Exploring consumer opinions on vegetarian food by sentiment analysis method

Tran Thi Huong Thao[1,2], Ho Trung Thanh[1,2*]

[1]University of Economics and Law, Ho Chi Minh City, Vietnam
[2]Vietnam National University, Ho Chi Minh City, Vietnam
[*]Corresponding author: thanhht@uel.edu.vn

ARTICLE INFO

DOI:10.46223/HCMCOUJS.econ.en.13.2.2256.2023

Received: April 25th, 2022

Revised: June 29th, 2022

Accepted: June 30th, 2022

JEL classification code:

C61; C63; C67

Keywords:

lexicon-based approach; product review; sentiment analysis; vegetarian food; Vietnamese market

ABSTRACT

The study's aim is to explore consumer opinions on vegetarian food in the electronic commerce area by the qualitative research method, in more detail, based on utilizing the lexicon-based approach to analyze sentiment. The sentiment of five aspects including price, package, shipment, brand, and quality is considered detecting from customers' reviews. Besides, the rating tags are also categorized as an opinion (positive, negative, neutral) based on the star number. The dataset includes 17,892 customer reviews. The findings found that the customers' sentiments are 20.8% positive, 0.7% negative, and 78.5% neutral for an average of five aspects. Product quality is the most concerning of the five aspects of customer comments. This aspect also makes up 8,672 positive comments. Besides, quality and package aspects are more than 5,000 positively attached tags and shipment has the most negative tags taking 26 rating tags. Furthermore, the average of the stars is 4.94 stars, which is close to reaching the peak of 5.0 stars. This study demonstrates that the view of customers for aspects about vegan products according to text comments, the quality aspect has a significant number of positive comments than other aspects. Hence, the business might understand the significant factors to deal with them.

1. Introduction

In Statista's report, the global vegan food market made up a value of about USD 16 billion in 2021. According to the Global Vegan Food Market study, between 2022 and 2027, this market is further predicted to grow at a CAGR of 9%, increasing by $10.7 billion to approximately $26.1 billion in 2026 (Expertmarketresearch, n.d.). A vegetarian diet was considered an effective solution to reduce the risk of cardiovascular disease, ischemic heart disease, and numerous types of cancer (Dinu, Abbate, Gensini, Casini, & Sofi, 2017). Vegetarian diets are well-suited to preserving the environment, preventing pollution, and reducing global warming, according to research (Leitzmann, 2003). Hence, the market for vegan products is currently considered in view of the fact that a vegan diet not only enhances a healthier person but also gives support to the environment. Furthermore, vegan cuisine can be found in a variety of places, including supermarket chains, convenience stores, internet retail, and others, according to Fortune Business Insights statistics. Owing to the Covid-19 pandemic, e-commerce portals have become a popular way for clients to buy things. Moreover, the chief executive officer of Vejii said that online grocery

ordering is one of the fastest-growing e-commerce industries, as well as the largest online vegan marketplace based on Forbes.

Indeed, a growing number of Vietnamese are adopting a vegan lifestyle, with Ho Chi Minh City, the country's largest and most overcrowded city, ranking among Asia's 10 leading vegan-friendly cities (Petaasia, 2017). The increasing number of vegetarian restaurants reflects the growing popularity of a vegetarian diet in Vietnam (Ngo et al., 2021).

Consumer feedback analysis is crucial for maintaining product quality and meeting customer needs (B. Nguyen, Nguyen, & Ho, 2021; Vanaja & Belwal, 2018). Many customers are leaving reviews to assist other people to find the correct things to buy as well as provide feedback to the company. Furthermore, companies and organizations have spent a lot of time and money trying to figure out how customers feel about the products and services they offer (Farkhod, Abdusalomov, Makhmudov, & Cho, 2021). This is because the opinions of those surrounding a client have a significant effect on their decision-making procedure (Kotler, 2000). Sentiment analysis has been the subject of a great deal of research to date.

Sources of information on organic, vegetarian, small-scale, and local products are the meaning element that is prominent in the informants' choices of e-shops (Berg & Henriksson, 2020). Additionally, healthy eating is the most commonly associated with a healthy lifestyle, exercise, weight loss, and diet among users (Pilař, Stanislavská, Kvasnička, Hartman, & Tichá, 2021). Besides, the product line and strategic marketing might focus on the value proposition of the commodities about vegan, homemade as well as clean eating. Ngo et al. (2021) investigated five factors (animal welfare, health, mood, religion, and reference) orientating people toward making vegetarian decisions. Thus, the emotionally driven element is found to be more significant when making a decision to acquire vegan items. Furthermore, Cooper, Dedehayir, Riverola, Harrington, and Alpert (2022) employed Leximancer, a text analytics software, to determine the themes of vegan product discussion. They pointed out that in additional topics, the researchers should find sentimental differences between vegans (positive or negative).

Some previous research explored customer sentiment in regard to aspects of type products, such as accessories and cell phones with some aspects of battery, body, display, price, network, service, sound (Anoop & Asharaf, 2020); entertainment with two groups of display and buggy functionality & performance (Alqaryouti, Siyam, Monem, & Shaalan, 2020); drug with aspects of effectiveness and side effects (Han, Liu, & Jing, 2020). Other previous studies compiled feedback on aspects of both the product and the service, for instance, restaurant with aspects of ambiance, experience, location, person, price, restaurant, service, style options, and sustenance (de Kok et al., 2018); fashion, toy, book, laptop, voucher, and so on with aspects of quality, shipment, price, design and satisfy (Le & Huh, 2021; B. Nguyen et al., 2021). Vegan products, on the other hand, have not been investigated in terms of client sentiments. With the possibility for growth prospects as well as the transition to vegetarian products, this study aims to provide more understanding of customer experiences on vegan food that are free of animal ingredients and not tested on animals by applying the sentiment method. Customers' opinion gives towards some aspects of this merchandise. To find out customer insights about vegetarian products involves four main tasks: (1) Which are the most popular keywords and their frequency with customers to write product reviews relevant to vegan food? (2) How do the consumer's emotions change over time? (3) What is the average rating star and demonstrate which cards are used the most? (4) Which aspect is the most important to customers and what is the point of view for each aspect?

The reality and scientific significance of the work in this paper are the following: (1) This research has conducted the customer experience analysis through text comments, ranking stars, and rating tags in customer reviews based-on sentiment analysis; (2) This study contributes to the area of vegan food in the Vietnamese market; (3) The business can have a deeper understanding of the customer's experience and behavior after the customers purchase products or services on the online store. From each aspect, businesses find out the points of customer displeasure in order to determine the solution; (4) This paper provides buyers and potential customers with a sense of how previous customers' sentiments things, allowing them to make appraisals, and purchase decisions.

The remainder of this paper consists of five parts arranged as follows. Section 2 provides an overview of previous research fundamentals. The research design, and sentiment analysis process are presented in Section 3. Section 4 demonstrates the results of the experimental work. The discussion of this study is shown in Section 5. And lastly, Section 6 presents conclusions from this work, its limitations, and further research proposal.

2. Background and related work

Customer experience has gotten a lot of attention in marketing research and practice over the last decade. Companies like Uber, Starbucks, and Zappos utilized digital transformation to capture customer experience owing to the customers being at the center of their business according to Voxco, and marketing researchers refer to it as the basic premise for marketing management (Lemon & Verhoef, 2016). Customer experience, according to (Holbrook & Hirschman, 1982), is defined as the entire event that a customer encounters when dealing with a brand. The customer's moods are frequently affected by this experience.

In recent years, consumers have become more eager for veggie options (Cooper et al., 2022; Choi & Joung, 2018; Lu & Gursoy, 2017), and many people around the world have decided to live a vegetarian lifestyle (Ngo et al., 2021; Rivera & Shani, 2013). This growing interest in organic or vegetarian merchandise among people is correlated to their yearning for a cleaner and greener lifestyle (Nilashi et al., 2021; Verain, Dagevos, & Antonides, 2015). In other words, nowadays vegetarian and organic goods play an essential role in consumer choice. Obviously, there has been a recent increase in research on vegetarian food. Choi, Joung, Choi, and Kim (2021) showed some factors on customer satisfaction and behavioral intentions like price, convenience, and menu options. Customer decision making based on four factors atmosphere, service, food, and value in vegetarian restaurants through cluster analysis, text mining, and predictive machine learning approaches (Nilashi et al., 2021). Cooper et al. (2022) analyzed using the text analytics tool to look for prominent topics and related terms in conversations relating to veganism. This research proposed a method to show drivers and had a limitation needed to analyze sentiment for each factor.

Sentiment Analysis (SA) also referred to as opinion mining, started early in the studies (Das & Chen, 2007; Turney, 2002; Wiebe, 2000). SA is the research of people's attitudes, feelings, judgments, views, and emotions toward objects and their properties as expressed in the written word. The sources can come from a variety of customer material for instance blog posts, feedback, chat rooms, survey responses, social media, comments, and so on.

The main approaches to analyzing the opinion include classic machine learning, lexicon-based, and hybrid.

Initially, sentiment classification has been executed at the document-by-document level. Pang, Lee, and Vaithyanathan (2002) applied Machine Learning (ML) approaches for analyzing opinions on movie reviews with three methods: Naive Bayes (NB), Support Vector Machine (SVM), and Maximum Entropy (ME) classifiers. Recently, Clara, Adiwijaya, and Purbolaksono (2020) studied the beautiful product with three aspects including price, package, and fragrance by using RF (Random Forest), TF-TDF, and n-gram. In addition, some works used an unsupervised learning method LDA (Latent Dirichlet Allocation) to find the aspect and identify which aspect is related to which opinion expression in a review of cell phone accessories (Anoop & Asharaf, 2020) and movie (Farkhod et al., 2021).

Deep learning distributed vectors to learn latent characteristics and has recently been demonstrated to outperform several machine learning approaches on similar tasks. Wang, Pan, Dahlmeier, and Xiao (2016) established a Recursive Neural Conditional Random Fields model by combining CRF with a Recursive Neural Network. Mai Long and Le (2018) used a sequence-labeling approach in conjunction with Bidirectional Recurrent Neural Networks (BRNN) and Conditional Random Fields (CRF) to extract opinion targets and assess their sentiment simultaneously so as to analyze smartphone reviews on YouTube and e-commerce. Le and Huh (2021) utilized customer reviews on Tiki with various products and services to solve the emotional analysis issue through a deep learning method with long short-term memory neural network LSTM (belong to RNN), combined with CBOW, and TF-IDF techniques.

To identify aspects of a lexicon approach, most early ABSA research used a set of filters on high-frequency noun phrases (Hu & Liu, 2004; Yauris & Khodra, 2017). A noun, adjective, verb, or adverb can all be used to indicate an aspect (Nguyen & Duong, 2019; Taboada, Brooke, Tofiloski, Voll, & Stede, 2011). The use of adjectives as indications of text semantic orientation has been the focus of much lexicon-based studies (Hu & Liu, 2004; Wiebe, 2000). Vu, Pham, Luu, and Ha (2012) approached through lexicon with VietSentiWordNet and showed syntax rules to analyze mobile phones with functionality, performance, support, and update features. In the Amazon context, Sharma, Nigam, and Jain (2014) applied a Wordnet dictionary with seed lists to determine the semantic orientation of customer reviews for books, watches, and phones. Most recently, Hung and Cao (2019) proposed a new lexicon with collocations and concepts, and word of mouth to polarize cosmetic products.

The study of Cooper et al. (2022) only used Text Analytics and stopped analyzing the topic, while we applied a sentiment method that can analyze aspects and sentiment of products and services. Our study utilizes a lexicon-based approach that does not require any prior leveling and training, such as machine learning or deep learning systems to comprehensive vegan products on e-commerce platforms. Additionally, we see that Nilashi et al. (2021) did not judge the price aspect and Choi et al. (2021) did not consider the brand aspect in their study. But we think that they are essential factors in consumer reviews as mentioned in the 3.2 section. Thus, we analyze five aspects (quality, price, brand, shipment, package) by sentiment method.

3. Research methodology

3.1. Method

The qualitative method is performed in this work. The authors conduct searches and research to aggregate regional and abroad investigation data about vegetarian food to clarify the essential contents as well as finding in-depth insights into research gaps. This approach is also

used to empirical the proposed model from data understanding, collection of rich data, preprocessing steps, cleaning phases, analysis process, visualization of the result, and implications.

About the data analysis method, we use a lexicon-based method for the sentiment classification. The lexicon-based technique involves estimating content orientation leaned on the semantic orientation of the document's words or phrases. They are assigned to tokens based on polarity such as - 1, 0, 1 for negative, neutral, and positive or the score may be assigned based on the intensity of polarity. The polarity of sentiment based on the standard level to category positive, negative, or neutral.

Corpus-based and dictionary-based methods are primarily two techniques used in this approach. The first method which addresses the challenge of identifying opinion words, relies on co-occurrence statistics or grammatical patterns in text corpora, as well as a set of positive and negative seed words. Additionally, the polarity values of each lexicon are stored in a dictionary in a dictionary-based approach. Using a few seed texts to boost based on a dictionary's antonym and synonym structure is a basic technique in this method. Collecting a tiny list of seed words with defined positive or negative polarities is the first task manually. Later, the set is enhanced by looking up antonyms and synonyms in lexical resources to expand the dictionary. After, the overall polarity score is calculated by adding the polarity scores of each word of the text found in the dictionary.

3.2. Data collection and analysis

This research has been carried out through four phases including data preparation, data preprocessing, analysis data, and visualization. The model proposal is depicted in Figure 1.

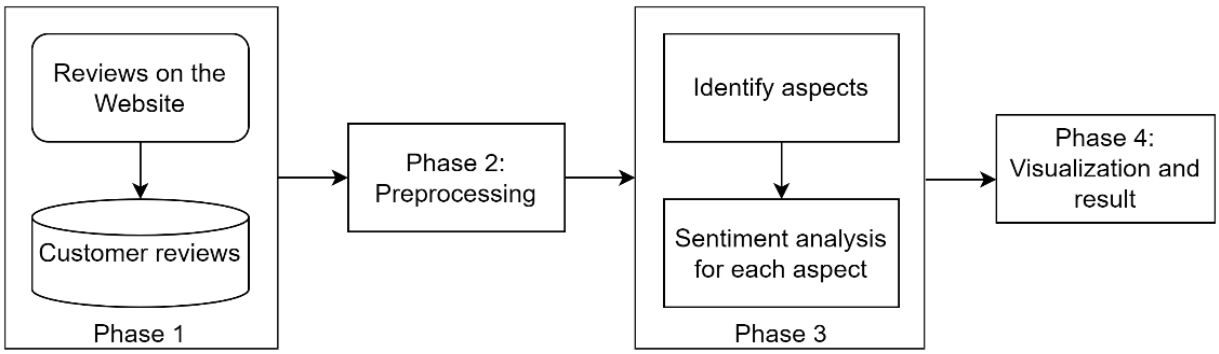

Figure 1. The proposal of the research stages

Source: Author's proposal

Phase 1: Data preparation

iPrice's research utilized data from SimilarWeb indicates that Shopee's average visit rate is remarkably high, reaching 57%, followed by Lazada 16% and Tiki 13%. Besides, the Shopee is picked to scratch the data thanks to the largest number of comments for the products.

The Cocoon Original Vietnam is a domestic vegan cosmetic brand, which was established by The Nature Store Cosmetic Company Limited - Vietnam. Cocoon is the first beauty brand in Vietnam to fulfill the cruelty-free criteria by the organization Leaping Bunny, The Vegan Society, and PETA. Thus, data samples for vegan products were chosen from the official Cocoon store. There are 28 vegan cosmetics from 06 classifications: coffee, eucalyptus & mint, grapefruit, rose, squash, and turmeric.

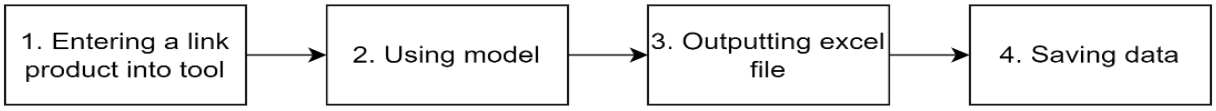

Figure 2. The steps of data preparation

Source: Author's proposal

The authors used the Octoparse tool to crawl product reviews. The data retrieval process consists of four small steps in Figure 1: Step 1: The link is picked from the Cocoon website to input the tool; Step 2: The model is created by the authors on a tool online, which is based to extract data; Step 3: Excel is the type of data that is selected for the output file, and Step 4: The data is saved in a file to the next phase. Dataset has nine columns: Material, Product, ID, Name, Datetime, Star, Price, Review, and Tags.

Phase 2: Data preprocessing

The steps of the preprocessing data

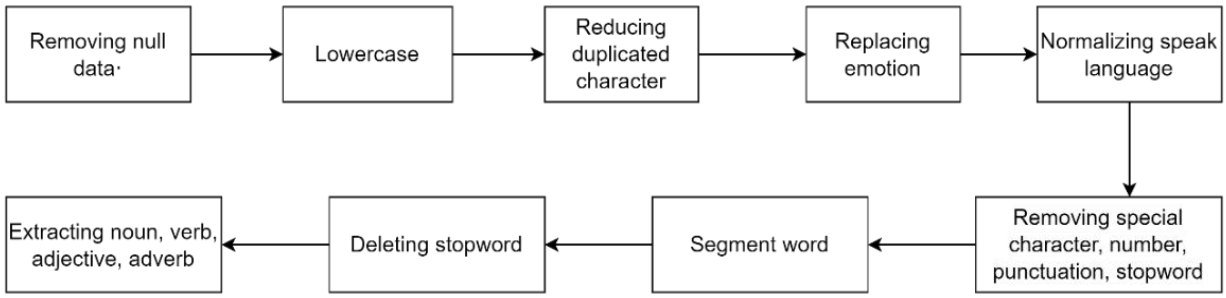

Figure 3. The steps of the preprocessing data

Source: Author's proposal

There are several steps to preprocess data which are shown in Figure 3. Firstly, the "Remove null data" step is to delete reviews with only star reviews without comments or rating tags. Following that, the Unicode Code Table distinguishes between upper and lower case. As a result, each review will be lower to deal with this problem. Then, the customers give a lot of ratings that contain 02 consecutive letters such as "Giao hàng nhanh (fast delivery)", "Sản phẩm chất lượng tốt (good quality product)". Nguyen and Nguyen (2020) thought that even though it's the same word, the recognition system is different. Hence, they are reduced the repeated characters. In addition, these days, several responses include not only "words" but also "symbols," as Bahri, Bahri, and Lal (2018) point out. The writer will convert some emojis, based on the Wikipedia "List of emoticons" and Nguyen and Nguyen (2020). The negation terms are very important to process sentimental analysis. Dang, Nguyen, and Nguyen (2018) gave that the word "không (no)" is sometimes abbreviated such as "khong", "khog","khg","kh","ko","k" etc. These are needed to replace for the next process. Besides, it is very important to omit punctuation marks such as colon, exclamation mark, dot, commas, etc. to clear text (Nguyen, Toan, & Dien, 2016). They need to be eliminated in order to tidy up the data. Stopwords are normal words, they can bring confusion for word segmentation. In Vietnamese, there are a lot of stopwords that have high occurrence frequencies but no value in meaning such as: "thì", "mà"," nên", "là", etc. (Nguyen, Ngo, Phan, Dinh, & Huynh, 2008). The authors compile a lot of stopwords from papers, documents, open-

source Github, and literature, etc. to compile the dictionary. Therefore, they are detected and removed to enhance the processing data. Lastly, the extracted words combine adjectives, adverbs, nouns, and verbs as (Nguyen & Duong, 2019; Taboada et al., 2011), which are considered sentiment words or aspect words. The pre-processing stage helps the obtained dataset to be cleaner, and processed more efficiently.

Phase 3: Analyzing data

In sentiment analysis, aspect groups are critical for the term extraction procedure. Thus, the frequency of terms in customer comments is computed to find the regular words by buyers' interest. Furthermore, Kotler's model contains five stages of the buying decision process which also showed some essential factors like appearance, price, brand, quality, and social media review (Kotler, 2000). In addition, the previous studies involving e-commerce field extracted aspects such as price (cost), quality, satisfy, shipment, design, package, fragrance, and so on (Clara, Adiwijaya, & Purbolaksono, 2020; Le & Huh, 2021; Nandal et al., 2020). Based-on the previous studies, this study continued surveying vagan food and identified food's features as well as customers' behavior relevant to the dataset to propose the five aspects for the model. Price represents the amount of money that the customers or clients must pay for the service, idea, or merchandise. Packing procedure plays a central role in protecting the product's integrity for the buyer. Shipment is defined as the act of delivering merchandise, letters, parcels to customers' houses or areas of work. Brand is considered as the assistance that an enterprise provides to its consumers both before and after they purchase products. Product quality is the characteristics of products that meet customer needs.

In customer reviews, these rating tags are optional for the customer and are used to evaluate the performance of each element. The tags are associated with 1-star and 2-star ratings which are considered negative, while positive aspects have at least 4-star ratings and 3-star tags are a neutral class (Fang & Zhan, 2015).

The authors performed the analysis with a structure-based approach to realize aspects and polarity. In order to have input data into the system, the authors need to build a list of lexicons for the dataset from sources. The vocabularies are based on some research such as (Dang et al., 2018; Le & Huh, 2021; Nguyen & Duong, 2019; Vu et al., 2012). Besides, for these descriptions of products, the authors identify the feature word and put them in the dictionary list. In addition, the authors refer to several open-source files on GitHub which are mentioned in the reference. The list of polar words is classified per aspect based on research by Yauris and Khodra (2017). Furthermore, a list of negations is saved separately. If an opinion word is accompanied by the negation, the polarity of that word is reversed. Anoop and Asharaf (2020) implemented the algorithm to find the aspect first and then its sentiment. In this research, the author also locates the aspect text. Next, the algorithm traverses right and left in seek of a term that expresses an opinion (Piryani et al., 2017). Positive words are considered to receive a positive value of 1, negative words are considered to have a negative value of -1, words with a neutral aspect receive a value of 0. If the total value of positive poles is positive, then that aspect is positive sentiment, value less than zero is negative sentiment, and the rest are neutral sentiment. Based on that, the authors have proposed and implemented the algorithm presented in detail as below.

```
Data JSON file: topic word for each feature
                opinion word (negative, positive, neutral) for each feature
Result: Sentiment on Aspect
For each review
        Scan the comment text
        If the line contains feature word:
                For each aspect do
                Pos <= Sum positive opinion count
                Neg <= Sum negative opinion count
                Neu <= Sum neutral opinion count
                Sum = Pos.1 + Neg.(-1) + Neu.(0)
                If Sum > 0:
                        Opinion = "Pos"
                Else if Sum < 0:
                        Opinion = "Neg"
                Else:
                        Opinion = "Neu"
        Else:
                Opinion = "Neu"
End
```

Figure 4. Pseudocode search for the sentiment on each aspect

Source: Author's proposal

Phase 4: Visualization

The overview of the product review will demonstrate the opinion of each aspect, the average of stars, the number of rating tags, the total count of customers, the revenue of each material,... The slicers are utilized to classify by time, materials, reviews, stars. To deeper understand the emotions of five aspects, the performer displays some column charts, line charts with percentages and values, and images by word cloud.

4. Research results

4.1. Summary overview of Cocoon's product

After summing the statistics, there are 17,892 valid comments in Figure 5. Total revenue from these customers is about 4 billion VND. Sorted by components, rose and winter melon are the two ingredients with the largest sales and revenue. The average star rating of the comments is very high: 4.94 stars. It shows that the number of customers with 5-star reviews is so high. Customers attach rating tags to each comment very little. Up to 11,133 comments were not tagged, accounting for 62% of the total number.

Figure 5. Overview of the result of the product review

Source: The author's data analysis

Looking at the bar chart images, the product quality evaluation of customers is very good. It takes about 50% of the total and is significantly more proportional than other aspects. The lowest is the price and brand of satisfaction for products and services. As for the negative opinion, the product price received the least negative reviews, with 15 reviews. The number of reviews for bad transportation is the highest in the five dimensions which are 266 reviews in total. Therefore, this aspect needs to be more thoughtful about issues that cause bad feelings for customers. Based on the total number of good and bad emotions, it is clear that product quality, packing, shipment, price, and customer services are in the order of customer concern.

4.2. A detailed description of each aspect

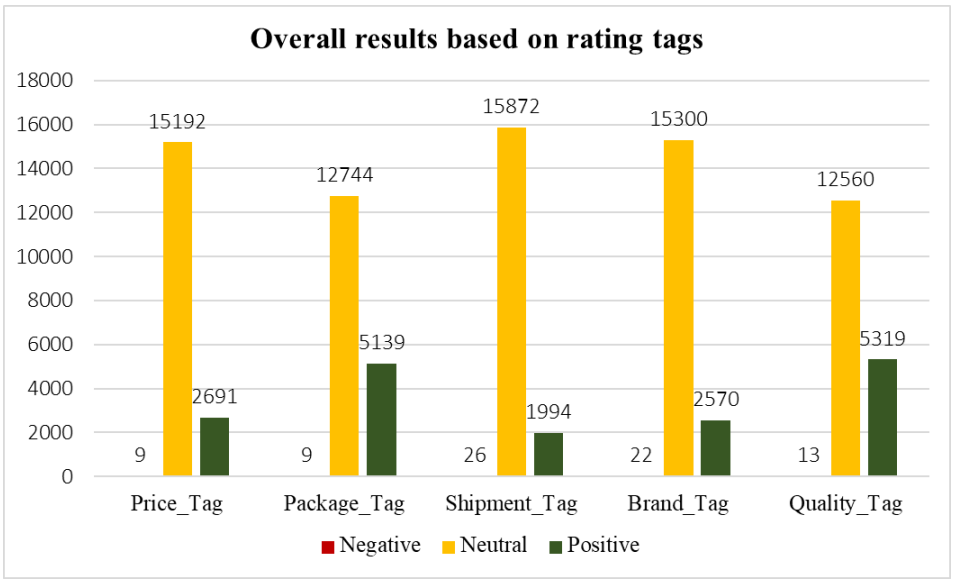

Figure 6. Overall results on five aspects based on rating tags

Source: The author's data analysis

It is obvious that most customers do not use many tags and do not even use tags to evaluate products in Figure 6. Regarding the positive rating tags, the quality and packaging are superior to the other aspects. Regarding product quality, there are 5,319 positive tags and 13 negative tags. Next, the packaging has 09 bad tags and 5,139 good tags. In addition, regarding shipping, there are 1,994 positive tags and 26 negative tags. In this aspect, the number of negative comments tags is the highest, consisting of less than 1%. In a nutshell, the number of tags is used quite a little and customers usually select positive tags.

4.3. Aspect-based sentiment analysis over time

Over time, the number of reviews received first with 16,107 reviews in 2021; second with 1,167 reviews in 2020; third with 614 reviews in 2022; and fourth with at least four reviews in 2019.

Figure 7. The total count of comments on five aspects by the year

Source: The author's data analysis

Based on the line graph in Figure 7, there is a variation in the number of positive, negative, and neutral comments over the years. The number of neutral reviews is the highest, above positive, and negative ones every year. By and large, the quantity of comments has increased with time, as have the opinions about each aspect. In 2021, it has the biggest difference in customer emotions. In 2019, there were only 04 comments and they were all neutral. The first period from 2019 to 2020 has a slight change in the number of sentiments per aspect, and the period from 2020 to 2021 has a sharp increase in volatility when the number rises much more rapidly. It is predicted that 2022 will sell in more quantities than the previous year and have a clearer polarization.

4.4. Word cloud for aspects

We used word cloud dashboards to clearly represent each aspect of all the goods studied in the case study after accumulating the texts in Figure 8.

Figure 8. Word cloud for five aspects

Source: The author's data analysis

It is clear that the first aspect clearly illustrates the significant words such as "sale", "giá" (price), and "rẻ" (cheap). The "sale" word is the best important in the price aspect of all words. Looking at the packing aspect, the keyword shown most clearly is "đóng-gói" (pack) which is the biggest size in the packaging aspect. In general, the words displayed in the graphic, are mostly positive emotions. This shows even more clearly that this aspect receives a lot of good customer reviews. The graphics of aspects via the word cloud technique, it has shown more clearly the level of use of words by customers. Overall, words with positive words are used a lot more than negative terms and neutral texts.

5. Discussion

In terms of vegan products, this study was more specific about the sentiment of each aspect of theory and method with regard to the role of aspect-based sentiment analysis from customers' feedback in Vietnam's online market while previous studies concentrated on discovering the interesting themes in consumer opinions (Cooper et al., 2022; Pilař et al., 2021). Further, this research provides the finding with emerging trends and potential merchandise for each factor in market opinions through each period. Over and above the significant increase in the number of customer comments also shows that vegetarian products are being used more and more by customers. In sum, there are important theoretical implications of using the method to analyze the customers' opinions over time.

According to the results obtained, the rate of positive comments is many times larger than the rate of negative comments. There are still very few customers who are not satisfied with goods and services. The order of magnitude of customer interest is quality, package, shipment, price, and brand. Some adjustments need to improve the buyers' experience by providing better products and services. Following that, we have the following suggestions for those aspects.

Vegetarian merchandise is a vegan and nature-friendly product, so the **packaging** process may be different from other products by using packaging materials made from biodegradable plastic and plant fibers to shorten the time to decompose. Air bubble foam is a popular material used for shock-absorbing packaging, which protects the product very well, but it takes about 10 to 100 years to degrade depending on environmental conditions. Therefore, the store ought to consider switching to other materials like kraft honeycomb wrapping paper, carton paper, and biodegradable plastic.

For E-commerce platforms, **shipment** plays a role in distributing products to users. Thus, product suppliers need to consider when there is a lot of traffic, there will be a balance of forces to contribute to early delivery to the supplier. Currently, Shopee supports tracking their orders on the app, which has helped customers see the status of the orders. Besides, it is more convenient for customers to provide additional support by sending an email. When the delivery person is about to convey the goods to customers, about thirty minutes before, it should be an automatic message sent to the customer's phone to notify in advance to prepare for a faster pick-up time and reduce the waiting time of the shipper.

Regarding the **price**, there are many different pricing strategies for businesses to choose to implement in line with their business goals. In Cocoon's case, the company can use the "Product bundle pricing" campaign. Currently, on the official Cocoon store, some merchandise is sold in several combos, but all kinds of them are not diverse so that not only businesses increase revenue but also customers have more choices. Combos can be formed by purchasing history in the past, which is very practical for the buyer.

All orders must be checked carefully, especially when there is a promotion going on. The attention to each customer shows them personalization, creating the image of the **brand** in their

hearts. For example, the store has a special promotion program for customers who buy products and give gifts on sale occasions. Customers who usually want to give gifts must add plus products to their shopping cart, then they will have gifts when packing the products. If the customer does not add to the cart, the customer service department can contact to confirm whether the customer wants to receive the gift or not.

The **quality** product is many times researched and tested to make up new merchandise. Therefore, this factor has been considered by the company itself. Besides, this aspect is also rated the best and most satisfied by their customers. Hence, the authors do not recommend this aspect to the company.

6. Conclusion and future work

6.1. Results

With the trend of using vegan products in the world, Vietnam is also a country that favors vegetarian products thanks to the benefits not only for health but also for the environment. We looked back at and referred to previous studies and carried out the experimental study using a lexicon-based approach which classifies the sentiment of each aspect in customer comments. In this case study, Cocoon cosmetics was studied with data collected from Shopee websites with 32,216 comments, through preprocessing the remaining 17,892 comments. Through the process of researching and analyzing the viewpoints of five aspects: price, package, shipment, brand, and quality based on the proposed model. A star rating is close to reaching a maximum of five stars. Regarding text comments, the quality aspect has a higher number of positive opinions than other aspects, making up 8,672 good comments. In contrast, the shipping aspect received the most negative reviews but also only accounted for 266 comments. The most frequency is "hàng" (merchandise), "sản-phẩm" (product), "dùng/sử-dụng" (use), "giao" (delivery), and "mua" (buy) that are used by customers when they show their opinions about the merchandise and services.

6.2. Limitations and future work

The number of stars is a general comment for a product that has not highlighted detail about each of the elements. In the implementation process, there are still many stages of manual implementation and it is difficult to scale up to develop for classifications in the next. Therefore, building an automatic implementation model will reduce project execution time.

This research direction needs to be built that automatically retrieves data and builds the dictionary owing to this manual phase. Customers' reviews can attach some images and short videos, and they will also receive coins if any reviews meet the standard of the website. Thus, the new aspects can be expanded to comprehend customer interest in the future.

References

Alqaryouti, O., Siyam, N., Monem, A. A., & Shaalan, K. (2020). Aspect-based sentiment analysis using smart government review data. *Applied Computing and Informatics*, Vol. ahead-of-print No. ahead-of-print. doi: 10.1016/j.aci.2019.11.003

Anoop, V. S., & Asharaf, S. (2020). Aspect-oriented sentiment analysis: a topic modeling-powered approach. *Journal of Intelligent Systems, 29*(1), 1166-1178. doi:10.1515/jisys-2018-0299

Bahri, S., Bahri, P., & Lal, S. (2018). A novel approach of sentiment classification using emoticons. *Procedia Computer Science, 132*, 669-678. doi:10.1016/j.procs.2018.05.067

Berg, J., & Henriksson, M. (2020). In search of the 'good life': Understanding online grocery shopping and everyday mobility as social practices. *Journal of Transport Geography, 83*(C), 102633-10264. doi:10.1016/j.jtrangeo.2020.102633

Choi, H. Y., & Joung, H. W. (2018). Exploring restaurant selection attributes of vegetarian customers. *Culinary Science & Hospitality Research, 24*(7), 83-88. doi:10.20878/cshr.2018.24.7.009

Choi, H., Joung, H. W., Choi, E. K., & Kim, H. S. (2021). Understanding vegetarian customers: The effects of restaurant attributes on customer satisfaction and behavioral intentions. *Journal of Foodservice Business Research, 25*(3), 353-376. doi:10.1080/15378020.2021.1948296

Clara, A. Y., Adiwijaya, A., & Purbolaksono, M. D. (2020). Aspect based sentiment analysis on beauty product review using random forest. *Journal of Data Science and Its Applications, 3*(2), 67-77. doi:10.34818/jdsa.2020.3.58

Cooper, K., Dedehayir, O., Riverola, C., Harrington, S., & Alpert, E. (2022). Exploring consumer perceptions of the value proposition embedded in vegan food products using text analytics. *Sustainability, 14*(4), Article 2075. doi:10.3390/su14042075

Dang, T. V., Nguyen, K. V., & Nguyen, N. L. T. (2018). A supervised method for aspect-based sentiment analysis. In *The Fifth International Workshop on Vietnamese Language and Speech Processing (VLSP)* (pp 21-24). Ho Chi Minh City, Vietnam: University of Information Technology, VNU-HCM, Vietnam

Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment parsing from small talk on the web. *For Amazon: Sentiment Parsing from Small Talk on the Web, 53*(9), 1375-1388. doi:10.2139/ssrn.276189

de Kok, S., Punt, L., van den Puttelaar, R., Ranta, K., Schouten, K., & Frasincar, F. (2018). Review-aggregated aspect-based sentiment analysis with ontology features. *Progress in Artificial Intelligence, 7*(4), 295-306. doi:10.1007/s13748-018-0163-7

Dinu, M., Abbate, R., Gensini, G. F., Casini, A., & Sofi, F. (2017). Vegetarian, vegan diets and multiple health outcomes: A systematic review with meta-analysis of observational studies. *Critical Reviews in Food Science and Nutrition, 57*(17), 3640-3649. doi:10.1080/10408398.2016.1138447

Expertmarketresearch. (n.d.). *Global vegan food market, expert market research*. Retrieved April 03, 2022, from https://www.expertmarketresearch.com/reports/vegan-food-market

Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data, 2*(1), 1-14. doi:10.1186/s40537-015-0015-2

Farkhod, A., Abdusalomov, A., Makhmudov, F., & Cho, Y. I. (2021). LDA-based topic modeling sentiment analysis using Topic/Document/Sentence (TDS) model. *Applied Sciences, 11*(23), 11091-11106. doi:10.3390/app112311091

Fortune Business Insights. (2022*). Vegan food market size, share & Covid-19 impact analysis, product type (Vegan meat, ... stores, online retails, and others), and regional forecast, 2021-2028.* Retrieved April 03, 2022, from https://www.fortunebusinessinsights.com/vegan-food-market-106421

Gary, D. (2021). *The online vegan marketplace is helping independent brands thrive.* Retrieved April 03, 2022, from https://www.forbes.com/sites/garydrenik/2021/04/06/the-online-vegan-marketplace-is-helping-independent-brands-thrive/?sh=405daf4345df

Github. (2017). *VietSentiWordNet a quick and simple method to find opinion for Vietnamese text*. Retrieved January 25, 2022, from https://github.com/sonvx/VietSentiWordNet

Github. (2019). *Building a program to build Vietnamese stopwords based on IDF using scikit-learn*. Retrieved January 12, 2022, from https://github.com/ltkk/vietnamese-stopwords/blob/master/make_stopwords.py

Github. (2020). *Aspect sentiment analysis Vietnam*. Retrieved January 25, 2022, from https://github.com/TorRient/ASA-VN

Han, Y., Liu, M., & Jing, W. (2020). Aspect-level drug reviews sentiment analysis based on double BiGRU and knowledge transfer. *IEEE Access, 8*, 21314-21325. doi:10.1109/ACCESS.2020.2969473

Holbrook, M. B., & Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research, 9*(2), 132-140

Hu, M., & Liu, B. (2004). Mining opinion features in customer reviews. *AAAI, 4*(4), 755-760.

Hung, C., & Cao, Y. X. (2019). Sentiment classification of Chinese cosmetic reviews based on integration of collocations and concepts. *The Electronic Library, 30*(1), 155-169. doi:10.1108/EL-04-2019-0093

Kotler, P. (2000). *Marketing management: The millennium edition* (Vol. 199). Upper Saddle River, NJ: Prentice Hall

Le, V. N. B., & Huh, J. H. (2021). Applying sentiment product reviews and visualization for BI systems in Vietnamese E-commerce website: Focusing on Vietnamese context. *Electronics, 10*(20), 2481-2501. doi:10.3390/electronics10202481

Leitzmann, C. (2003). Nutrition ecology: The contribution of vegetarian diets. *The American Journal of Clinical Nutrition, 78*(3), 657S-659S. doi:10.1093/ajcn/78.3.657S

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing, 80*(6), 69-96. doi:10.1509/jm.15.0420

Liu, B. (2020). Sentiment analysis: Mining opinions, sentiments, and emotions. In *Studies in natural language processing* (pp. 01-17). Chicago, IL: University of Illinois, Cambridge University Press.

Lu, L., & Gursoy, D. (2017). Does offering an organic food menu help restaurants excel in competition? An examination of diners' decision-making. *International Journal of Hospitality Management, 63*, 72-81. doi:10.1016/j.ijhm.2017.03.004

Mai Long & Le, B. (2018). Aspect-based sentiment analysis of vietnamese texts with deep learning. In *Asian conference on intelligent information and database systems* (pp. 149-158). Cham, Switzerland: Springer.

Nandal, N., Tanwar, R., & Pruthi, J. (2020). Machine learning based aspect level sentiment analysis for Amazon products. *Spatial Information Research, 28*(5), 601-607. doi:10.1007/s41324-020-00320-2

Ngo, T. V. Q., Vo, D. T. A., Ngo, P. A., Nguyen, A. D. M., Le, T. M., To, L. T. P., & Nguyen, P. T. T. (2021). Factors influencing on consumer's decision on vegetarian diets in Vietnam. *The Journal of Asian Finance, Economics, and Business, 8*(2), 485-495. doi:10.13106/jafeb.2021.vol8.no2.0485

Nguyen, B., Nguyen, H. V., & Ho, T. (2021). Sentiment analysis of customer feedback in online food ordering services. *Business Systems Research, 12*(2), 46-59. doi:10.2478/bsrj-2021-0018

Nguyen, K. N. D., & Duong, T. H. (2019). One-document training for Vietnamese sentiment analysis. *International Conference on Computational Data and Social Networks,* 189-200. doi:10.1007/978-3-030-34980-6_21

Nguyen, N. D. (2021). *E-commerce report quarter III/2021: E-commerce race between Vietnam, Thailand and Malaysia*. Retrieved February 20, 2022, from https://iprice.vn/xu-huong/insights/bao-cao-tmdt-quy-iii2021-cuoc-dua-tmdt-giua-viet-nam-thai-lan-va-malaysia

Nguyen, P. H., Ngo, T. D., Phan, D. A., Dinh, T. P., & Huynh, T. Q. (2008). Vietnamese spelling detection and correction using Bi-gram, Minimum Edit Distance, SoundEx algorithms with some additional heuristics. *IEEE International Conference on Research, Innovation and Vision for the Future in Computing and Communication Technologies,* 96-102. doi:10.1109/RIVF.2008.4586339

Nguyen, Q. K. P., & Nguyen, K. V. (2020). Exploiting Vietnamese social media characteristics for textual emotion recognition in Vietnamese. *International Conference on Asian Language Processing (IALP)*, 277-279. doi:10.48550/arXiv.2009.11005

Nguyen, T. L. T., Nguyen, X., & Dinh, D. (2016). Vietnamese plagiarism detection method. *Proceedings of the Seventh Symposium on Information and Communication Technology,* 44-51. doi:10.1145/3011077.3011109

Nilashi, M., Ahmadi, H., Arji, G., Alsalem, K. O., Samad, S., Ghabban, F., & Alarood, A. A. (2021). Big social data and customer decision making in vegetarian restaurants: A combined machine learning method. *Journal of Retailing and Consumer Services, 62*(C), 102630-102648. doi:10.1016/j.jretconser.2021.102630

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, *10*, 79-86. doi:10.3115/1118693.1118704

Petaasia. (2017). *Hồ Chí Minh named one of the top 10 most vegan-friendly cities in Asia 2017 [Ho Chi Minh named one of the top 10 most vegan-friendly cities in Asia 2017]*. Retrieved April 03, 2022, from https://www.petaasia.com/media/ho-chi-minh-named-one-top-10-most-vegan-friendly-cities-asia

Pilař, L., Stanislavská, L. K., Kvasnička, R., Hartman, R., & Tichá, I. (2021). Healthy food on Instagram social network: Vegan, homemade and clean eating. *Nutrients, 13*(6), 1991-2010. doi:10.3390/nu13061991

Piryani, R., Gupta, V., Singh, V. K., & Ghose, U. (2017). A linguistic rule-based approach for aspect-level sentiment analysis of movie reviews. In *Advances in computer and computational sciences* (pp. 201-209). doi:10.1007/978-981-10-3770-2_19

Research and Markets. (2020). *Vietnam skin care products market by type, demographics, age group and sales channel: Opportunity analysis and industry forecast, 2021-2027.* Retrieved January 20, 2022, from https://www.researchandmarkets.com/reports/5157263/vietnam-skin-care-products-market-by-type#src-pos-12

Rivera, M., & Shani, A. (2013). Attitudes and orientation toward vegetarian food in the restaurant industry: An operator's perspective. *International Journal of Contemporary Hospitality Management, 25*(7), 1049-1065. doi:10.1108/IJCHM-07-2012-0116

Sharma, R., Nigam, S., & Jain, R. (2014). Determination of polarity of sentences using sentiment orientation system. *International journal of Advances in Computer Science and Technology (IJACST) WARSE, 3*(3), 182-187

Statista. (2021). *Value of the worldwide vegan food market from 2020 to 2021 with a forecast for 2025*. Retrieved April 03, 2022, from https://www.statista.com/statistics/1280275/value-of-the-global-vegan-food-market

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics, 37*(2), 267-307. doi:10.1162/COLI_a_00049

Turney, P. D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 417-424. doi:10.48550/arXiv.cs/0212032

Vanaja, S., & Belwal, M. (2018). Aspect-level sentiment analysis on e-commerce data. *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)*, 1275-1279. doi:10.1109/ICIRCA.2018.8597286

Verain, M. C., Dagevos, H., & Antonides, G. (2015). Sustainable food consumption. Product choice or curtailment? *Appetite, 91*, 375-384. doi:10.1016/j.foodqual.2021.104267

Voxco. (n.d.). *What is the connection between digital transformation and customer experience*. Retrieved April 03, 2022, from https://www.voxco.com/blog/what-is-the-connection-between-digital-transformation-and-customer-experience

Vu, T. T., Pham, T. H., Luu, T. C., & Ha, T. Q. (2012). *A feature-based opinion mining model on product reviews in Vietnamese*. In R. Katarzyniak, T. F. Chiu, C. F. Hong & N. T. Nguyen (Eds.), *Semantic methods for knowledge management and communication* (vol. 381). doi:10.1007/978-3-642-23418-7_3

Wang, W., Pan, S. J., Dahlmeier, D., & Xiao, X. (2016). Recursive neural conditional random fields for aspect-based sentiment analysis. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 616-626. doi:10.18653/v1/D16-1059

Wiebe, J. (2000). Learning subjective adjectives from corpora. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence* (pp. 735-740). Palo Alto, CA: American Association for Artificial Intelligence

Wikipedia. (2022). *List of emoticons*. Retrieved January 16, 2022, from https://en.wikipedia.org/wiki/List_of_emoticons

Wikipedia. (2022). *Rules for placing accent marks in the national language script*. Retrieved January 16, 2022, from https://vi.wikipedia.org/wiki/Quy_tắc_đặt_dấu_thanh_trong_chữ_quốc_ngữ

Yauris, K., & Khodra, M. L. (2017). Aspect-based summarization for game review using double propagation. *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*, 1-6. doi:10.1109/ICAICTA.2017.8090997