JDI
JOURNAL OF DEVELOPMENT AND INTEGRATION
TẠP CHÍ PHÁT TRIỂN VÀ HỘI NHẬP
Số 75 (2024) 75-81 | jdi.uef.edu.vn

# Exploring the customer experience of Vietnam Airlines via sentiment analysis

**Ton Nguyen Trong Hien** *
Van Lang University

KEYWORDS

Natural Language Toolkit,
Sentiment analysis,
Skytrax,
VADER,
Vietnam Airlines.

ABSTRACT

The sentiment analysis of social media has gained much attention due to the succinct nature of comments, enabling access to public opinions. Using this technique, the present study aimed to explore customer sentiments regarding their experiences with Vietnam Airlines, the national carrier. According to the Skytrax evaluation, Vietnam Airlines achieved a 4-star rating out of 5, indicating the need for service improvements to attain perfection. The significance of this research lies in providing a comprehensive perspective on customer experience evaluations. To achieve this, we utilized the Valence Aware Dictionary for Reasoner (VADER) to classify sentiments expressed in social networking forum data. The dataset comprised 670 reviews from the Skytrax forum, which underwent a data pre-processing phase. The results exhibited favorable accuracy in detecting ternary and multiple sentiment classes. The comparative analysis between visual customer ratings and their corresponding comments yielded similar outcomes. Additionally, the results highlighted the need for a re-evaluation of the in-flight entertainment system.

## 1. Introduction

In contemporary society, characterized by constant advancements, people's travel demands have witnessed a notable surge. As a result, passenger transportation services have experienced rapid development across various modes, including road, rail, sea, and air. Of particular significance, air travel offers advantages such as swift transportation, frequent flight schedules, and relatively affordable fares, catering to the diverse travel requirements of different societal segments. Vietnam Airlines (VNA) was established as a prominent air transport entity under state administration. As Vietnam's pioneering airline, VNA has achieved the prestigious Skytrax certification for three consecutive times, solidifying its position as a reputable 4-star international carrier (The British organization Skytrax, renowned for its independent and leading global survey program, has been a prominent authority in evaluating the service quality of airlines and airports). Similar to many other businesses, VNA strives to maintain its reputation and expand its operations by upholding and enhancing its service standards. However, a crucial management concern arises in comprehending customer opinions to address the identified areas of improvement. Given the substantial volume of data involved, traditional

* Corresponding author. Email: *hien.tnt@vlu.edu.vn* (Ton Nguyen Trong Hien)

https://doi.org/10.61602/jdi.2024.75.10
Received: 13/12/2023; Revised: 15/01/2024; Accepted: 22/01/2024
ISSN (print): 1859-428X, ISSN (online): 2815-6234

approaches like surveys or manual reading present considerable challenges in acquiring and analyzing this information effectively.

Acquiring customer feedback serves as a valuable approach for businesses to gain insights into the strengths and weaknesses of products or services. This enables them to promptly meet customer expectations and deliver exceptional offerings. In today's era of remarkable scientific and technological advancements, particularly the proliferation of the Internet and platforms like YouTube, Twitter, or forums, individuals have not only gained the ability to share information but also to express their opinions and perspectives on various products, services, and societal matters. Consequently, social networking forum has been considered as a highly significant and extensive source of valuable information. Undoubtedly, the forum serves as a valuable repository of real-time user-generated data. This platform provides a space where individuals can freely express their viewpoints, ideas, and feedback based on their experiences. By scrutinizing these user comments, enterprises can gain insights into customers' perceptions of their products or services.

However, attaining a comprehensive understanding is far from a straightforward task. The sheer magnitude of data generated presents significant complexities when relying on manual processes. Moreover, social media data compounds the challenge with its inherent linguistic diversity, rendering automated approaches equally arduous. Additionally, social networks host contentious discussions covering a wide range of subjects. To address these complexities, this paper employs VADER sentiment analysis as a means to discern public sentiment by analyzing their feedback concerning the services offered by VNA.

## 2. Literature review

### *2.1. Customer experience of airline services and components of customer experience*

Customer experience of airline services encompasses a range of experiences from the moment they begin the ticket booking process until they complete their journey (Laming & Mason, 2014). This includes the search, booking, check-in, boarding, in-flight services, airport procedures, and post-flight experience. This experience can influence customers' feelings towards the airline and their future decisions when choosing an airline. Components of customer experience include:

+ Seat Comfort: This refers to the comfort level of the seats provided on the aircraft, including factors such as seat size, legroom, recline, and overall ergonomics.

+ Cabin Staff Service: This encompasses the quality of service provided by the cabin crew, including their friendliness, responsiveness to passenger needs, professionalism, and overall customer interaction during the flight.

+ Food & Beverages: This relates to the quality, variety, and presentation of food and beverages offered during the flight, including meal options, freshness, taste, and availability of complimentary or for-purchase items.

+ Ground Service: This includes the customer experience during pre-flight and post-flight activities on the ground, such as check-in procedures, baggage handling, security screening, boarding processes, and overall efficiency and friendliness of ground staff.

+ Value for Money: This assesses whether the price paid for the airline service aligns with the perceived benefits and overall experience provided to the customer, taking into account factors such as ticket cost, additional fees, service quality, and overall satisfaction with the flight experience.

Unsurprisingly, previous studies have extensively addressed the issue of service improvement within the context of VNA. The research literature reveals a focus on investigating the practical significance of this topic, particularly through the synthesis and evaluation of the quality of VNA's ground services. These assessments are typically conducted through the distribution of behavioral survey questionnaires to customers at Noi Bai, as well as utilizing survey and evaluation sources in Da Nang and Tan Son Nhat. However, it is worth noting that the existing study primarily delves into the factors influencing the quality of ground services, without further exploration of other dimensions (Phan, 2018). The airport terminal appears to be a subject of significant interest, as evidenced by the attention it has received in various studies. For instance, Tran (2019) explored solutions to enhance customer satisfaction regarding the quality of check-in services specifically at the VNA check-in counter located within the Domestic Terminal of Tan Son Nhat International Airport. Additionally, Wang and Pham (2020) recommended the adoption of a proactive service model by VNA to assess and improve their existing service network at international airports. These studies exemplify the growing body of research focused on addressing key aspects of airport terminal operations and service enhancements within the context of VNA.

The existing body of literature reveals a notable gap in research concerning a comprehensive review of VNA's services. A comprehensive examination encompassing other dimensions of VNA's service offerings including their meals, services and so on.

### *2.2. Sentiment analysis*

Every day, a substantial amount of data is generated across various platforms, including social media, blogs, websites, and multimedia. The widespread adoption of social media platforms such as Facebook, TikTok, and YouTube has notably contributed to the exponential growth in data generation. As a result, the field of sentiment analysis has garnered increasing attention from researchers. Sentiment analysis, a method of text analysis, aims to identify the polarity, whether positive or negative, expressed within a given text, encompassing entire documents, paragraphs, sentences, or clauses. Its primary objective is to assess the attitudes, sentiments, evaluations, and emotions of the speaker or writer through the computational analysis of subjectivity present in textual content. Several notable studies have made valuable contributions to the field of sentiment analysis in recent years, including the works of Bokaee et al. (2022) and He et al. (2022). The primary objective of text summarization is to obtain concise summaries of textual content. This summarization technique finds applications in various domains, such as searching for documents related to specific topics, generating overviews of news articles, condensing emails, extracting summaries of medical information, and providing briefs of scientific articles. The process of text summarization involves several steps, including topic identification, interpretation, and summary generation. While traditional classification algorithms can be employed to train sentiment classifiers using manually labelled text data, such as the following: Naive Bayes, Multi-nominal NB, etc.; it is worth noting that manual labeling is a resource-intensive and time-consuming task.

In recent years, there has been a growing focus on the social aspects within the context of libraries, including investigations into the emotions and sentiments experienced and expressed by library users. Many of these studies leverage established sentiment analysis tools such as Python's Natural Language Toolkit and VADER Sentiment Analyzer which have been primarily trained on product and service reviews (Jongeling et al., 2015; Elbagir & Yang, 2020). This approach is considered more effective than traditional algorithmic analysis methods (Pai et al., 2022; Passi & Motisariya, 2022); "This approach made the review sentiment process of thousands of data faster and more accurate" (Budianto et al., 2022) as the current approach effectively addresses the challenges associated with pre-analysis difficulties. However, it is crucial to be cautious of potential issues inherent in unstructured text analysis, particularly when dealing with free language usage, slang, and unconventional notations found on social networks. These factors can pose challenges and increase the likelihood of missed or misinterpreted analyses. Thus, VADER is an established model that has been pre-trained using rule-based values specifically tailored to sentiments expressed on social media platforms. This model enables the evaluation of text messages by providing assessments not only on the polarity of sentiments (positive or negative) but also on the intensity of those emotions (Hutto & Gilbert, 2014). VADER utilizes a comprehensive dictionary of terms for evaluation, which includes various examples such as:

Negations: Modifiers that reverse the meaning of a phrase (e.g., "not good").

Contractions: More complex forms of negations (e.g., "wasn't good").

Punctuation: Punctuation marks that indicate increased intensity (e.g., "It's great!!!").

Slang: Variations of slang words, such as "watcha" (what are you), "hella" (very/extremely)

Based on the aforementioned points, it is evident that the topic under consideration, specifically the services provided by VNA, has not been thoroughly explored. The existing literature lacks a comprehensive examination of VNA's services. Therefore, in light of this gap in the research, and by considering the appropriate methodology, we have determined the following research topic: Exploring the customer experience of VNA via sentiment analysis utilizing the VADER application.

## 3. Methodology

This study employs text-mining techniques to analyze the data. The research is divided into distinct phases (Figure 1). The first phase focuses on data collection and pre-processing. In the next phase, the study incorporates the data analysis.

### *3.1. Data preparation & pre-processing*

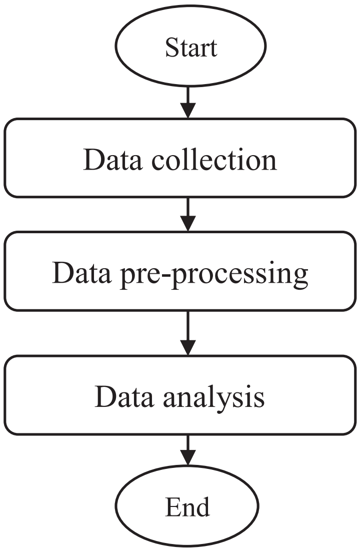


**Figure 1. Research steps**

*Source: Author analyzes*

The choice to study customer experience perceptions of VNA is justified as it is the national airline of Vietnam and plays a significant role in the country's aviation industry. This selection also stems from the fact that Vietnam Airlines' services are not yet entirely perfect and require improvement based on customer feedback and perceptions.

The data for this study is sourced from the reputable international air transport rating and organization unit, Skytrax, available on their website (https://www.airlinequality.com). Specifically, the data is extracted from an aviation forum page within the website where passengers share their experiences and feedback regarding the services provided by airlines. The purpose of this platform is to assist prospective passengers in making informed decisions based on the experiences and preferences of others. To ensure data reliability, Skytrax requires customers submitting reviews to provide a copy of their flight ticket and supporting photographic evidence. This verification process conducted by the Skytrax admin helps to ascertain the authenticity of the reviewer's experience, thereby mitigating the presence of undesired untrusted comments. Consequently, the dataset can be considered relatively reliable due to this measure taken to validate the reviews. The data used for analysis is derived from publicly available datasets without extracting any user information.

Before the analysis, pre-processing steps were implemented on the data. This involved removing whitespace and duplicate entries, as well as normalizing the text by converting all letters to lowercase.

At the time of extraction, it was 01:19 PM on September 15th, 2023.

* The data collected from the website consists of publicly available information and is not intended for commercial purposes. No personal or sensitive data is gathered, and thus it is considered valid (Liu & Chen, 2021).

### *3.2. Data analysis*

The analysis involves the application of data mining techniques using the Python programming language. The data that are present in the online social medias are highly unstructured, and thus sentiment analysis was applied. In sentiment analysis, there are two approaches: machine learning and lexicons. Dhaoui et al. (2017) argue that the results are nearly equivalent. Personally, it is believed that in certain complex contexts, such as free-form topic comments on social media, machine learning methods are applied. However, in situations where the discussion topics are predictable (related to aviation objects in this case), lexicon methods are employed. In this context, one such technique is VADER, which is an open-source, rule-based opinion analysis library available in NLTK (Natural Language Toolkit) was utilized.

## 4. Results and Discussion

### *4.1. Results*

The data extraction process yielded a total of 670 comments. Preliminary analysis of the data revealed descriptive statistics, indicating an average rating of 6/10 across multiple criteria, namely Food and beverage, Comfortable Seating, Staff, Expenses, and Inflight Entertainment. According to the collected data, the average rating score for ticket classes is 6/10. The analysis further indicates that the in-flight entertainment criterion received the lowest rating, with a score of 2/5 points, while the remaining criteria obtained a rating of 3/5 points each. In light of these findings, the evaluation details for different ticket classes are presented in Table 1.

To enhance the depth of analysis, VADER enables the calculation of sentiment in text and furnishes the probability of classifying a given input sentence as positive, negative, or neutral. The Sentiment Intensity

**Table 1. Distribution for evaluation of airfare classes**

| First class | Business class | Premium economy | Economy class |
|---|---|---|---|
| 2 | 176 | 53 | 439 |

*Source: Author collects*

Analyzer in NLTK generates sentiment scores based on a lexicon-based approach, wherein each word in the text is assigned a sentiment score. These scores are computed using a blend of lexical, grammatical, and syntactic heuristics. The SentimentIntensityAnalyzer in the NLTK library provides score values following the formula:

$$\frac{x}{\sqrt{x^2 + \alpha}}$$

In VADER, alpha is set to be 15 which approximate the maximum expected value of x (Amin, Hossain, Akther & Alam, 2019).

The sentiment scores generated by VADER can be interpreted as follows: The "Negative" score corresponds to the probability of the text being classified as negative. The "Neutral" score indicates the probability of the text being categorized as neutral. Finally, the "Positive" score represents the probability of the text being identified as positive.

Traditional scoring methods adhere to the following principles:

Positive Score: Words or phrases expressing positivity, such as "happy" or "excellent", receive high scores.

Negative Score: Words or phrases associated with negativity, such as "sad" or "terrible", receive low scores.

Neutral Words: Words or phrases lacking strong emotional connotations, such as "the" or "is" are assigned average scores. In the present analysis, out of the 670 comments examined, approximately 80% exhibit neutral sentiment. Positive comments account for around 20% of the dataset, while negative comments constitute merely 5% (see Figure 2).

Researchers aiming to establish consistent criteria for classifying sentences as positive, neutral, or negative can find value in the following standardized thresholds, as identified in the study by Hutto and Gilbert (2014):

Positive sentiment: Sentences with a compound score greater than or equal to 0.05 are classified as positive.

Neutral sentiment: Sentences with a compound score between -0.05 and 0.05 (inclusive) are considered neutral.

Negative sentiment: Sentences with a compound score less than or equal to -0.05 are categorized as negative.

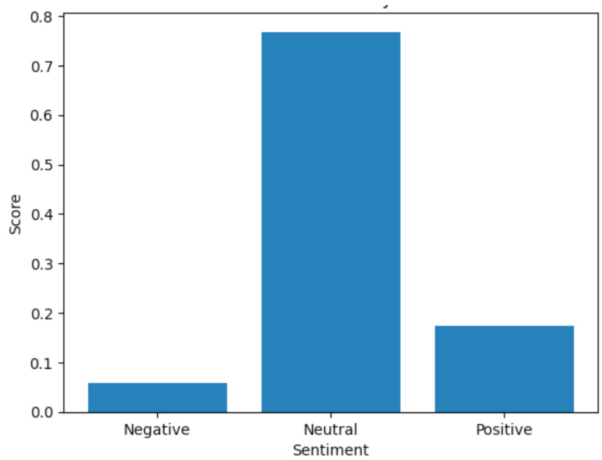


**Figure 2. Sentiment analysis**

*Source: Author analyzes*

```
# Determine the sentiment label
if compound_score >= 0.05:
    sentiment_label = "Positive"
elif compound_score <= -0.05:
    sentiment_label = "Negative"
else:
    sentiment_label = "Neutral"

# Print the sentiment label
print(f"Sentiment Label: {sentiment_label}")
```

```
Positive Score: 0.173
Negative Score: 0.059
Neutral Score: 0.768
Sentiment Label: Positive
```

**Figure 3. Results displayed from the Jupyter Notebook computing platform**

*Source: Author codes*

Figure 3 - The results of the analysis indicate the following scores: Positive Score: 0.173, Negative Score: 0.059, and Neutral Score: 0.768, which demonstrate:

Positive Score: The positive score represents the probability or intensity of positive sentiment in the analyzed text. In this case, the positive score is 0.173, indicating a moderate level of positive sentiment.

Negative Score: The negative score represents the probability or intensity of negative sentiment in the analyzed text. In this case, the negative score is 0.059, indicating a relatively low level of negative sentiment.

Neutral Score: The neutral score represents the probability or intensity of neutral sentiment in the analyzed text. In this case, the neutral score is 0.768, indicating a relatively high level of neutral sentiment.

Overall, the sentiment analysis result suggests that the analyzed text has a highly positive sentiment with a moderate level of positive sentiment, a low level of negative sentiment, and a relatively high level of neutral sentiment.

To further advance the analysis, we will proceed with the classification of keywords based on their categorization as nouns and adjectives, considering the frequency of their appearance.

Noun:

Prominently, the keywords Seat, Flight, Food, Staff, and Service stand out, with the issue of seating on the aircraft being mentioned the most, with 174 instances of repetition. Flight, Food, Staff, and Service correspondingly have frequencies of 120, 118, 110, and 96 in terms of appearance (Figure 4).

Adjectives:

Positive

We examine the words that indicate the occurrence level of a positive action or event within a sentence. The keyword "Comfort" appears most frequently, with 73 occurrences, followed by "Good" and "Valuable" with 57 and 55 occurrences, respectively. "Great" and "Friendly" occur 12 times each (Figure 5).

Negative

12 instances of emphasizing the old, 6 instances of repeating the slow progress, and 5 instances of the keyword "unworkable" appearing (Figure 6).

### *4.2. Discussion*

Compared to previous studies by Phan (2018), Tran (2019), Wang & Pham (2020) emphasize that the quality of VNA ground services still requires significant improvement. They argue that all airport clusters face issues, and overall, VNA encounters persistent challenges in ground services, specifically in the check-in process, boarding process, baggage handling, and document mishandling.

Contrary to this research, no ground service issues were identified, as social media platforms are currently

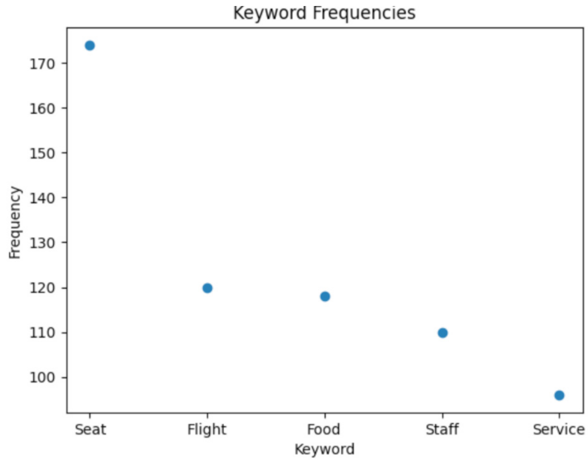


**Figure 4. Results displayed from Jupyter notebook computing platform**

*Source: Author analyzes*

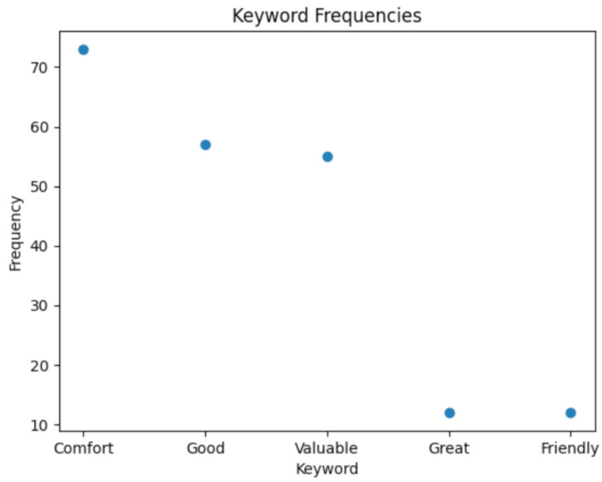


**Figure 5. Visualization of positive keywords**

*Source: Author analyzes*

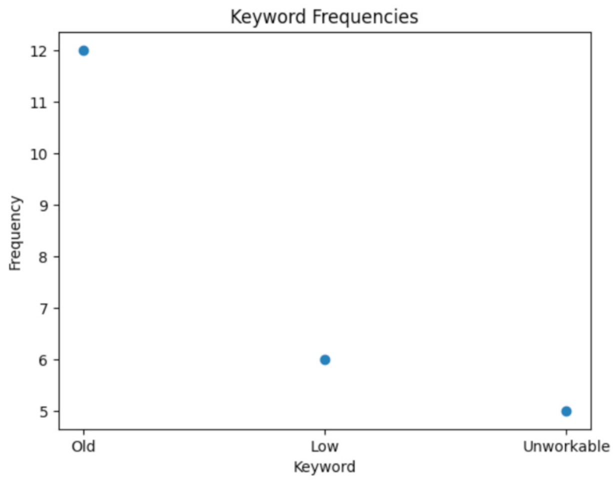


**Figure 6. Visualization of negative keywords**

*Source: Author analyzes*

inclined towards evaluating the in-flight experience as the primary focus. Based on the aforementioned analysis, we can make an informed speculation that there is evidence to suggest that issues related to comfortable seating, service experience, and on-board food are rated fairly well, and the service staff are perceived as friendly. The noun representing the entertainment system does not frequently appear, but negative keywords imply possible concerns about outdated equipment or malfunctioning.

## 5. Conclusion

The present study involved an analysis of customer comments in comparison with the star ratings on the VNA Skytrax forum page. The findings revealed a notable consistency between the star ratings and the corresponding comments provided by customers. The effective utilization of VADER sentiment analysis for processing large-scale data sets demonstrated VNA's ongoing efforts to address service-related concerns, particularly in the context of entertainment equipment replacement or repair. It is recommended that future research focuses on analyzing big data derived from social media platforms, as these platforms provide a valuable space where individuals openly share their genuine thoughts and experiences, offering an advantage over traditional survey methods. However, it should be noted that this study has certain limitations, primarily stemming from the relatively limited size of the data set. Despite the inclusion of all customer comments across various ticket classes during the analysis, future studies would benefit from obtaining larger data sets to enhance the generalizability and replicability of findings.

***Acknowledgment***

***We would like to express our gratitude to the reviewers for their valuable comments on the manuscript. We also would like to extend our sincere gratitude to Van Lang University (Vietnam) and to Muban Chom Bueng Rajabhat University (Thailand) for providing support for this research.***

### REFERENCES

Amin, A., Hossain, I., Akther, A., & Alam, K. M. (2019). Bengali Vader: A sentiment analysis approach using modified Vader. *2019 International Conference on Electrical, Computer and Communication Engineering* (ECCE). DOI: https://doi.org/10.1109/ecace.2019.8679144.

Bokaee Nezhad, Z., & Deihimi, M. A. (2022). Twitter sentiment analysis from Iran about COVID-19 vaccine. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews,* 16(1), 102367. DOI: https://doi.org/10.1016/j.dsx.2021.102367.

Budianto, A. G., Wirjodirdjo, B., Maflahah, I., & Kurnianingtyas, D. (2022). *Sentiment Analysis Model for klikindomaret android app during pandemic using vader and transformers NLTK library.* 2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). DOI: https://doi.org/10.1109/ieem55944.2022.9989577.

Dhaoui, C., Webster, C. M., & Tan, L. P. (2017). Social media sentiment analysis: lexicon versus machine learning. *Journal of Consumer Marketing*, 34(6), 480–488. https://doi.org/10.1108/jcm-03-2017-2141.

Elbagir, S., & Yang, J. (2020). Sentiment analysis on Twitter with Python's Natural Language Toolkit and vader sentiment analyzer. *IAENG Transactions on Engineering Sciences*. https://doi.org/10.1142/9789811215094_0005.

He, L., Yin, T., & Zheng, K. (2022). They may not work! an evaluation of eleven sentiment analysis tools on seven social media datasets. *Journal of Biomedical Informatics*, 132, 104142. https://doi.org/10.1016/j.jbi.2022.104142.

Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media,* 8, 1.

Jongeling, R., Datta, S., & Serebrenik, A. (2015). Choosing your weapons: On sentiment analysis tools for Software Engineering Research. *2015 IEEE International Conference on Software Maintenance and Evolution (ICSME).* DOI: https://doi.org/10.1109/icsm.2015.7332508.

Laming, C., & Mason, K. (2014). Customer experience - An analysis of the concept and its performance in airline brands. *Research in Transportation Business & Management,* 10, 15–25. https://doi.org/10.1016/j.rtbm.2014.05.004.

Liu, W., & Chen, P. (2021). Justification of the behavior regulatory pattern on data scraping. *Computer Law & Security Review,* 43, 105578. DOI: https://doi.org/10.1016/j.clsr.2021.105578.

Pai, A. R., Prince, M., & Prasannakumar, C. V. (2022). *Real-time twitter sentiment analytics and visualization using vader*. 2022 2nd International Conference on Intelligent Technologies (CONIT). DOI: https://doi.org/10.1109/conit55038.2022.9848043.

Passi, K., & Motisariya, J. (2022). Twitter sentiment analysis of the 2019 Indian election. *IOT with Smart Systems*, 805–814. DOI: https://doi.org/10.1007/978-981-16-3945-6_79.

Phan, T. T. H. (2018). *Nâng cao chất lượng dịch vụ mặt đất của Vietnam Airlines.* Luận văn Thạc sỹ, Đại học Quốc gia Hà Nội, Việt Nam.

Tran, X. B. (2019). *Giải pháp nâng cao sự hài lòng của khách hàng đối với chất lượng dịch vụ check-in tại quầy thủ tục Vietnam Airlines ở Nhà ga Quốc nội Sân bay Quốc tế Tân Sơn Nhất.* Luận văn Thạc sỹ, Đại học Kinh tế Thành phố Hồ Chí Minh, Việt Nam.

Wang, T.-C., & Pham, Y. T. (2020). An application of Cluster Analysis Method to determine Vietnam Airlines' ground handling service quality benchmarks. *Journal of Advanced Transportation,* 2020, 1–13. DOI: https://doi.org/10.1155/2020/4156298.