregate consumer preferences from online product reviews. International Journal of Research in Marketing, 27(4), pp.293-307. [11] Dehkharghani, R., Mercan, H., Javeed, A. and Saygin, Y., 2014. Sentimental causal rule discovery from Twitter. Expert Systems with Applications, 41(10), pp.4950-4958.

[10] Chen, A., Lu, Y. and Wang, B., 2017. Customers" purchase decisionmaking process in social commerce: a social learning perspective. International Journal of Information Management, 37(6), pp.627-638.

[11] Uysal, A.K. and Gunal, S., 2014. The impact of preprocessing on text classification. Information Processing & Management, 50(1), pp.104- 112.

[12] Zhang, W., Yoshida, T. and Tang, X., 2011. A comparative study of TF* IDF, LSI and multi-words for text classification. Expert Systems with Applications, 38(3), pp.2758-2765.

[13] Zhao, T. and Eskenazi, M., 2016. Towards end-to-end learning for dialog state tracking and management using deep reinforcement learning. arXiv preprint arXiv:1606.02560.

[14] Ahani, A., Nilashi, M., Ibrahim, O., Sanzogni, L. and Weaven, S., 2019. Market segmentation and travel choice prediction in Spa hotels through TripAdvisor"s online reviews. International Journal of Hospitality Management, 80, pp.52-77.

[15] Rita, P., Oliveira, T. and Farisa, A., 2019. The impact of e-service quality and customer satisfaction on customer behavior in online shopping. Heliyon, 5(10), p.e02690.

[16] Lucini, F.R., Tonetto, L.M., Fogliatto, F.S. and Anzanello, M.J., 2020. Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. Journal of Air Transport Management, 83, p.101760.

[17] Bai, T., Zhao, W.X., He, Y., Nie, J.Y. and Wen, J.R., 2018. Characterizing and predicting early reviewers for effective product marketing on e-commerce websites. IEEE Transactions on Knowledge and Data Engineering, 30(12), pp.2271-2284.

[18] Zhao, Y., Xu, X. and Wang, M., 2019. Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. International Journal of Hospitality Management, 76, pp.111-121. [

[19] Liu, Y., Jiang, C. and Zhao, H., 2019. Assessing product competitive advantages from the perspective of customers by mining usergenerated content on social media. Decision Support Systems, 123, p.113079.

[20] Jin, J., Ji, P. and Gu, R., 2016. Identifying comparative customer requirements from product online reviews for competitor analysis. Engineering Applications of Artificial Intelligence, 49, pp.61-73.

[21] Sun, Q., Niu, J., Yao, Z. and Yan, H., 2019. Exploring eWOM in online customer reviews: Sentiment analysis at a fine-grained level. Engineering Applications of Artificial Intelligence, 81, pp.68-78.

[22] Thompson, J.J., Leung, B.H., Blair, M.R. and Taboada, M., 2017. Sentiment analysis of player chat messaging in the video game StarCraft 2: Extending a lexicon-based model. Knowledge-Based Systems, 137, pp.149-162. [25] Kumar, S., Yadava, M. and Roy, P.P., 2019. Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. Information Fusion, 52, pp.41-52.

[23] Pradip Suresh Mane, Ashok Kumar Jetawat, Pravin Jagannath Nikumbh", 2020, "Web Page Recommendation using Random Forest with Fire Fly Algorithm in Web Mining", International Journal of Engineering and Advanced Technology (IJEAT), Vol. 9, Issue-3, February 2020, pp. 499-505.

[24] Kolhe L.N., Khairnar V., Jetawat A.K. :Prediction-Based Parallel Clustering Algorithm for M-Commerce. In: Fong S., Akashe S., Mahalle P. (eds) Information and Communication Technology for Competitive Strategies. Lecture Notes in Networks and Systems, vol 40. Springer, Singapore, pp. 31-39