THE QUEST
ISSN 1908-3211 (Online) | ISSN 1908-322X (Print)
Journal of Multidisciplinary Research and Development
P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest

Utilizing Google Map Reviews and Sentiment Analysis: Knowing Customer Experience in Coffee Shops

Sahagun, Mary Anne M.[1]*, Flores, Jun P. [2] , Jocson, Joefil C.[3]

[1,2] *Don Honorio Ventura State University, Bacolor, Pamanga, 2001, Philippines*
[3] *Nueva Ecija University of Science and Technology, Nueva Ecija, Philippines. 3104*
*Corresponding author's email: masahagun@dhvsu.edu.ph

Abstract. Electronic word of mouth (eWOM) is a good source of information, and this includes customer reviews. Through this review, consumers make informed decisions. In this study, the researchers utilized Google Maps Reviews of customers of three known coffee shops. A google map review scraper was used to extract all customer's reviews and star ratings. In order to extract important information from reviews, opinion mining was done. MATLAB R2022a was used for sentiment analysis and opinion pre-processing. Each coffee shop's most popular words are represented using the Bigram model and the bag-of-words technique. This allows for the visual identification of the unique characteristics of these coffee businesses. According to the study's findings, coffee shop B had the most positive average percentage sentiment score (73%), while coffee shop C had the least negative average sentiment score. The Bigram model shows that customers enjoy the coffee these three coffee shops serve. However, when it comes to taste, location, bread, and pastries, coffee shop C has the most words. Lastly, the correlation values for star ratings vs sentiment scores for coffee shops A and B are r=0.4726 and r=0.4812. There is absolutely no association between sentiment score and star ratings for coffee shop C.

Keywords: Bag—of—word; E—worm; Customer review; Opinion Mining; Sentiment analysis; Vader algorithm

1. Introduction

EWOM, also known as electronic word-of-mouth, is the peer-to-peer sharing of thoughts and recommendations on goods and services over the internet. A reliable information source is an eWOM (Dixit et al., n.d.). Because they reflect people expressing their evaluations of a good or service, customer reviews are regarded as eWOM. Customers are satisfied when they

P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest

share their positive experiences. (Pereira et al., 2017). Several authors have demonstrated the impact of eWOM on purchase intention. (Ohk & M. Kim, 2018), (Yusuf et al., 2018). Ninety-one percent of customers said they checked online reviews before making purchases. (Cheung et al., 2009). Consumers who are preparing to buy a specific commodity would compare similar products that have the same features, such as price, function, and quality. Web customer evaluations are a great resource for competition mining because they provide a bunch of comparative data. In this paper, such tasks involving the process of competitive mining were through analyzing the sentiment reviews from customers' experiences. In this manner, this helps business owners and competitors to see what customers like and dislikes in a certain business.

Reviews serve as a platform for consumers to establish trust so that they may examine the experiences of past buyers and make wise judgments. Customers can leave opinions and comments about their personal experience on certain services, like google map reviews, a user needs to type the name of the business and location using Google Maps, and afterward customers can make comments and ratings of their experiences on products or services. Also, customers may leave ratings that provide strong convictions for prospective customers. Ratings alone, however, do not accurately reflect customer sentiment. To effectively assess and categorize sentiments in text, approaches for classifying opinions must be used. When a company monitors the online discussion, this kind of extraction aids in comprehending the social sentiment or products or services associated with their brand.

In this paper, the researchers utilized google Maps reviews of customers of three known coffee shops. A google map review scraper was used to extract all customer's reviews. Opinion mining was conducted to extract significant information from reviews. Pre—processing of opinions and sentiment analysis were conducted using MATLAB R2022a. The most frequent word review of each coffee shop is illustrated using bag-of-words. This study will help future buyers to make better decisions based on the analysis of feedback received. It will also allow business owners to meet consumer expectations better on the basis and see competitors' strong points.

THE QUEST
ISSN 1908-3211 (Online) | ISSN 1908-322X (Print)
Journal of Multidisciplinary Research and Development
P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest

Related Works

Identifying user sentiments based on positive, negative, and neutral connotations is the goal of sentiment analysis, also known as opinion mining. Three categories of opinion mining exist: document level, sentence level, and phrase level. (B. Liu, 2012). Numerous real-world issues can be solved with sentiment analysis. Many businesses are adopting sentiment analysis as an example. The study that determined the polarity of smartphone product reviews only based on the positive and negative orientation of the review (Wahyudi & Kristiyanti, 2016). Sentiment Analysis was used to predict election results for three major political parties (Boutet et al., 2012). tracking the popularity and desirability of a brand (Greco & & Polli, 2019). Studying product launches (Rathore & & Ilavarasan, 2020). A decision aid for purchasing products and services (Feldman, 2013).

Some earlier relevant studies served as inspiration for this study. The frequent usage of the multiword phrase was not taken into consideration, notably in coffee shops, even though several works on sentiment analysis have been done using different methodologies. In this paper, customer experiences in three known Coffee Shops inside the Clark Economic Zone, Philippines were considered, and opinion mining was performed. Gained text opinion was analyzed using sentiment analysis. Furthermore, this study includes a bag of n-gram models and analyses text data using multiword phrases.

2. Methodology

The approach for sentiment analysis can be machine learning-based and lexicon-based. For this study, the lexicon-based method was used. It uses a sentiment dictionary with opinion words and matches them with data to determine polarity. Sentiment scores are assigned to the opinion words describing how positive, negative, and objective the words contained in the dictionary are. Data was collected from Google Map Reviews and were scraped using Bolster application. Data pre-processing and data cleaning were performed before analyzing. Sentiment Analysis was performed using command function in text analytics in MATLAB R2022a. A customer review with sentiment score of "0" is a neutral polarity, a

THE QUEST ISSN 1908-3211 (Online) | ISSN 1908-322X (Print)
Journal of Multidisciplinary Research and Development
P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest

sentiment score of greater than "0" is a positive polarity, and a value less than "0" is a negative polarity. Bag—of—words and bigram model were used to further visualize the sentiments of the customers. Correlation Analysis was also performed to see if there is a correlation with Sentiment scores of customer's reviews and their star rating.

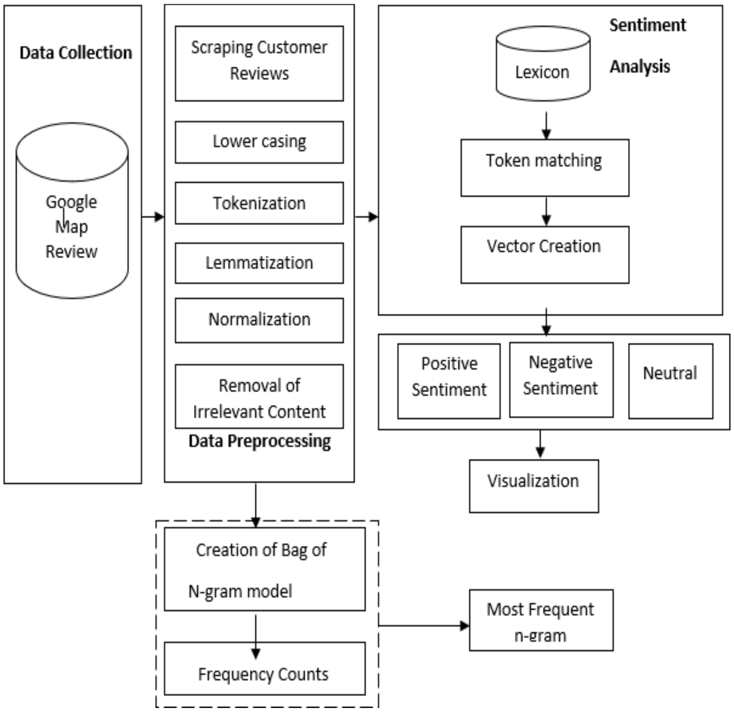


Figure 1. System Architecture

2.1. *Data Collection and Pre—processing*

The researchers considered three coffee shops inside the Clark Economic Zone, Philippines. Coffee Shop "A" is an internationally known coffee shop that started in 2009, coffee Shop B just opened last 2021 during the pandemic, and coffee shop C opened just this 2022. Customer Reviews in google map reviews of each coffee shop were scrapped using a bolster application. The following number of customer reviews were taken: 100 reviews for coffee shop A, 96 reviews for coffee shop B, and only 51 reviews available in google map review for coffee shop C.

Pre-processing of data includes lowering all cases of reviews using Microsoft

THE QUEST
ISSN 1908-3211 (Online) | ISSN 1908-322X (Print)
Journal of Multidisciplinary Research and Development
P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest

excel, followed by tokenizing the “text data”, using the command function "tokenized document" in MATLAB R2022a. Lemmatization was improved by removing stop words using the command function as seen in figure 2. Upon normalization of cleaned documents, short words and longwords were removed using the syntax:

newDocuments= removeShortWords(documents,len)

where: len = Maximum length of words to remove, specified as a positive integer. The function removes words with len or fewer characters. Data cleaner application using MATLAB R2022a was used to process the modification of data remove or correct information in preparation for analysis.

Figure 2. Pre-processed data of Coffee Shop A

2.2 *Sentiment Analysis*

The second phase is sentiment analysis. The system analyses the pre-processed data to identify instances of sentiment. It uses NRC Emotion lexicon which includes annotations for 14,182 unigram words for English.

Cleaned and tokenized documents are evaluated with a sentiment lexicon with words annotated with a sentiment score or intensity of emotion ranging from -1 to 1 where scores close to 1 indicate strong positive sentiment, scores close to -1 indicate strong negative sentiment and scores close to zero indicate neutral sentiment. Sentiment Scores are computed using the command function ‘vaderSentimentScores’ or by using the Valence Aware Dictionary and Sentiment Reasoner (VADER) algorithm.

Each word in the lexicon has an emotion vector ($\vec{E}$)containing a Boolean value (b) for each sentiment (s) and emotion (e):

$$\vec{E} = \vec{E}_e + \vec{E}_S \qquad \text{eq. (1)}$$

where: $\vec{E}_e \in \{b_0, \ldots, b_7\}$ and $\vec{E}_S \in \{b_8, b_9\}, \forall\, b_i \in \{0,1\}$

P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest

Lexicon-based approach; opinion words are divided into two categories. Positive opinion words are used to express some necessary things, and negative opinion words are used to describe unnecessary things.

2.2 Visualization

To facilitate analysis of customer's review or feedback about their experiences, the sentiment analysis system has a data-visualization component that creates sentiment word clouds. The system groups together positive and negative comments, in this manner, it is possible to track the customers' likes and dislikes. Figure 3-5 shows the bag of words of a cloud of the three coffee shops.

2.3 Bag of N-gram Model

A bag-of-n-grams model records the number of times that each n-gram appears in each document of a collection. An n-gram is a collection of n successive words. does not split the text into words. The performance of faculty can also be seen on the most frequent n-gram.

3. Results and Discussion

Sentiment analysis was conducted using the function Vadersentimentscore in MATLAB R2022a to determine the sentiment scores. Figure 3 shows the result of comment reviews and sentiment scores. Coffee shop B has the highest positive average percentage sentiment score of 73%. Since both coffee shops B and C are newly opened, coffee shop C shows better sentiment scores having the lowest

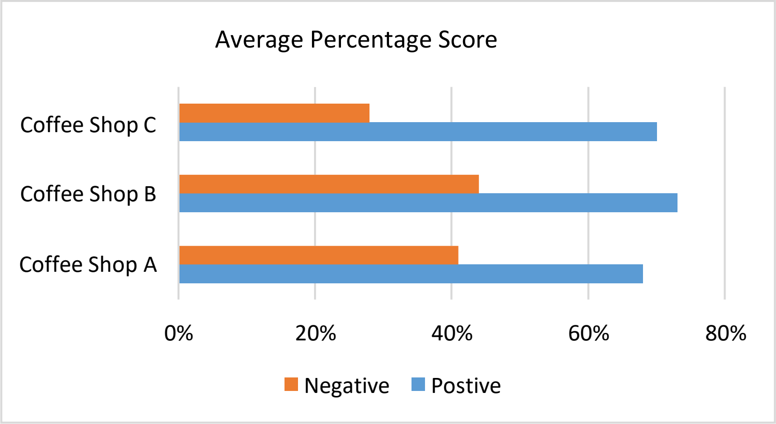


Figure 3. Summarized Average Sentiment Score

negative average percentage sentiment score of 28%. It is also a good thing to note that even if coffee shop A had been existing for almost 13 years its reviews

THE QUEST
ISSN 1908-3211 (Online) | ISSN 1908-322X (Print)
Journal of Multidisciplinary Research and Development
P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest

from customers are more likely the same as coffee shop B with a positive average percentage sentiment score of 68%. The use of the Bag of word model is also important to visualize the frequency number of times that words appear in each collected review.

This is an effective method for extracting information from large amounts of text data. The strength of coffee shop A reveals its strength seen in bag-of-word in Figure 4

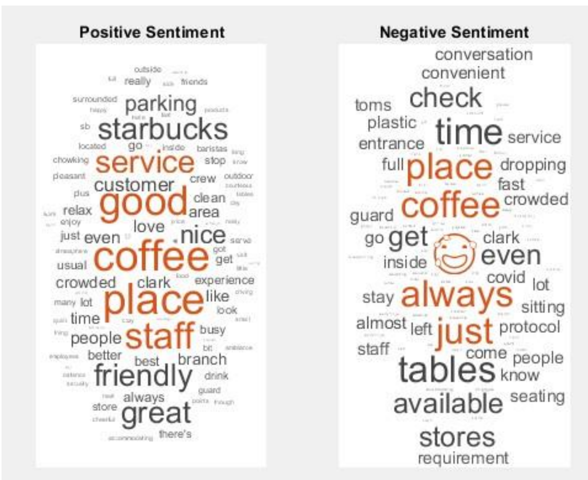


Figure 4 Coffee Shop A

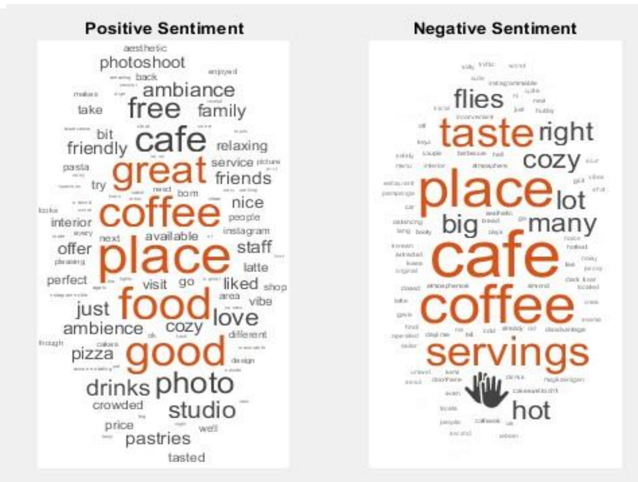


Figure 5 Coffee Shop B

The following words were the strength of the shop: “service”, “good”, “coffee”, “place”, and “staff”. From these bags of words, customers like the service of the shop, the area or place was also complemented, the coffee was mentioned it means customers are not just on areas, services but of course the product of the shop which is coffee, and lastly, customers appreciate the staff of the coffeeshop which complements also the word "service". Looking at its negative sentiment, "coffee" was considered a negative sentiment. "Tables" and "available" words are also seen as well as the word "crowded” This can be interpreted that the coffee shop is always crowded and finding a place or sit is often difficult. Coffee Shop B in figure 5 was appreciated by the word "great",and "coffee”. It can also be seen that 'food" and not just coffee were commented positively by customers. Among other words that can be seen are the words “photoshoot”, “family", "friends", and "ambiance”. This is because these coffee shops are Korean-concept café and more likely a studio café that offers free photoshoots for family and friends. One that is disturbing under the negative sentiment was the word “flies”,

In Figure 6, the word “recommended" is a good indicator of coffee shop C, aside from the “food”,” coffee”, and" place”, customers who experience the place commented

THE QUEST
ISSN 1908-3211 (Online) | ISSN 1908-322X (Print)
Journal of Multidisciplinary Research and Development
P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest

how “spacious” the place is, “pastries”, “pasta”, and "bread" were complimented.

However, in terms of "price" that is the negative sentiment of the customers of this shop.

Figure 6 Bag-of-Words of Coffee Shop C

Figure 7a Coffee Shop A Bigram Model

Bag-of-bigram Model

The bag-of-Bigram Model as seen in figure 7 was also considered to further understand the frequent words or phrases that leads to concerned topic or likes by customers. From the Bigram Model, the summary of frequent phrases per coffee shop can be observed. The strength of coffee shop A is three (3): service of the staff, the place, and the coffee. Coffee shop B is appreciated for two things: the free offer of photoshoots and they served coffee.

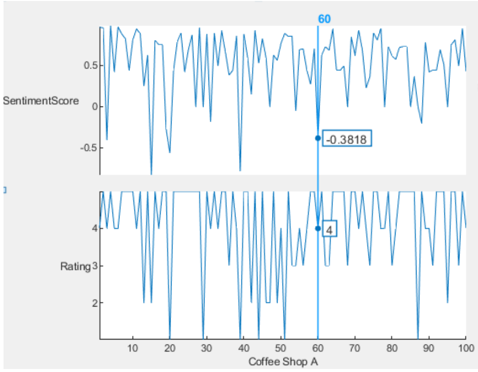
Sentiment Score versus Star Ratings

Figure 7b Coffee Shop B Bigram **Model**

Figure 7c Coffee Shop C bigram Model

A correlation analysis was performed for each coffee shop to determine whether the ratings obtained from the customer have a relationship with their sentiments. For coffee shop A in figure 8, for instance, customer number 60 rated the shop with 4 ratings.

THE QUEST
ISSN 1908-3211 (Online) | ISSN 1908-322X (Print)
Journal of Multidisciplinary Research and Development
P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest



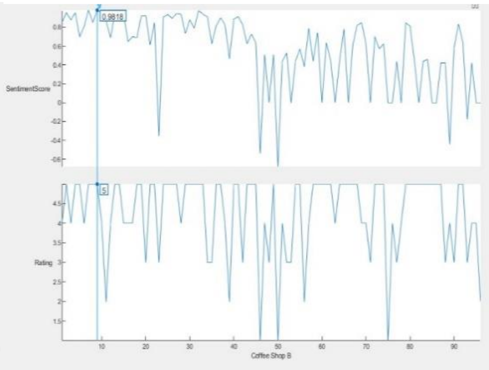


Figure 8 Sentiscore vs rating of Coffee Shop A **Figure** 9. Sentiscore vs rating of Coffee Shop B

The words “noise” and “suffer” overpower the positive word “heartwarming”. While customer number 5 rated the shop as "5" and the review made has a sentiment score of 0.9803. The Pearson correlation of sentiment score and rating of customer’s review has a value of r=0.4726 which has a positive moderate correlation between sentiment score and star rating. Figure 9 shows that review and rating of customer number 8. The star rating was "5" , and the sentiment score of the review was 0.9818. For this review, it seems that the star ratings and comments of the customer are correlated, which is not all have the same case. The Pearson correlation r of coffee shop B is r=0.4812. There is a positive moderate correlation between the sentiment score and star ratings of the customers. Coffeeshop C has the least review data as compared with the first two coffee shops since it just opened this year. It can be observed from figure 10, that there is a correlation between the star ratings of “5” was given by customer number 5 with the sentiment score of review of 0.9545. Looking into customers nos. 30-35, the graph of star ratings and sentiment score of the points are not the same. The Pearson correlation between the sentiment score and star ratings of customer reviews for this shop is -0.00865. Meaning, that there is no correlation at all.

4. Conclusion

Electronic word of the mouth contains important data that when data mining is performed, information is extracted. The customer's experiences in the three coffee shops in this study were scraped using bolster application through google Maps reviews. The sentiment analysis using the Vader algorithm revealed the true review of the customers based on their personal experiences as compared with the star rating. The bigram model helped to visually see the things customers like and the discomfort the customers

THE QUEST ISSN 1908-3211 (Online) | ISSN 1908-322X (Print)
Journal of Multidisciplinary Research and Development
P-ISSN: 1908-322X and E-ISSN: 1908-3211
Vol. 1 No. 2 (2022)
Received: 11-16 / Revised: 12-15/ Accepted:12-30-22
https://neust.journalintellect.com/quest

experienced. Each coffee shop has its peculiarity. All customer reviews appreciate the coffee served. However, for coffee shop A the place and the friendly service are most like, for coffee shop B the place and the free photoshoot, and for coffee shop C the taste, pasta, pastries, and place. Lastly, sentiment scores of comment reviews versus the star rating of the customers had a moderate correlation for coffee shops A and B, while for coffee shops C there is no correlation at all.

References

B. Liu. (2012). Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5, no, 1-167.

Boutet, A., Kim, H. H., & Yoneki, and E. (2012). What's in twitter: i know what parties are popular and who you are supporting now! Proceedings of the 2012 IEEE/ACM In- Ternational Conference on Advances in Social Networks Analysis and Mining, 132-139.

Cheung, C. M. K., Lee, M. K. O., & D.R. Thadani. (2009). The impact of positive electronic word- of-mouth on consumer online purchasing decision. World Summit on Knowledge Society, 501-510.

Dixit, S., Badgaiyan, A. J., & Khare, A. (n.d.). An integrated model for predicting consumer's intention to write online reviews. J. Retailing Consum.Serv46, 112-120.

Feldman, R. (2013). Techniques and applications for sentiment analysis. Communications of the ACM, 56(4), 82-89.

Greco, F., & & Polli, A. (2019). Emotional text mining: Customer profiling in brand man- agement. International Journal of Information Management.

Ohk, K., & M. Kim. (2018). Who's leading China's E—commerce industry? The antecedents and consequences of E—WOM focusing on one person media. J. Theor. Appl. Inf. Technol., 96 (5), 1323-1333.

Pereira, H. G., Cardoso, M., & P. Dionísio. (2017). The determinants of website purchases: the role of e-customer loyalty and word-of-mouth. Int. J. Electron. Market. Retailing, 8 (2) (201.

Rathore, A. K., & & Ilavarasan, P. V. (2020). re-and post-launch emotions in new product development: Insights from twitter analytics of three products. International Journal of Information Management, 50, 111-127.

WAHYUDI, M., & KRISTIYANTI, D. A. (2016). Sentiment analysis of smartphone product review using support vector machine algorithm- based particle swarm optimization. Journal of Theoretical & Ap— Plied Information Technology, 91, n.

Yusuf, A. S., Razak, A., Hussin, C., & Busalim, A. H. (2018). Influence of e-WOM engagement on consumer purchase intention in social commerce. J. Serv. Market, 32 (4), 493-504