Evaluating Airline Passengers' Satisfaction during the COVID-19 Pandemic: A Case Study of AirAsia Services through Sentiment Analysis and Topic Modelling

Lee Jie Yu[1], Nor Hasliza Md Saad[2]*, Zhu Kun[3], and Ghada ElSayad[4]
[1]Business Analytics, School of Management, Universiti Sains Malaysia
Penang, Malaysia 11800
[2]Operations Management & Business Analytics Sections, School of Management, Universiti Sains Malaysia
Penang, Malaysia 11800
[3]Knowledge Engineering, School of Computer Sciences, Universiti Sains Malaysia
Penang, Malaysia 11800
[4]Business Information Systems Department, College of Management and Technology, Arab Academy for Science, Technology and Maritime Transport (AAST)
Cairo, Egypt 4470351
Email: [1]ljyjieyu112032@gmail.com, [2]norhasliza@usm.my, [3]zhu.kun@student.usm.my, [4]ghadaelsayad@aast.edu

***Abstract*—AirAsia has emerged as a dominant force among prominent low-cost airlines in recent years. However, the COVID-19 pandemic outbreak has severely impacted airline services, including AirAsia. There is a strong need for airline services to monitor customer experience and satisfaction from online customer reviews on the website to keep pace with changing customer perceptions toward their service quality. A growing number of travelers choose to express their experiences and emotions on online customer review platforms, resulting in substantial online airline service evaluations. The research analyzes 796 online customer reviews from Skytrax, a well-known online airline review website. The information hidden in customer-generated reviews is analyzed with the text mining technique, including topic modeling and sentiment analysis. The research uses the Latent Dirichlet Allocation (LDA) model for topic analysis and the Valence Aware Dictionary for Sentiment Reasoning (VADER) model for sentiment analysis. The sentiment ratio for AirAsia's online reviews is approximately 59% positive and 41% negative. Only four reviews are neutral. The findings indicate that the online review of AirAsia has a greater proportion of positive sentiments than negative sentiments. In addition, the topic modeling shows hidden topics with the top high-probability keywords concerned with interior and seat, baggage, online service, staff service, flight schedule, and refund. The research demonstrates using sentiment analysis and topic modeling on customer review data as a more thorough alternative to survey-based models for researching airline service. The research contributes to the methodological advancements in text mining analysis and expands the current knowledge of customer review data.**

***Index Terms*—Airline Service, Customer Satisfaction, COVID-19 Pandemic, Sentiment Analysis, Topic Modelling**

Received: March 04, 2024; received in revised form: May 22, 2024; accepted: May 22, 2024; available online: Aug. 21, 2024.
*Corresponding Author

I. INTRODUCTION

THE aviation industry is pivotal in economic development. It provides a fast and efficient transportation network that facilitates global business, creates employment opportunities, supports the supply chain, and drives international trade. Moreover, it significantly impacts related industries such as tourism and hospitality. In 2018, the aviation industry contributed 3.5% to Malaysia's GDP [1]. The industry is divided into Full-Service Carriers (FSC) and Low-Cost Carriers (LCC). Malaysia Airlines (MAS) and Malindo Air operate as FSCs, while AirAsia, Firefly, and Maswings are LCCs. In 2018, AirAsia emerged as the dominant player in the Malaysian airline industry, capturing a 50% market share and accounting for 60% of the total seat capacity. In contrast, MAS held a 29% market share [2].

The global enforcement of lockdown measures in response to the COVID-19 pandemic significantly impacted the airline industry. These measures led to widespread flight cancellations, and airlines were required to refund ticket fees to customers. The Malaysian aviation industry suffered losses of RM13 billion in 2020 due to ongoing travel restrictions [3]. It also recorded a significant decrease in passenger numbers. These losses forced airlines to implement cost-cutting measures such as reducing flight frequencies, grounding aircraft, and implementing unpaid leave for employees [4]. Additionally, since the pandemic, airlines have adapted their operations to comply with new health and safety regulations, further increasing their operational costs [5]. The airline industry's recovery from the pandemic's impact is expected to be slow and challenging. Therefore, assessing customer satisfaction is crucial for airlines in the post-pandemic era.

Many airline companies emphasize their service strategy as a way to satisfy existing customers' needs and attract new customers. Providing excellent services is crucial. Service quality directly influences customer satisfaction, and positive experiences can foster long-term loyalty [6]. Customers also expect courteous treatment when resolving issues with poor service quality. Consequently, customer satisfaction is a strategic objective for companies aiming to increase customer retention and profitability. When products or services meet customers' expectations, customers tend to feel satisfied, leading to greater brand loyalty [7]. This loyalty, in turn, increases the company's revenue, primarily through repeat purchases. Therefore, it is essential for companies to consistently deliver superior service quality to enhance customer satisfaction, which significantly affects retention. In addition, satisfied customers are more likely to share their positive experiences with others, both online and within their social networks [8]. Hence, it is imperative that decision-makers for airline companies understand how customers perceive the brand's services. This understanding can help them effectively manage positive and negative customer feedback and engage consumers with their product or service offerings [9].

In today's digital age, customers increasingly use online reviews to share their experiences and express satisfaction with products or services [10–12]. These voluntary reviews help customers to evaluate and decide whether to trust businesses, significantly influencing their purchasing decisions. For instance, travelers can enhance their decision-making by leveraging other passengers' experiences shared on online platforms like Skytrax. These insights enable travelers to make informed choices and enhance their travel experiences. Online reviews are also a valuable source of information for companies, offering insights into product or service characteristics and prevailing market conditions [13]. Thus, online reviews are an essential tool that can help companies to remain competitive and meet customer expectations in a rapidly evolving marketplace.

The research assesses customer satisfaction with AirAsia services and interprets the derived insights into actionable managerial strategies that maintain or improve service quality. The research aims to provide valuable insights that can empower AirAsia to make informed decisions that will ultimately improve customer satisfaction and strengthen its competitive position in the Malaysian airline industry. The research collects unstructured data in the form of online reviews that customers post on Skytrax after using travel services of AirAsia to determine these insights. Next, the research employs sentiment analysis to interpret the text reviews. The research analyzes the frequency of words appearing in online reviews and uses topic modelling to identify factors influencing customer satisfaction and highlight dimensions where AirAsia can improve.

The findings of the research contribute to expanding the existing literature while also providing valuable insights from an organizational and customer perspective. By leveraging text mining analysis techniques, the research helps AirAsia to interpret and respond to customers' feedbacks, strengthening the relationship between the company and its customers. From an organizational perspective, analyzing data from customers' online reviews enables the company to understand and respond to customers' voices, address feedbacks and suggestions, and enhance customer–brand relationships. Improvements in service quality lead to a better customer experience, ultimately fostering satisfaction and loyalty. Thus, the research explores the customer experience and offers valuable data and analysis methods derived from customer reviews.

A. Online Reviews

Online reviews are a form of Internet word-of-mouth that contains valuable information from consumers [14]. Customers increasingly share their experiences and concerns about a service or product on social media platforms [15–17]. Many online forums allow consumers to read and provide reviews. Amazon and Consumer Reports are popular product review websites [18]. Meanwhile, Tripadvisor, Booking, and Expedia are popular websites for traveling reviews [19–21]. Then, Yelp and Zomato are known for their restaurant reviews [22–24].

Online reviews are usually publicly available. Consequently, management and other stakeholders, competitors, and investors can use this information to understand customers' intentions and impel targeted customers to seek a particular service or product [25]. Previous research [26] performs a sentiment analysis of educational data using 66,000 online reviews of massive open online courses. Another research [27] provides valuable insights into the nature of online hate speech in the Indonesian context by using online review data from more than 13,000 online tweets to investigate why certain tweets are labelled as hate speech. Then, by analyzing 10,198 online consumer reviews for 516 DVD products on Amazon, previous reserach [28] finds that lower ratings lead to reduced motivation to read individual reviews. Consumers tend to trust average ratings as a proxy for collective wisdom, so inconsistent ratings reduce the helpfulness of reviews.

B. Sentiment Analysis

Sentiment analysis is a technique that evaluates people's opinions, sentiments, and emotions towards goods, services, organizations, occasions, and themes. Sentiment analysis is also known as opinion mining [29]. The goal is to discover people's opinions in text, which are classified as positive, negative, or neutral [28]. Sentiment analysis functions as a categorization system that seeks to classify people's opinions and dispositions while emphasizing relevant facts. It can be used to quickly extract insights from a considerable amount of text data [30–33]. In sentiment analysis, people's opinions or feelings are analyzed using natural language processing, text analysis, and statistics [34, 35]. Sentiment analysis, as applied in the research, can help airline companies extract customer opinions and feelings from online reviews that are helpful for determining which airline companies provide superior services.

Sentiment analysis can be applied to a wide range of industries, including e-commerce [35], hospitality [36], politics [37], and aviation [38]. For example, previous research [39] applies a hybrid sentiment analysis model to data from Twitter regarding US airlines. The sentiment analysis approach effectively reduces the need for large amounts of labelled data. Furthermore, it reduces labor costs in practical applications, which increases accessibility and efficiency. Another previous research [40] analyzes online reviews from Skytrax, focusing on themes and emotions related to the airline experience. Sentiment analysis shows that 'staff service' has a positive impact on satisfaction, especially when paired with the meal and food themes identified in the theme modelling. In another previous research [41], it measures customer satisfaction in the airline industry through online customer reviews. It analyzes more than 55,000 online customer reviews from 400 airlines and passengers in 170 countries by means of sentiment analysis. Ultimately, it identifies 27 dimensions of satisfaction that are described using 882 adjectives.

C. Topic Modeling

In fields like machine learning and natural language processing, topic modeling is a statistical method used to identify abstract subjects within a text collection [42]. It is helpful for analyzing unstructured data since it can discover latent semantic relationships between documents and words [43]. Latent Dirichlet Allocation (LDA) is a popular topic model method that can provide the topic of each document in the document set in the form of a probability distribution. LDA allows topic clustering or text classification to be carried out in accordance with the topic after extracting their topic by analyzing some documents [44].

In the past several years, topic modelling has been predominantly used to analyse massive papers or collections of comment sentences [45], and the topics that are learnt are primarily directed at the entire product brand. However, the majority of user comments are focused on specific aspects or themes of the product, such as service [42], cost performance [45], or transportation [24]. This finding demonstrates that users are more concerned with specific aspects of the product than with its overall rating. Therefore, topic modelling using LDA can uncover previously unnoticed topic patterns, giving companies a clearer picture of consumer preferences and concerns. This information leads to more effective marketing strategies and service improvement plans.

II. Research Method

Section A presents the data collection method, and section B explains the data exploration process. Then, section C presents the sentiment analysis and topic modeling process, as shown in Fig. 1.

A. Data Collection

Skytrax is a leading international rating system that classifies airlines and airports by the quality of the product and the staff service standards. The research collects data from the Skytrax review website (https://www.airlinequality.com/review-pages/), a consumer forum in which passengers can voice their experiences or feelings. The research uses 796 reviews for AirAsia, collected from August 2010 to June 2022. The data from Skytrax's AirAsia airline review website are

Cite this article as: L. J. Yu, N. H. Md Saad, Z. Kun, and G. ElSayad, "Evaluating Airline Passengers' Satisfaction during the COVID-19 Pandemic: A Case Study of AirAsia Services through Sentiment Analysis and Topic Modelling", CommIT Journal 18(2), 137–150, 2024.

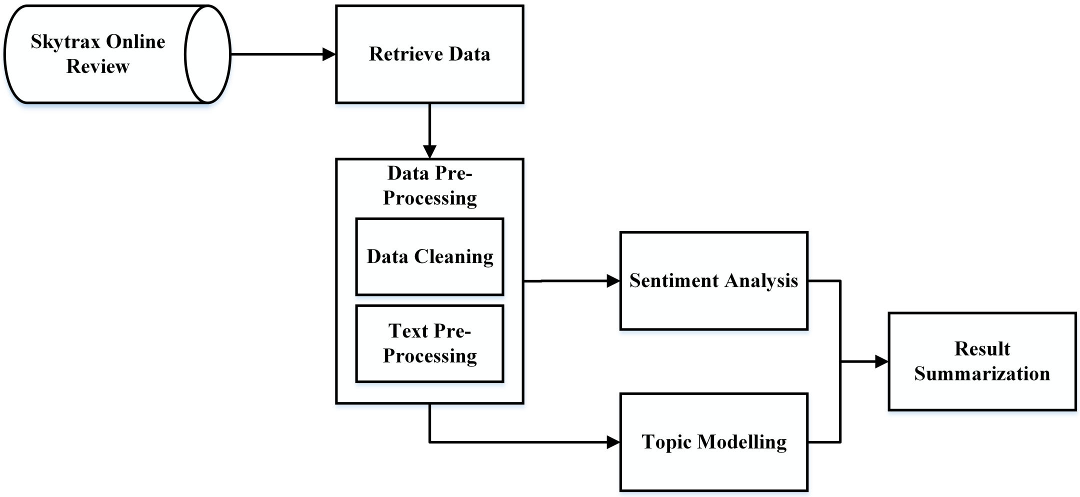

Fig. 1. Process of sentiment analysis and topic modeling.

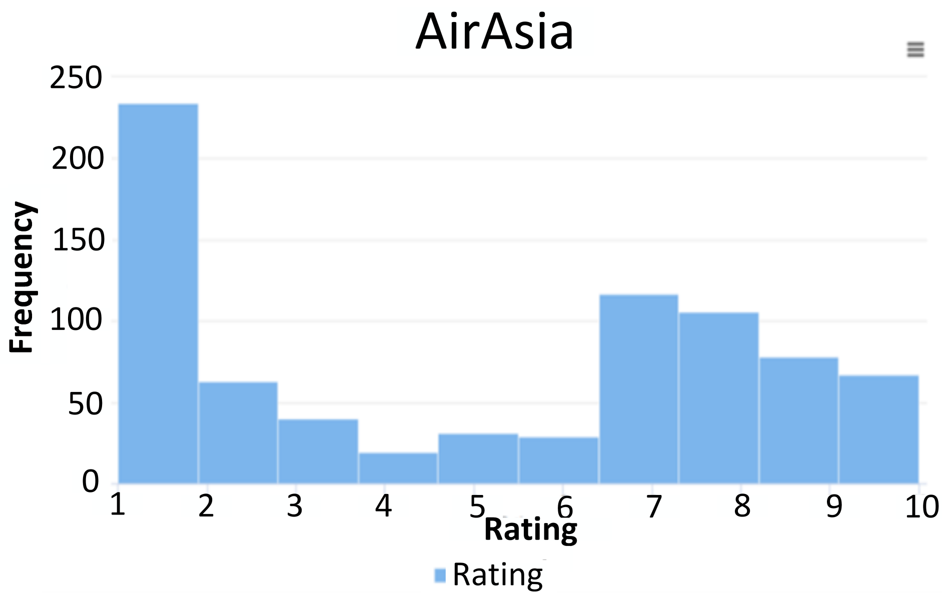

Fig. 2. Histogram of ratings.

analyzed to identify the factors that lead to customer satisfaction. According to the ratings histogram, 234 customers give a 1 out of 10 rating, meaning that most customers are unsatisfied with the AirAsia services. However, some customers give a highly positive review – 67 gave a rating of 10, and 78 give a rating of 9, as shown in Fig. 2. Then, the sample reviews are shown in Table I.

The bar chart in Fig. 3 shows the customers' various countries of origin. From these customers, 141 come from Malaysia, followed by Australia, which has 130 respondents. The third highest country is the United Kingdom with 74 respondents.

B. *Data Pre-processing*

Data pre-processing is a crucial phase in data mining that converts raw data into a valuable and practical format to facilitate or enhance performance [46]. However, the raw data are generally incomplete [47] because data collection techniques frequently lack strict regulation, leading to out-of-range data, the inability to combine data, and missing values [48]. Data that have not been thoroughly checked for these problems may yield misleading results during analysis [49].

TABLE I
SAMPLE OF PASSENGER’S REVIEWS.

Airline Name	Overall Rating	Review Date	Review	Type of Traveller	Seat Type	Route
AirAsia	7	12th April 2024	A flight that left early, unusual and great for travellers. Loading a little cumbersome but mainly from disorganised passengers. Staff affable and helpful. The seats narrow but in good condition	Solo leisure	Economy Class	Penang to Langkawi

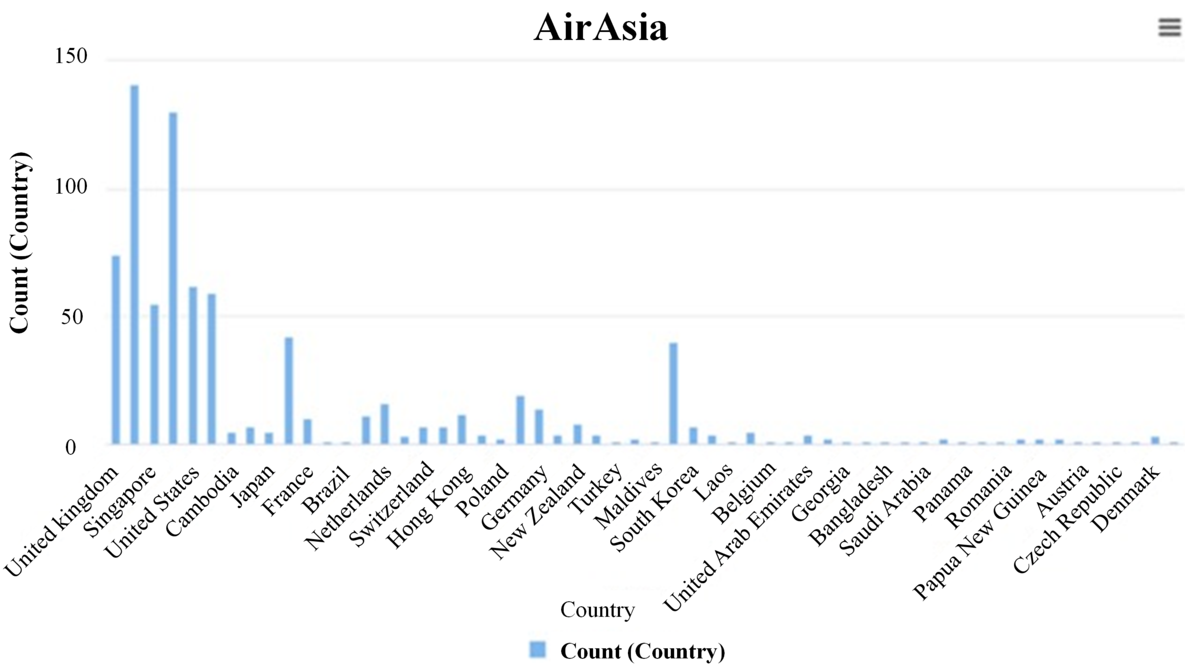

Fig. 3. Bar chart of customers’ countries of origin.

The dataset from Skytrax goes various stages of pre-processing to produce clean and usable data for the research.

Text-based data collected from websites contain a wide range of attributes. These include irrelevant details, such as unnecessary columns, URLs, and missing values [50]. Data cleaning involves removing unwanted data from extracted data, with the goal of preparing data in the appropriate format for analysis. Text pre-processing involves cleaning the text data and transforming it into a form that can feed the model [51]. Text pre-processing aims to extract key phrases or features to increase the association between words and documents or between words and classes [52]. Each document is represented as a feature vector, and the text is split into individual words [53]. Before pre-processing the text, the Nominal-to-Text operator must transform the polynomial data into text [24]. This operator changes the type of selected nominal attributes to text. After that, the data are connected to the Process Document from the Data operator. It is necessary to double-click the Process Document from the Data operator and add the Tokenize operator to split the document’s text into a sequence of tokens. There are numerous options for splitting the points that are displayed in the parameter.

The research uses the default non-letter character, resulting in tokens composed entirely of a single word [42]. The next step is to standardize all words to lowercase. Hence, the Transform Cases operator is applied. Then, the Filter Stopwords (English) operator is used to filter certain words from the document. Conjunctions, prepositions, and pronouns like ‘the’, ‘in’, ‘a’, ‘with’, ‘behind’, ‘she’, and ‘he’, which do not contribute to the document’s meaning and have no bearing on the classification process, are considered

Cite this article as: L. J. Yu, N. H. Md Saad, Z. Kun, and G. ElSayad, “Evaluating Airline Passengers’ Satisfaction during the COVID-19 Pandemic: A Case Study of AirAsia Services through Sentiment Analysis and Topic Modelling”, CommIT Journal 18(2), 137–150, 2024.

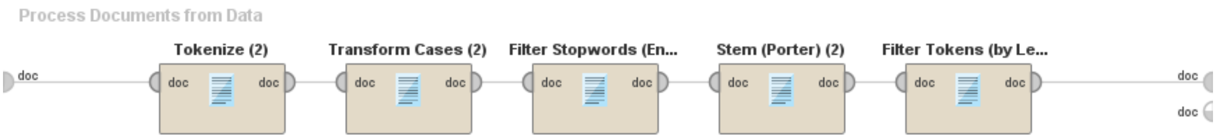

Fig. 4. Text pre-processing process.

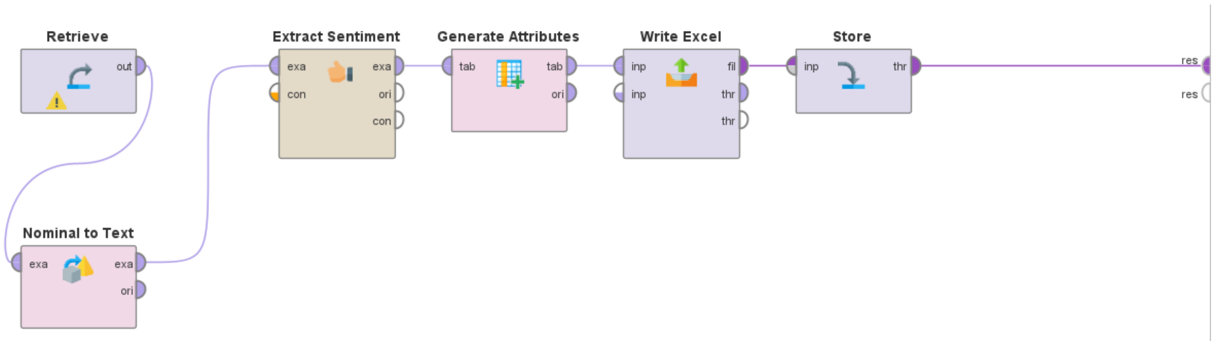

Fig. 5. Sentiment analysis process.

stop words. The next step, stemming, involves removing affixed from features and reducing inflected words to their stem [47]. Other alternatives in RapidMiner can be used, including Snowball, Porters, Lovins, Dictionary, German, and Arabic. The research reduces the length of words until they meet a predetermined minimum length using the Stem (Porter) operator. The last step involves filtering tokens based on length using the Filter Token (by length) operator. The minimum character is set to 3 and the maximum to 25 to avoid considering a single character. Figure 4 depicts the text pre-processing procedures. Tokenization, the initial stage of text pre-processing, involves dividing a sentence into words, which are sometimes referred to as tokens [24].

C. Sentiment Analysis

The Valence Aware Dictionary for Sentiment Reasoning (VADER) is used in sentiment analysis. VADER is a lexicon and rule-based sentiment analysis tool used to score the text [54]. It is particularly sensitive to the sentiments expressed on social media. It relies on a dictionary that maps lexical features to emotion intensities (i.e., the sentiment score). Based on the sentiment score [12], the VADER algorithm outputs the polarity of sentiment into three classes: positive, negative, and neutral. All lexicon ratings have been standardized to a value between -1 and +1. Positive sentiment has a sentiment score above zero, while negative sentiment has a score below zero. Neutral sentiment is represented by a score of zero.

The first step for sentiment analysis involves inputting the clean data from the repository and using the Nominal-to-Text operator to convert the type of the chosen nominal attributes to text [55]. Next, the Extract Sentiment operator determines the sentiment for each text and calculates its polarity. There are some options to expose additional results depending on the chosen method. However, the model chosen in the research is VADER. Then, the Generate Attribute operator is added to construct new user-defined attributes using mathematical expressions [49]. The sentiment analysis process expression is shown in Fig. 5. It generates one of the three sentiments based on the score.

D. Topic Modeling

In the research, it is necessary to retrieve the clean data from the repository and then change the attribute type to ‘text’ for topic modelling. Next, the Select Attributes operator is used to select the desired attribute. Then, the Set Role operator is added. This operator alters the roles of one or more attributes and allows the users to specify which values the computer are used as labels of the sentences [56]. The next step is to add Extract Topics from Data (LDA). This operator can capture significant inter- and intra-document statistical structures to identify hidden topics in the processed data using the LDA method [57]. Users can set the number of topics desired in the parameters. The research categorized the AirAsia online customer reviews into six topics. From the topics, it is possible to determine what aspects AirAsia can enhance to improve

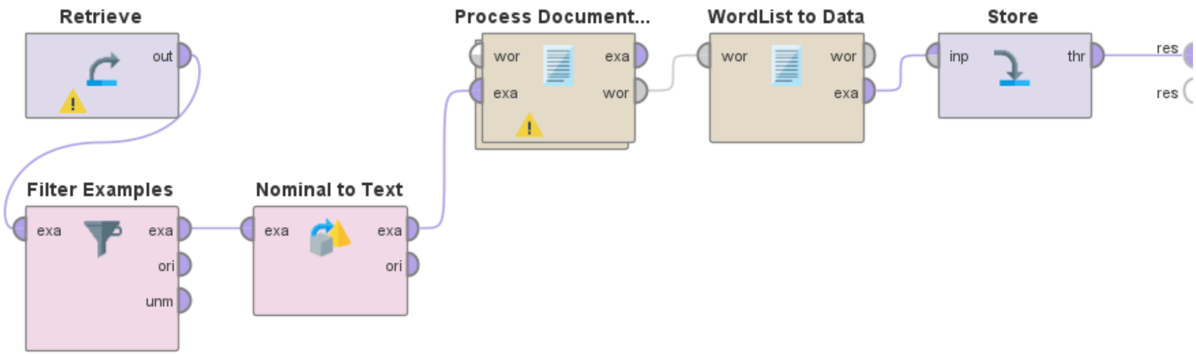

Fig. 6. Topic modelling process.

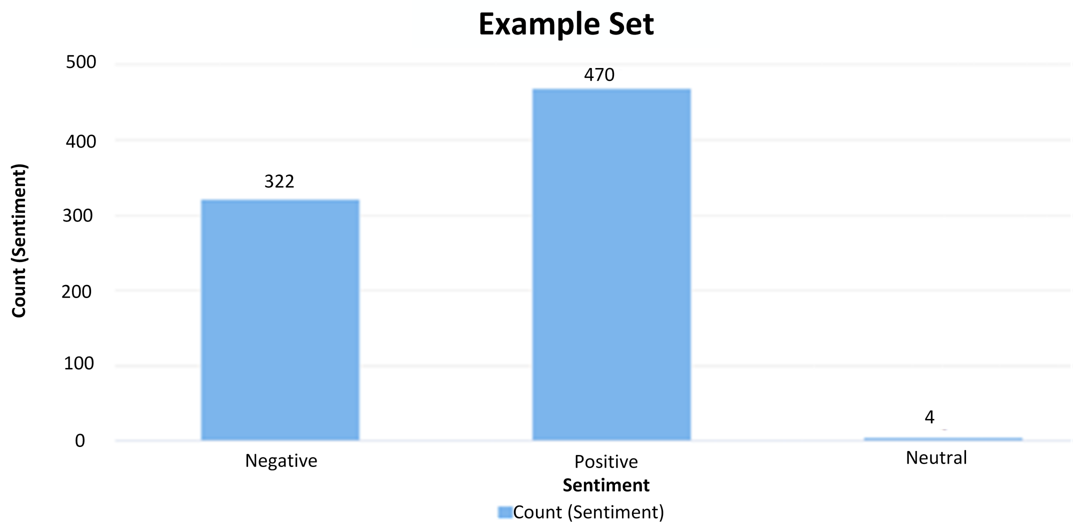

Fig. 7. The result of sentiment analysis.

customer satisfaction with service quality. Figure 6 shows the topic modelling process in RapidMiner.

III. Results and Discussion

A. Sentiment Analysis Results

The VADER tool is used to determine positive, negative, and neutral sentiment in the research. VADER assigns a polarity score to each word in a given social media text and combines these scores to generate an overall sentiment score. VADER's predefined lexicon and rules are used to classify sentiment for each text, making it possible to assess the emotional tone of the text data accurately.

Figure 7 presents the overall sentiment count for the total number of Skytrax online reviews in each category (negative, positive, and neutral) for AirAsia. Then, Table II depicts the sentiment counts of three classes. Out of 796 reviews, 470 are in the positive class, 4 are neutral, and 322 are classified as negative. AirAsia's online review has a higher percentage of positive than negative sentiment. The sentiment ratio for AirAsia's online reviews is approximately 59% positive and 41% negative, with only 4 reviews being neutral.

Table II shows the top 13 most frequently used words for each of the three sentiments. The positive and negative sentiment results yield around 2,000 examples, and the neutral sentiment only has 52 examples. Hence, the most frequent words are sorted. The words 'Kuala' and 'Lumpur' repeatedly appear in these three sentiment lists. This finding indicates that AirAsia is of strong importance in this region.

Next, Fig. 8 demonstrates that words expressing positive emotions, such as 'flight', 'seat', 'service', and 'good', are frequently mentioned in the dataset. The 'good' keyword indicates satisfaction after using

Cite this article as: L. J. Yu, N. H. Md Saad, Z. Kun, and G. ElSayad, “Evaluating Airline Passengers’ Satisfaction during the COVID-19 Pandemic: A Case Study of AirAsia Services through Sentiment Analysis and Topic Modelling”, CommIT Journal 18(2), 137–150, 2024.

TABLE II
SENTIMENT ANALYSIS RESULTS.

Positive Sentiment				Negative Sentiment				Neutral Sentiment			
Row No.	Word	In Documents	Total	Row No.	Word	In Documents	Total	Row No.	Word	In Documents	Total
750	flight	393	915	705	flight	266	699	49	ticket	2	4
1718	seat	206	356	1805	time	153	216	2	airasia	3	3
2043	time	225	330	47	airasia	122	205	3	airlin	2	3
305	check	169	264	295	check	110	203	6	asia	3	3
43	airasia	179	264	856	hour	121	196	24	kuala	2	3
822	good	174	234	51	airlin	115	182	10	cabin	2	2
52	airlin	151	216	487	deal	104	180	11	cancel	1	2
118	asia	132	211	1571	servic	111	167	21	hassl	2	2
207	board	147	202	1546	seat	90	163	25	lumpur	2	2
1154	lumpur	170	199	198	board	85	137	30	monei	1	2
1068	kuala	169	198	460	custom	86	130	35	problem	1	2
1749	servic	148	188	982	kuala	107	124	1	advertis	1	1
1876	staff	139	181	1062	lumpur	106	123	4	amritsar	1	1

Fig. 8. Word cloud for positive sentiment.

AirAsia’s service, and ‘friendly’ indicates positive feelings when the customer can solve their problem with the help of AirAsia staff. This finding confirms that the good attitude of staff positively affects customers in the airline service. Thus, airline companies should focus on training their staff to ensure service quality in the future. Moreover, the words ‘meal’ and ‘food’ appear in the positive sentiment category, indicating that the customers have a positive perception of the airline’s in-flight meal. The word ‘comfort’ also expresses positive feelings regarding the airplane interior, seat, or staff service.

In the negative sentiment category, ‘delay’, ‘time’, and ‘cancel’ are frequently mentioned, excluding ‘flight’. Customers experience negative emotions when there is a service or flight delay. In other words, customers often feel unsatisfied with the delay and cancel the flight, as shown in Fig. 9.

Figure 10 shows the word cloud for neutral sentiment. In this category, ‘ticket’, ‘cabin’, and ‘money’ are often mentioned. This finding suggests that customers value AirAsia ticket and cabin service prices when selecting an airline company.

B. Topic Modeling Result

The topic modeling technique is used to categorize the number of subjects numerous times and select the optimal descriptive subset. To identify which dimension of AirAsia can be improved, the subject is divided into six categories.

Each topic includes 10 keywords, and a topic name is established to describe the keywords contained in each. In Topic 0, ‘Interior and Seat’, which may include terms such as ‘good’, ‘service’, ‘seat’, and ‘staff’. Topic 1 is entitled ‘Baggage’ and includes the keywords of ‘baggage’, ‘boarding’, ‘luggage’, and ‘check-in’. Topic 2 includes the keyword ’Online Services’, including ‘online’, ‘check’, ‘baggage’, and

Cite this article as: L. J. Yu, N. H. Md Saad, Z. Kun, and G. ElSayad, “Evaluating Airline Passengers’ Satisfaction during the COVID-19 Pandemic: A Case Study of AirAsia Services through Sentiment Analysis and Topic Modelling”, CommIT Journal 18(2), 137–150, 2024.

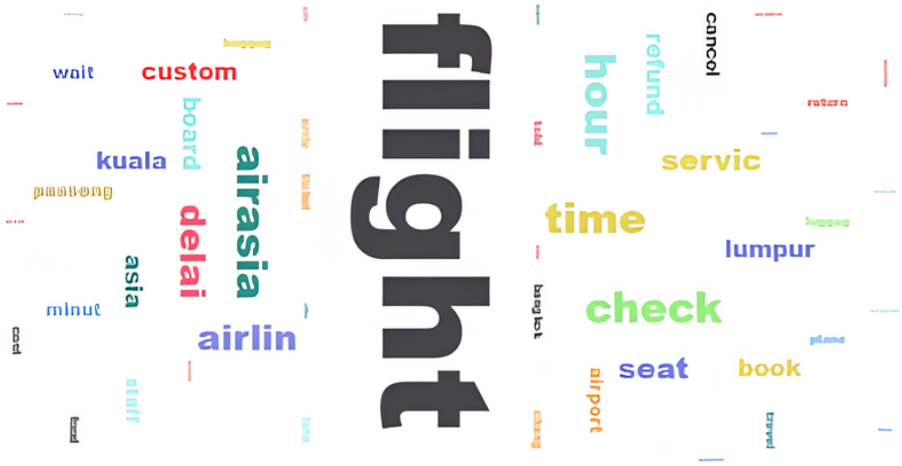

Fig. 9. Word cloud for negative sentiment.

Fig. 10. Word cloud for neutral sentiment.

‘pay’. Topic 3 is ‘Staff Service’ and includes the keywords of ‘airline’, ‘service’, and ‘flight’. Topic 4, ‘Flight Schedule’, includeas terms like ‘time’, ‘hours’, ‘minutes’, and ‘delayed’. Topic 5 is ‘Refund’. The keywords include ‘cancelled’, ‘get’, ‘back’, and ‘ticket’.

In Fig. 11, it is notable that Topic 5 has the greatest weight. Meanwhile, Topic 1, ‘Baggage’, has a low total weight. This finding implies that customers are more focused on the flight schedule issue than baggage.

C. Findings

The research derives various keywords related to ‘service’ that connect to different parts of the airline. It includes ‘seat’, ‘time’, ‘check’, ‘board’, ‘food’, ‘cabin’, ‘meal’, and ‘staff’. In the context of AirAsia, these keywords indicate that many customers regard certain services as essential. Negative sentiment indicates that the customers feel less pleased than expected, which may prevent them from purchasing tickets and using AirAsia in the future.

The keywords ‘time’, ‘delay’, ‘hour’, ‘refund’, and ‘cancel’ are used with negative sentiment. Thus, it can be indicated that customers are particularly concerned about flight time. Most of them have had a bad experience with delayed and cancelled flights. After the Malaysian government reopens the border, the demand for domestic and international flights increases, especially during the holidays because many Malaysians return to their hometown after two years of lockdown. However, many flights are delayed and cancelled during Hari Raya, Labour Day, and Hari Aidilfitri due to the shortage of spare aircraft. AirAsia Malaysia operates with 40 aircraft compared to almost 100 during the pandemic. Several factors influence the lack of flights, including frequent problems and technical difficulties that arise during standard inspections, in

Cite this article as: L. J. Yu, N. H. Md Saad, Z. Kun, and G. ElSayad, "Evaluating Airline Passengers' Satisfaction during the COVID-19 Pandemic: A Case Study of AirAsia Services through Sentiment Analysis and Topic Modelling", CommIT Journal 18(2), 137–150, 2024.

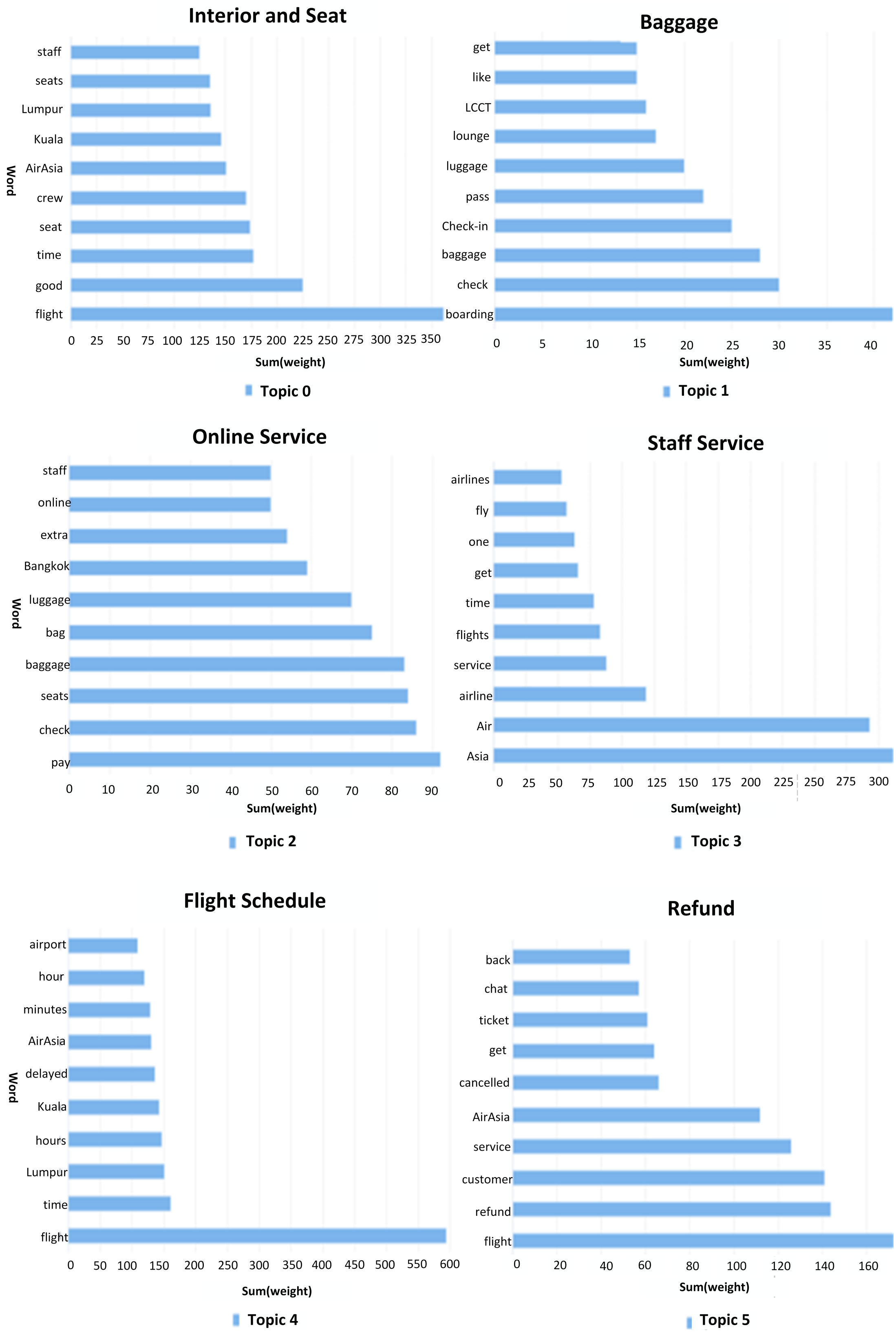

Fig. 11. Output of topic modeling.

addition to Malaysia’s lengthy waiting lists for aircraft repair facilities. Hence, flights have been rescheduled frequently. Some customers claim that AirAsia has retested a flight multiple times, which affects their satisfaction.

Furthermore, the issue of refunds concerns many customers. AirAsia has cancelled or rescheduled many flights due to the COVID-19 pandemic, and customers have requested refunds. However, many customers have claimed that they are never refunded, and some customers have taken a long time to request a refund. Other complaints have been directed to AirAsia Virtual Allstar (AVA), an AI-based chatbot, which is the only way customers can reach customer service. Customers must wait in line in the system due to the large volume of refund issues, and it may take a long time to reach human customer support representatives.

Some customers complain about the challenge of contacting customer support to request a refund. AirAsia offers three types of compensation: a refund, account credit, or a rescheduled flight. The airline encourages the customers to accept the refund in the form of account credit since it happens more quickly than receiving the money back. The lengthy refund process is likely to affect customer satisfaction.

The research findings indicate that AirAsia can improve factors that include the interior, seat, baggage, online service, staff service, flight schedule, and refund process. Currently, AirAsia has transformed itself into a digital airline. They offer various online services, including online ticket purchase, onboard check-in, online baggage purchase, and support from AVA. Some customers have made negative comments regarding the AVA system because it can only answer simple questions and complicated issues such as refunds and rescheduling flights. Customers report that they cannot solve their problem using AVA and must spend significant time and energy contacting and communicating with it. Therefore, it is necessary to improve the quality and speed of online services.

A quality service system should be established and operated to ensure the quick and efficient delivery of assistance to customers. Moreover, among the services provided by AirAsia, customers express positive perceptions regarding the comfort of the interior and seats, the taste of in-flight meals, and the attitude of staff. These factors may persuade customers to choose AirAsia. Since customers expect comfortable and friendly services, the company should strive to continuously maintain the quality of service provided. Thus, employees should receive service training to maintain consistency in service delivery.

IV. Conclusion

In short, the research identifies factors that affect customer satisfaction and aspects that AirAsia should improve. The research method involves analyzing online reviews posted by customers through sentiment analysis and topic modeling. The airline company can easily interpret the data and strive to improve on the problematic or unfavourable elements to the target market. As a result of negative sentiment, most customers are concerned about delayed flights, cancelled flights, and the refund process. Thus, these issues affect customer satisfaction. The results show six aspects that AirAsia should improve on. Three related to service include online services, staff services, and baggage services. It is clear that customers want the best possible experience when flying with AirAsia.

The research acknowledges some limitations that can influence the accuracy of its findings. The analysis relies on collected text data, which may not fully represent all demographics or feedback types. Additionally, sentiment analysis tools, while valuable, may not perfectly capture the subtleties or context within the text. Finally, the lack of a broader context surrounding the user-generated data can limit how sentiment is interpreted. To address these limitations, future researchers should explore more nuanced sentiment analysis techniques and incorporate additional data sources like demographics or flight information. Furthermore, dynamic topic modeling can reveal evolving trends in customer feedback. The development of sentiment-based recommendation systems and qualitative research methods can provide deeper insights and personalised recommendations to improve customer satisfaction and loyalty for AirAsia.

Even with its limitations, the research provides valuable guidance and suggestions to future researchers to conduct similar studies on related topics. The findings of the research aim to enhance AirAsia’s approach to customer satisfaction and aid the company in fulfilling customers’ needs. Consequently, the research recommends that AirAsia enhance the performance of its flights and services while also increasing customer satisfaction levels.

Author Contribution

Writing—original draft, L. J. Y. and N. H. M. S.; Methodology, L. J. Y., N. H. M. S., and Z. K.; Formal analysis, L. J. Y. and N. H. M. S.; Analysis result review, L. J. Y., N. H. M. S., Z. K., and G. E. All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] IATA, "The importance of air transport to Malaysia," 2018. [Online]. Available: https://www.iata.org/en/iata-repository/publications/economic-reports/malaysia--value-of-aviation/

[2] CAPA, "Malaysia aviation: Growth slows, profits under pressure," 2019. [Online]. Available: https://shorturl.at/tfRvK

[3] S. F. S. Yahya, "The impact of COVID-19 on Malaysia's airline industry," 2020. [Online]. Available: https://shorturl.at/QnT4t

[4] B. Vinod, "The COVID-19 pandemic and airline cash flow," *Journal of Revenue and Pricing Management*, vol. 19, no. 4, pp. 228–229, 2020.

[5] T. H. Hassan and A. E. Salem, "Impact of service quality of low-cost carriers on airline image and consumers' satisfaction and loyalty during the COVID-19 outbreak," *International Journal of Environmental Research and Public Health*, vol. 19, no. 1, pp. 1–16, 2022.

[6] K. Tegar, A. Lestari, and S. W. Pratiwi, "An analysis of airlines customer satisfaction by improving customer service performance," in *Proceedings of the Conference on Global Research on Sustainable Transport (GROST 2017)*. Atlantis Press, 2017, pp. 619–628.

[7] M. Giatsoglou, M. G. Vozalis, K. Diamantaras, A. Vakali, G. Sarigiannidis, and K. Chatzisavvas, "Sentiment analysis leveraging emotions and word embeddings," *Expert Systems with Applications*, vol. 69, pp. 214–224, 2017.

[8] A. K. Kar, "What affects usage satisfaction in mobile payments? Modelling user generated content to develop the "digital service usage satisfaction model"," *Information Systems Frontiers*, vol. 23, no. 5, pp. 1341–1361, 2021.

[9] M. Siering, A. V. Deokar, and C. Janze, "Disentangling consumer recommendations: Explaining and predicting airline recommendations based on online reviews," *Decision Support Systems*, vol. 107, pp. 52–63.

[10] B. A. Sparks, H. E. Perkins, and R. Buckley, "Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior," *Tourism Management*, vol. 39, pp. 1–9.

[11] Q. Ye, R. Law, B. Gu, and W. Chen, "The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings," *Computers in Human Behavior*, vol. 27, no. 2, pp. 634–639, 2011.

[12] K. Zhu and N. H. Samsudin, "Attention-based spatialized word embedding Bi-LSTM model for sentiment analysis," *Pertanika Journal of Science & Technology*, vol. 32, no. 1, pp. 79–98, 2024.

[13] N. Korfiatis, P. Stamolampros, P. Kourouthanassis, and V. Sagiadinos, "Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews," *Expert Systems with Applications*, vol. 116, pp. 472–486, 2019.

[14] M. Zhang, B. Fan, N. Zhang, W. Wang, and W. Fan, "Mining product innovation ideas from online reviews," *Information Processing & Management*, vol. 58, no. 1, 2021.

[15] I. Bueno, R. A. Carrasco, R. Ureña, and E. Herrera-Viedma, "A business context aware decision-making approach for selecting the most appropriate sentiment analysis technique in e-marketing situations," *Information Sciences*, vol. 589, pp. 300–320, 2022.

[16] M. Ferrer-Estévez and R. Chalmeta, "Sustainable customer relationship management," *Marketing Intelligence & Planning*, vol. 41, no. 2, pp. 244–262, 2023.

[17] M. J. Noh and K. T. Lee, "An analysis of the relationship between quality and user acceptance in smartphone apps," *Information Systems and E-Business Management*, vol. 14, pp. 273–291, 2016.

[18] X. Fang and J. Zhan, "Sentiment analysis using product review data," *Journal of Big Data*, vol. 2, pp. 1–14, 2015.

[19] V. Chang, L. Liu, Q. Xu, T. Li, and C. H. Hsu, "An improved model for sentiment analysis on luxury hotel review," *Expert Systems*, vol. 40, no. 2, 2023.

[20] L. A. D. A. Lari, S. Iyanna, and F. Jabeen, "Islamic and Muslim tourism: Service quality and theme parks in the UAE," *Tourism Review*, vol. 75, no. 2, pp. 402–413, 2020.

[21] J. M. Luo, H. Q. Vu, G. Li, and R. Law, "Topic modelling for theme park online reviews: Analysis of Disneyland," *Journal of Travel & Tourism Marketing*, vol. 37, no. 2, pp. 272–285, 2020.

[22] A. Brasoveanu, M. Moodie, and R. Agrawal, "Textual evidence for the perfunctoriness of independent medical reviews," in *KiML*, 2020, pp. 1–9.

[23] Y. Cheng, L. Yao, G. Xiang, G. Zhang, T. Tang, and L. Zhong, "Text sentiment orientation analysis based on multi-channel CNN and bidirectional GRU with attention mechanism," *IEEE Access*, vol. 8, pp. 134 964–134 975, 2020.

[24] F. M. Khan, S. A. Khan, K. Shamim, Y. Gupta,

and S. I. Sherwani, "Analysing customers' reviews and ratings for online food deliveries: A text mining approach," *International Journal of Consumer Studies*, vol. 47, no. 3, pp. 953–976, 2023.

[25] A. Bilal, H. T. Mirza, and I. Hussain, "Identifying significant textual features in titles of Google play store applications and their influence on user review rating," *Knowledge and Information Systems*, vol. 65, no. 3, pp. 1159–1178, 2023.

[26] A. Onan, "Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach," *Computer Applications in Engineering Education*, vol. 29, no. 3, pp. 572–589, 2021.

[27] M. A. Ibrahim, S. Arifin, I. G. A. A. Yudistira, R. Nariswari, A. A. Abdillah, N. P. Murnaka, and P. W. Prasetyo, "An explainable AI model for hate speech detection on Indonesian Twitter," *CommIT (Communication and Information Technology) Journal*, vol. 16, no. 2, pp. 175–182, 2022.

[28] S. Lee, S. Lee, and H. Baek, "Does the dispersion of online review ratings affect review helpfulness?" *Computers in Human Behavior*, vol. 117, 2020.

[29] B. Dahal, S. A. P. Kumar, and Z. Li, "Topic modeling and sentiment analysis of global climate change tweets," *Social Network Analysis and Mining*, vol. 9, pp. 1–20, 2019.

[30] I. Jun, Z. Feng, R. Avanasi, R. A. Brain, M. Prosperi, and J. Bian, "Evaluating the perceptions of pesticide use, safety, and regulation and identifying common pesticide-related topics on Twitter," *Integrated Environmental Assessment and Management*, vol. 19, no. 6, pp. 1581–1599, 2023.

[31] A. K. Juswil, S. Nofrima, and H. A. Saputra, "Political communication patterns through social media: A case of an Indonesian presidential staff Twitter account," *CommIT (Communication and Information Technology) Journal*, vol. 16, no. 2, pp. 203–212, 2022.

[32] N. H. M. Saad, A. S. Z. Abidi, Z. Yaacob, M. M. M. Ali, and Z. Kun, "Twitter sentiment analysis on Meta: A lexicon-based analysis using Rapidminer ," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 6, pp. 259–270, 2023.

[33] A. N. Sutranggono and E. M. Imah, "Tweets emotions analysis of community activities restriction as COVID-19 policy in Indonesia using support vector machine," *CommIT (Communication and Information Technology) Journal*, vol. 17, no. 1, pp. 13–25, 2023.

[34] S. Zheng, J. Wang, C. Sun, X. Zhang, and M. E. Kahn, "Air pollution lowers Chinese urbanites' expressed happiness on social media," *Nature Human Behaviour*, vol. 3, no. 3, pp. 237–243, 2019.

[35] X. Xu, B. Li, Y. Shen, B. Luo, C. Zhang, and F. Hao, "Short text classification based on hierarchical heterogeneous graph and LDA fusion," *Electronics*, vol. 12, no. 12, pp. 1–14, 2023.

[36] Y. Wen, Y. Liang, and X. Zhu, "Sentiment analysis of hotel online reviews using the BERT model and ERNIE model–Data from China," *PLoS ONE*, vol. 18, no. 3, pp. 1–14, 2023.

[37] D. T. Hermanto, A. Setyanto, and E. T. Luthfi, "Algoritma LSTM-CNN untuk Binary klasifikasi dengan Word2vec pada media online," *Creative Information Technology Journal*, vol. 8, no. 1, pp. 64–77, 2021.

[38] A. Knorr, "Big data, customer relationship and revenue management in the airline industry: What future role for frequent flyer programs?" *Review of Integrative Business and Economics Research*, vol. 8, no. 2, pp. 38–51, 2019.

[39] M. Chiny, M. Chihab, O. Bencharef, and Y. Chihab, "LSTM, VADER and TF-IDF based hybrid sentiment analysis model," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, pp. 265–275, 2021.

[40] H. J. Kwon, H. J. Ban, J. K. Jun, and H. S. Kim, "Topic modeling and sentiment analysis of online review for airlines," *Information*, vol. 12, no. 2, pp. 1–14, 2021.

[41] F. R. Lucini, L. M. Tonetto, F. S. Fogliatto, and M. J. Anzanello, "Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews," *Journal of Air Transport Management*, vol. 83, 2019.

[42] I. A. Aditya, F. N. Haryadi, I. Haryani, I. Rachmawati, D. P. Ramadhani, T. Tantra, and A. Alamsyah, "Understanding service quality concerns from public discourse in Indonesia state electric company," *Heliyon*, vol. 9, no. 8, pp. 1–14, 2023.

[43] S. Q. Ong, M. B. M. Pauzi, and K. H. Gan, "Text mining and determinants of sentiments towards the COVID-19 vaccine booster of Twitter users in Malaysia," *Healthcare*, vol. 10, no. 6, pp. 1–12, 2022.

[44] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[45] B. Sun, H. Mao, and C. Yin, "How to identify

product defects and segment consumer groups on an online auto forum," *International Journal of Consumer Studies*, vol. 46, no. 6, pp. 2270–2287, 2022.

[46] R. Hubbard, "P-values," in *International encyclopedia of statistical science*. Springer, 2011, pp. 1144–1145.

[47] A. I. Kadhim, "An evaluation of preprocessing techniques for text classification," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 16, no. 6, pp. 22–32, 2018.

[48] A. Ghermandi and M. Sinclair, "Passive crowdsourcing of social media in environmental research: A systematic map," *Global Environmental Change*, vol. 55, pp. 36–47, 2019.

[49] S. Mishra and D. Collier, "A framework for generating annotated social media corpora with demographics, stance, civility, and topicality," 2020. [Online]. Available: http://arxiv.org/abs/2012.05444

[50] C. Singla, F. N. Al-Wesabi, Y. S. Pathania, B. S. Alfurhood, A. M. Hilal, M. Rizwanullah, M. A. Hamza, and M. Mahzari, "An optimized deep learning model for emotion classification in Tweets," *Computers, Materials & Continua*, vol. 70, no. 3, pp. 6365–6380, 2022.

[51] E. Haddi, X. Liu, and Y. Shi, "The role of text pre-processing in sentiment analysis," *Procedia Computer Science*, vol. 17, pp. 26–32, 2013.

[52] J. Dehler-Holland, M. Okoh, and D. Keles, "Assessing technology legitimacy with topic models and sentiment analysis–The case of wind power in Germany," *Technological Forecasting and Social Change*, vol. 175, pp. 1–21, 2022.

[53] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp, "More than a feeling: Accuracy and application of sentiment analysis," *International Journal of Research in Marketing*, vol. 40, no. 1, pp. 75–87, 2023.

[54] C. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2014, pp. 216–225.

[55] A. I. Almulhim, "Understanding public awareness and attitudes toward renewable energy resources in Saudi Arabia," *Renewable Energy*, vol. 192, pp. 572–582, 2022.

[56] S. K. Trivedi, S. Vishnu, A. Singh, and M. Yadav, "Research trends in sustainable e-payment systems: A study using topic modeling approach," *IEEE Transactions on Engineering Management*, vol. 71, pp. 7511–7525, 2023.

[57] X. Wei and V. Taecharungroj, "How to improve learning experience in MOOCs an analysis of online reviews of business courses on Coursera," *The International Journal of Management Education*, vol. 20, no. 3, 2022.