Print ISSN: 2288-4637 / Online ISSN 2288-4645
doi:10.13106/jafeb.2022.vol9.no10.0189

Understanding the Sentiment on Gig Economy: Good or Bad?*

Fatin Aimi Naemah NORAZMI[1], Nur Syazwani MAZLAN[2],

Rusmawati SAID[3], Rahmita Wirza OK RAHMAT[4]

Received: September 15, 2022 Revised: November 26, 2022 Accepted: December 05, 2022

Abstract

The gig economy offers many advantages, such as flexibility, variety, independence, and lower cost. However, there are also safety concerns, lack of regulations, uncertainty, and unsatisfactory services, causing people to voice their opinion on social media. This paper aims to explore the sentiments of consumers concerning gig economy services (Grab, Foodpanda and Airbnb) through the analysis of social media. First, Vader Lexicon was used to classify the comments into positive, negative, and neutral sentiments. Then, the comments were further classified into three machine learning algorithms: Support Vector Machine, Light Gradient Boosted Machine, and Logistic Regression. Results suggested that gig economy services in Malaysia received more positive sentiments (52%) than negative sentiments (19%) and neutral sentiments (29%). Based on the three algorithms used in this research, LGBM has been the best model with the highest accuracy of 85%, while SVM has 84% and LR 82%. The results of this study proved the power of text mining and sentiment analysis in extracting business value and providing insight to businesses. Additionally, it aids gig managers and service providers in understanding clients' sentiments about their goods and services and making necessary adjustments to optimize satisfaction.

Keywords: Industrial Revolution 4.0, Gig Economy, Social Media, Sentiment Analysis

JEL Classifications Code: C88, F60, M10, M31

1. Introduction

The advent of the Industrial Revolution 4.0 (IR4.0) and the Internet of Things (IoT) have brought the world into a new era; the dawn of digital communication that changes how we live. In addition, the pandemic of COVID-19 has hastened the tide of change, forcing individuals, businesses, and governments to adapt to digitalization for their daily needs. This revolution means that individuals can send and receive information faster and easier, thus allowing the gathering of interests and preferences of individuals without the limitation of time and geographical location. At the same time, technological developments have resulted in significant changes in the way we work. For example, the emergence of the platform economy or gig economy has become a significant game-changer. It radically altered traditional employment, such as regular labor contracts and interactions with employers.

According to the MDEC (2022), the gig economy contributed USD2.7 Trillion to the global economy, and the market revenue is projected to be worth USD335 Billion by 2025. Additionally, 700,000 Malaysians have worked in the gig economy, and their earnings are estimated at RM1.3 Billion in income since 2016. Currently, many people choose to be involved in the gig economy because it offers a work-life balance and gives them the flexibility they need because they cannot work for a traditional employer due to personal circumstances. For example, 7 in 10 freelancers in America

*Acknowledgements:
Funding for this project comes from the Fundamental Research Grant Scheme (FRGS/1/2020/SS0/UPM/02/28; Grant No. 05-01-20-2354FR) provided by the Ministry of Higher Education (MOHE) Malaysia.
[1]First Author. Postgraduate Student, School of Business and Economics, Universiti Putra Malaysia, Malaysia. ORCID ID: 0000-0001-6224-6098. Email: aiminaemah@gmail.com
[2]Corresponding Author. Senior Lecturer, School of Business and Economics, Universiti Putra Malaysia, Malaysia. ORCID ID: 0000-0002-9493-0131. [Postal Address: 43400 UPM, Serdang Selangor, Malaysia] Email: nur.syazwani@upm.edu.my
[3]Professor, School of Business and Economics, Universiti Putra Malaysia, Malaysia. ORCID ID: 0000-0002-5150-7843.
Email: rusmawati@upm.edu.my
[4]Professor, Faculty of Computer Science and Technology, Universiti Putra Malaysia, Malaysia., ORCID ID: 0000-0002-9168-245.
Email: rahmita@fsktm.upm.edu.my

say that freelancing gives them the flexibility they need. Around 77% of freelancers agree that they have a better work-life balance than traditionally working (Upwork & Freelancers Union, 2019). Gig work is commonly found in location-based accommodation, transportation, and delivery services. Many gig companies have emerged recently to provide gig work to the population in Malaysia. Examples of e-hailing services include Grab, My Car, Ezcab, and Socar, delivery services such as Food Panda, Grab Food, BungkusIt, and Lalamove, and accommodation services Airbnb, Agoda, and Booking.com.

Along with the growth of the gig economy, substantial issues such as a disparity in supply and demand, gig labor protection, and employment standards in the gig economy also arise (Belanche et al., 2021; Black, 2020; Oyer, 2020). As the gig economy challenges labor practices, another negative impact is failing to meet customer expectations, indirectly affecting consumption decisions (Healy et al., 2017). For instance, an accommodation advertised on Airbnb may not meet the safety standards as this is entirely based on trust and ratings. Therefore, this leads to different sentiments expressed online via different Social Networking Sites (SNS). In another example, Grab is the most used by local riders in Malaysia and is currently the dominant player in the e-hailing industry. However, the demand for improved service quality was anticipated due to the rapid increase in the usage of e-hailing services. Therefore, users on social media often address issues such as price increases, unsatisfactory services, and passenger protection (Yan Chi et al., 2020). Consequently, customers give positive and negative feedback and express their views on social media, which helps other people buy or not buy that particular product or service again based on the social proof of social media ratings and comments. Indirectly, this affects the chance of securing reservations from potential customers resulting in uncertainty and unstable income for the businesses.

Besides, most research on the gig economy only focuses on certain industries, such as e-hailing industries (Adam et al., 2020; Ubaidillah et al., 2019; Yan Chi et al., 2020), e-commerce (Al-Adwan, 2019; Tripopsakul, 2018), hotel industry (Bae & Han, 2020; Valdivia et al., 2017; Zvarevashe & Olugbara, 2018), etc. Moreover, there is limited research that has been done across industries. In addition, due to the ambiguous legal employment status of people conducting gig work, the majority of extant research focuses on labor conditions in the gig economy. For example, Uber and Deliveroo (Sargeant, 2017) urge unionization and improved working conditions for gig workers (Anwar & Graham, 2021; Borzaga et al., 2019; Graham et al., 2017). Very little research has been made to look from a business perspective and the whole gig economy performance. Thus, a broader view of the gig economy and its incumbents is required.

In this paper, the perception of users toward the gig economy will be investigated. Grab, Foodpanda and Airbnb are among the first pioneer and the biggest gig businesses model that provide location-based services in Malaysia. These three gig companies provide location-based services such as online transportation, food delivery, and accommodation services. This study wants to focus on these three location-based services as these services are close to consumers and the most common services used by the population in everyday life. They also have driven the development and transformation of the gig economy in Malaysia. This paper aims to answer the following questions:

How are social media comments pre-processed and classified into sentiments?

What are the most appropriate classifiers for classifying social media comments into sentiments?

What are the users' perceptions of the gig economy in Malaysia?

The study involves the process of big data scraping from different SNS (YouTube, Reddit, and Twitter), focusing on the main business models (Grab, Foodpanda, Airbnb) of the gig economy in Malaysia. Moving forward, the outcomes of this research will be imperative for the growth of the gig economy, which can potentially lead to a significant amount of increased economic growth worldwide.

2. Literature Review

The spawn of the new economic form, the gig economy, is driven by the rise of the 4th Industrial Revolution (IR4.0). According to Ray (2011), the market volume of the gig economy reached around US$100 billion in 2010. The gig economy is a growing business model that encourages new thinking and perspectives. This new form of the economy offers an enormous and practical economic sense for consumers, environments, communities, and businesses that are sufficiently innovative and forward-thinking (Belk, 2014). However, security and privacy are the most critical issues in the gig economy, alongside a lack of regulations, uncertainty, and an unstable income stream. Thus, relevant stakeholders must build consumers' trust through online platforms. For instance, most decisions are now based on online reviews posted on the SNS, and the trustworthiness of these reviews may be questionable (Chang, 2017).

User satisfaction is the ultimate goal that any service wants to achieve. To reach this goal, users must be included in the development and planning of any service. Obtaining user comments, complaints, and ideas may ensure users' involvement. Decision-makers often employ traditional approaches such as questionnaires and surveys to get consumers' perspectives. Time, effort, and staff are all consumed by these conventional procedures. This approach has been more accessible in recent years as social media has evolved (Pournarakis et al., 2016). People express their thoughts and opinions on social media sites such as Facebook, Twitter, Instagram, YouTube, Reddit, etc. Decision-makers can obtain users' input by scraping data from various social media networks and mining their opinions. This method of mining public opinion is known as sentiment analysis.

The sentiments of the online community are vital for the growth and sustainability of the gig economy. These public perceptions can be predicted via sentiment analysis, where significant behavioral influences of sentiments can help companies and government agencies formulate strategies and policies (Collins et al., 2013). About 80% of the data in an organization is estimated to be text-based. Still, it is impractical to manually process huge amounts of textual data to extract meanings and sentiments from them (Gentzkow et al., 2019). Therefore, sentiment analysis is a type of Big Data analytics of textual data or text analytics used to extract sentiments and patterns from unstructured textual data. Sentiment analysis has been used in multiple research areas. For example, it is used to get users' opinions on certain catastrophic events such as natural disasters (Ragini et al., 2018), extreme political viewpoints (Abdullah & Hadzikadic, 2017), terrorism (Mirani & Sasi, 2016), and others. In e-commerce, it is used to get customers' product reviews (Fang & Zhan, 2015) and to use shopping apps (Kim & Yoo, 2021). In healthcare, sentiment is used to analyze users' opinions on the COVID-19 outbreak (Chakraborty et al., 2020). In the financial sector, sentiment is used to get customers' perspectives on financial companies (Park & Javed, 2020) and predict the stock market (Nguyen & Pham, 2018). In transportation, it is used to investigate users' perceptions of public transportation (Fen et al., 2020).

In the gig economy context, several studies have performed sentiment analysis; for example, in the online transportation industry, Baj-Rogowska (2017) performed sentiment analysis on Uber using data gathered from Facebook. The results showed that Uber had higher negative sentiments compared to positive sentiments. The results are consistent with Saragih and Girsang (2018) and Anastasia & Budi (2017). Meanwhile, in the food purchase and delivery services, Trivedi & Singh (2021) conducted a Twitter-based sentiment analysis of three companies (Zomato, Swiggy, and UberEATS). Zomato has received the highest positive sentiments (26%) compared to the other two companies, Swiggy, and UberEATS.

On the other hand, very little research has been done on social media-based sentiment analysis for accommodation services. One previous study has been found that performs sentiment analysis by using the data that has been scrapped from TripAdvisor webpages (Valdivia et al., 2017). Most past research still uses surveys and questionnaires to obtain consumers' perceptions of accommodation services (Agapitou et al., 2020; Wang et al., 2016; Zvarevashe & Olugbara, 2018). In this generation of social networking, there is a massive influx of source materials (termed Big Data) on platforms like Twitter, Facebook, LinkedIn, YouTube, and many other popular SNS (Fang & Zhan, 2015). As mentioned earlier, global social network users have grown exponentially and surpassed 2 billion users in 2016, who are spending an average of more than 2 hours a day. In Malaysia, more than half the population are active users of SNS, with an average of 3 hours daily on social networking, one of the highest in the region (Hootsuite, 2020). This phenomenon is new to us, and much fundamental research is needed to understand the behavioral patterns of the public through the analysis of these big data, especially in the context of the gig economy.

As evident from the studies, public perception has significantly impacted the outcomes in these various fields. A change in consumers' perceptions can lead to the betterment of the stakeholders (firms, consumers, governments) in many aspects (social, economic, environmental). To our best knowledge, no or limited research has been conducted to explore netizens' sentiments from the SNS's big data analytics that can guide the respective bodies in devising both betterment and remedial strategies for the gig economy in Malaysia. The technique employed to understand the behavioral patterns of the public are relatively new in the context of the gig economy and serve as the main contribution of this research. This study aims at filling the existing research gap in an emerging potential market.

Sentiment analysis is a type of Big Data analytics used to extract sentiments and patterns from unstructured textual data. It can be defined as the "identification, extraction, and processing of subjective information from the source material, including opinions, attitudes, and emotions" (Medhat et al., 2014). Anvar Shathik and Krishna Prasad (2020) gave an overview of sentiment analysis methodologies. They include lexicon-based approaches, supervised and unsupervised machine learning approaches, and hybrid approaches (a combination of lexicon and machine learning).

Mandloi and Patel (2020) analyzed participants' sentiments from Twitter data by using supervised machine learning approaches; the Naive Bayes Classifier, the Support Vector Machine, and the Maximum Entropy method and the results indicated that Naive Bayes has the greatest accuracy and may be considered baseline learning methods in the study. However, Vidya et al. (2015) found the best classifier is the Support Vector Machine with the greatest accuracy compared to Naïve Bayes and Decision Tree classifier methods after analyzing consumer sentiment from Twitter data to assess mobile carriers' brands; XL Axiata, Telkom, and Indosat. On the other hand, using the unsupervised machine, Krismawati et al. (2022) analyzed Twitter public sentiment on the economic growth rate during the third wave of the omicron variant of the COVID-19 virus in Indonesia learning, which is a Deep Learning Neural Network method to build the model and the results show that most tweets about economic growth have neutral sentiments.

Veluchamy et al. (2018) presented a comparative research classification model ~~n~~ to compare six different sentiment classification approaches, three supervised machine learning approaches, including SVM, Gradient Boosting, ~~and~~ LR algorithms, and three lexicon-based techniques, including VADER, Pattern, and SentiWordNet lexicons, to analyze Amazon reviews datasets. According to the results, the three supervised machine learning classifiers outperformed the lexicon-based classifiers on all parameters, including accuracy, precision, recall, and F1 score. Furthermore, the VADER lexicon model outperforms the other three lexicon-based models across the board. The LR algorithm is the best overall classifier among the six models, receiving the highest scores. The results obtained are indifferent to the research by Sham and Mohamed (2022), where they performed sentiment analysis to find the most effective sentiment analysis approach for climate change domains.

The VADER lexicon proved its efficiency in analyzing sentiments through the literature. In this study, the VADER lexicon will be used to assign the polarity of the text. The reason for choosing the VADER lexicon is that it is frequently used in the sentiment analysis field, specifically in the social media domain. The hybrid approach proved its efficiency through the literature and performed better in sentiment analysis. Therefore, this study adopted the hybrid approach, the lexicon with supervised machine learning. We will explore the usage of the lexicon with three machine-learning algorithms. The Vader lexicon and the machine learning algorithms Support Vector Machine, Light Gradient Boosted Machine, and Logistic Regression. The aim is to find the best model that can efficiently classify the gig economy text.

3. Methodology

The sentiment analysis process is summarized and presented in Figure 1. The following subsections provide a comprehensive breakdown of the steps. Phyton was used in the whole process.

3.1. Data Acquisition

Data acquisition is carried through the Twitter API, Reddit API, and YouTube API. These APIs allow users to interact with its data, i.e. tweets, posts, and comments. The data collection was carried out for 16 years, from 2006 until 2021. Comments and tweets related to Foodpanda, Airbnb, and Grab are gathered using annotations such as *foodpanda_my, foodpandamalaysia, grab Malaysia, grab_my, Airbnb,* and *airbnbmalaysia.* A total of 18,589 comments were collected. After screening the data, some other languages were found other than English. So, we used the Language detection tool in MALAYA NLTK to detect each language for each comment. Since the composition of Malay, other languages, and formal languages such as ‘Manglish’ and ‘Rojak’ are small compared to English, the translation into the English language was carried out using Google translate API.

3.2. Data Pre-processing

3.2.1. Data Cleaning

Data was cleaned by removing the retweets, duplicate tweets and comments, hyperlinks, mentions, special characters, symbols, and numerals. Lastly, the pre-processing process converted all the text strings to lowercase.

3.2.2. Tokenisation

Tokenisation is breaking down a string into smaller chunks like words, keywords, phrases, symbols, and other tokens. In this research, we tokenized them into words. The tokens are then used as an input following the processes, stemming and lemmatization by using Whitespace Tokeniser with NLTK.

3.2.3. Stemming and Lemmatisation

The stemming stage aims to get the basic word on the tokenized words. It works by cutting off the end of the beginning of the word by considering a list of common prefixes and suffixes that can be found in an inflected word. On the other hand, lemmatization considers the morphological study of words. This requires an algorithm to search through the dictionaries to connect the form to its

Figure 1: The Pipeline of Sentiment Analysis Process

lemma. In this research, we used a dictionary for the English language, Wordnet Lemmatizer, with NLTK.

3.3. Lexicon-Based Approach

Lexicon-Based sentiment analysis was carried out to investigate the sentiment and polarity of tweets by using Valence Aware Dictionary and Sentiment Reasoner (VADER) lexicon. The polarity of the texts was determined by searching the occurrences of the word in the VADER dictionary and then replacing the word position with the polarity value shown by the lexicon dictionary. The polarity of a whole text is calculated by adding the polarities of all the words in that text. VADER lexicon has been discovered to be quite effective compared to other tools. It operates extremely well on social media text. Moreover, it does not require training data since it is built from a generalizable, valence-based, and human-cured gold standard sentiment lexicon (Hutto & Gilbert, 2014). The standard thresholds for classifying sentences as either positive, neutral, or negative have been set below:

Positive sentiment: compound score ≥ 0.05

Neutral sentiment: (compound score > −0.05) and (compound score < 0.05)

Negative sentiment: compound score ≤ −0.05

3.4. Feature Extraction Technique

A feature extraction technique has been adopted in this study: the Term Frequency–Inverse Document Frequency (TF-IDF). The feature extraction process is important for sentiment classification in machine learning as large feature spaces can increase the complexity and reduce the ability to solve tasks (Koncz & Paralic, 2011). TF-IDF is a technique that involves comparing the words in the documents and their relevancy within the comprehensive documents.

3.5. Supervised Machine Learning Algorithm

Support Vector Machine (SVM), Light Gradient Boosted Machine (LGBM), and Logistic Regression (LR) are three supervised machine learning algorithms used in this work. These three machine learning classifiers were chosen because they are among the top most commonly used machine learning classifiers in sentiment analysis. Furthermore, 80% of the dataset was used to train the model, while another 20% became the test dataset used to test the trained model and evaluate the model's performance. Then, the model's performance was calculated using a confusion matrix.

3.6. Performance Metrics

Each model's accuracy, sensitivity, precision, and F1- score were compared using a confusion matrix. As a result, the most effective algorithm for predicting people's opinions about the gig economy was identified. Precision measures correctness, recall measures completeness, F1-score is the harmonic mean of precision, and recall and accuracy are the overall performance of a single classifier (Kandhro et al., 2019).

Equations 1–4 represent the equations of Accuracy, Precision, Recall, and F1-score, respectively, while Table 1 is the Confusion Matrix. Based on Table 1, True Positive (TP) is the situation in which the actual class and the expected class are both labeled as positive, whereas True Negative (TN) is the circumstance in which the actual class and the anticipated class are both labeled as negative. False Positive (FP) refers to the circumstance where negative class data is incorrectly anticipated to be positive class data, while False Negative (FN) refers to the condition where positive class data is incorrectly expected to be negative class data.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + F} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

4. Results and Discussion

This paper used a lexicon-based approach, Vader Lexicon, to assign the sentiment polarity to the social media comments. Figure 2 illustrates the total number of comments for each polarity. The polarity of each comment is assigned to either score ≤ −0.05 (negative), score > −0.05 and <0.05 (neutral), or score ≥ 1 (positive). From the collected data within 16 years, there are 3074 (19%) negative comments, 4661(29%) neutral comments, and 8411(52%) positive comments regarding the gig economy in Malaysia. Notice that the total number of comments is distinct for each polarity. The total number of positive comments is far higher than

Table 1: Confusion Matrix

	Actual Positive	Actual Negative
Predicted Positive	TP	FP
Predicted Negative	FN	TN

neutral and negative comments. The very high percentage of positive sentiments pictured a good and promising future for the gig companies to stay in our economy.

In Figure 3, the sentiment regarding the gig economy starting from 2006 shows a gradual increase in the trend, which means the gig services are there but have not been very popular and well-known by people. However, from 2017 onwards, these gig services began to gain more attention from their consumers. Grab company announced its merger with Uber in Southeast Asia during this period. This deal has driven Grab to become the largest online-to-offline mobile platform in Southeast Asia and a major player in Malaysia's online transportation and food delivery services. In addition, Foodpanda has undergone a rebrand in this period, changing all its branding from neon orange to neon pink, with thousands of pink driver shirts, delivery bags, and store identification stickers distributed.

Furthermore, Foodpanda has updated its app and front end in conjunction with its rebranding to add new features like live order tracking to stand out from its competitors. These new strategies taken by gig companies resulted in more people have used these gig services, and more sentiment has been given online to these gig companies. Moreover, from 2019 onwards, we can see a spike in the sentiments

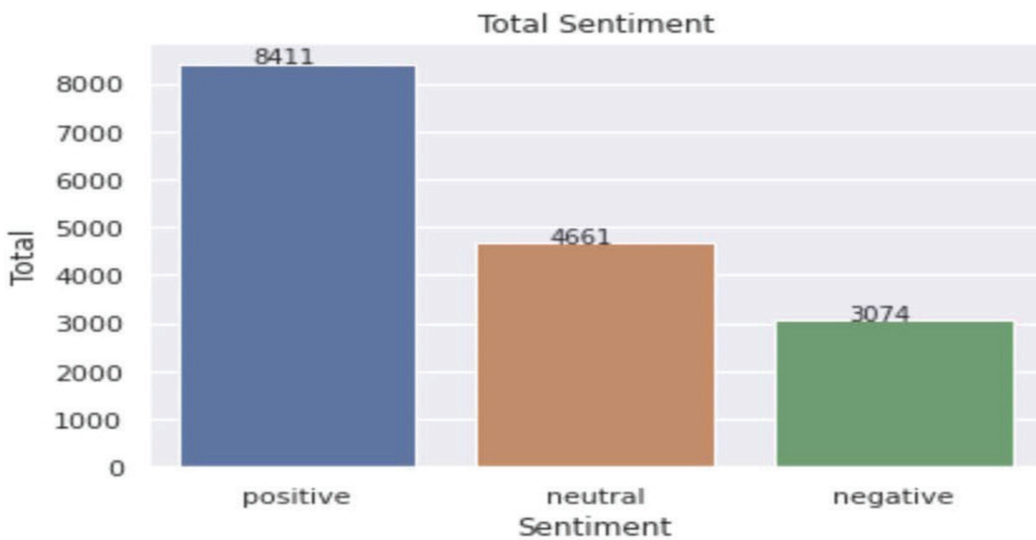

Figure 2: Total Number of Comments for Each Polarity

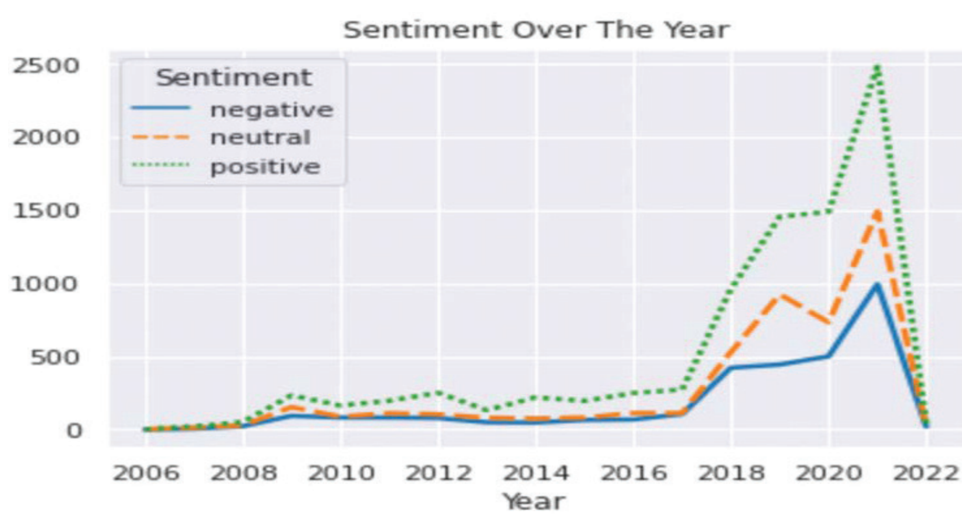

Figure 3: Sentiment Over the Year

shared by social media users. This is because the outbreak of Covid caused most people to use social media as a tool for communication. In addition, the physical barriers and the stay-at-home order imposed by the government forced people to get to know and use gig services. So, it explains the sharp increment in positive, negative, and neutral sentiments towards gig services on social media.

Throughout the analysis, frequencies of words were found; Figure 4 shows the cloud of the words, and the size of the world reflects its frequency. For example, the most frequent words in the collected dataset for positive sentiments were 'thank' and 'good', which convey a positive meaning. On the other hand, the most frequent words that convey a negative sense are 'cancel' and 'pay'. In addition, the most frequent bigrams in the negative and positive polarities were found. Figures 5 and 6 show these frequencies.

Figure 5 shows that the most positive comment is 'God bless', which appears 75 times, followed by time champion 65 times. This indicates that some people think punctuality in gig economy services is awesome. People are thrilled with the punctuality and effectiveness of the service providers, for example, riders that can deliver the food in a short period, drivers that can arrive on time, and hosts that can solve the accommodation problem in a short period. Moreover,

Figure 4: Word Cloud of Positive and Negative Sentiment

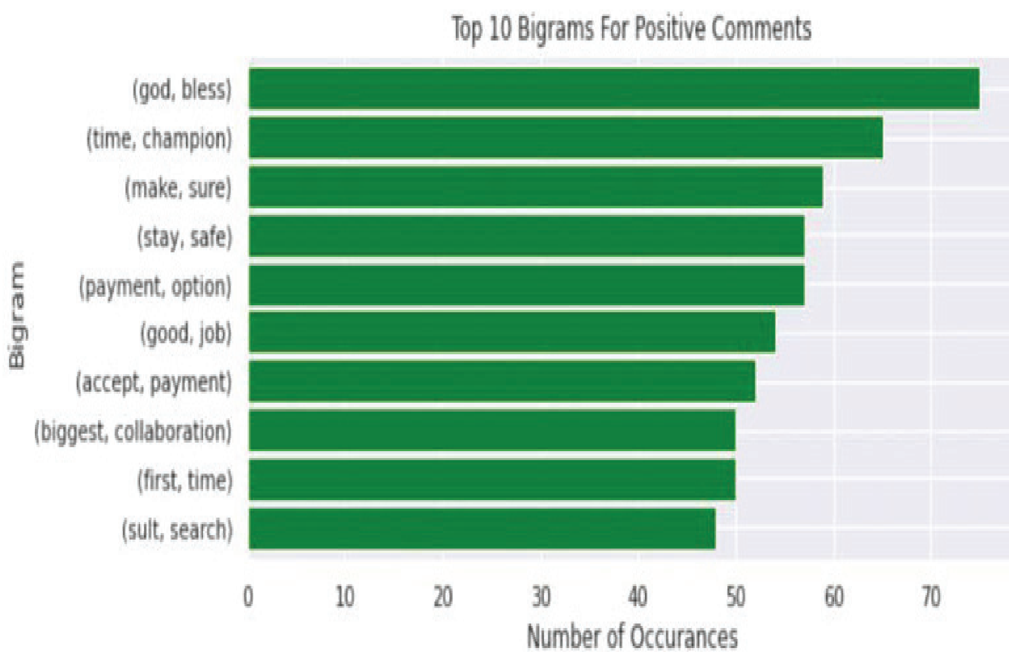

Figure 5: Frequency Plot of Positive Polarity

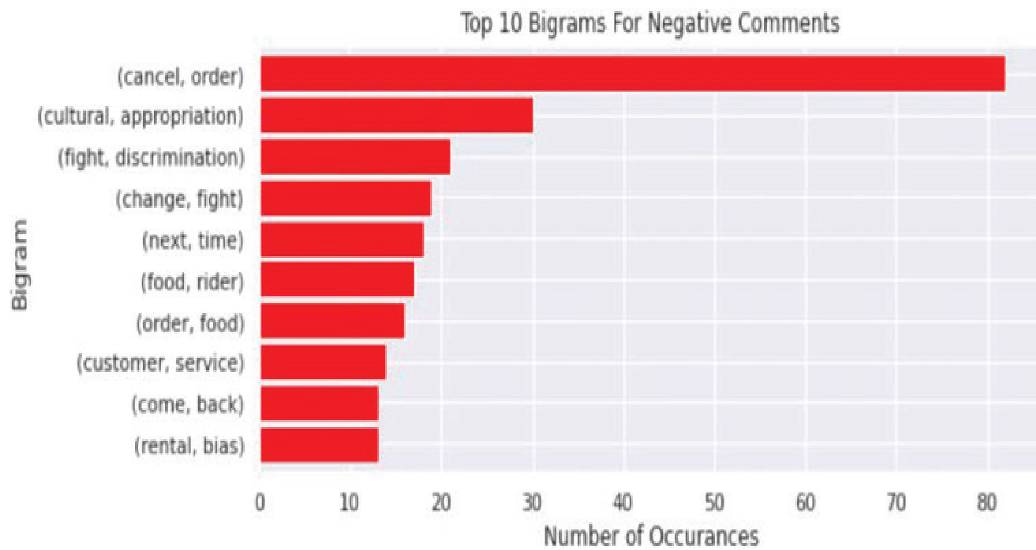

Figure 6: Frequency Plot for Negative Polarity

the helpful and friendly rider, host, or driver is one of the positive perceptions of users towards the gig economy in Malaysia. Furthermore, consumers are attracted more to businesses that offer customers a wide range of payment options, as people talk positively 58 times on social media. Usually, people avoid using gig services because they do not want to become trapped in the payment issue.

Introducing various payment methods has resulted in increased positive perceptions and, as a result, increased business for many companies. Payments can be made in multiple ways, including cash on delivery, online money wallets, debit, and credit cards, and more. On top of it, most social media users have suggested that Airbnb should enhance its payment method by taking cryptocurrency payments. This request should be considered and worth exploring and implementing for gig companies. Besides, the biggest collaboration has appeared 50 times with positive sentiments. This shows that people have more positive perceptions of a company that collaborates with other big or start-up businesses. For example, Foodpanda collaborated with the Federal government as part of the 'E-Commerce' and 'Shop Malaysia Online' campaign to assist the digitalization of micro, small, and medium enterprises (SMEs) and increase consumers' digital spending (MOF, 2020). Consumers are delighted with the collaboration campaign as they will have various product choices on one platform, and more discounts and vouchers will be available. This is also vital to maintain a competitive advantage between the sellers.

Figure 6 shows the negative users' perceptions of the gig economy regarding bad service, culture, discrimination, and biases. The most negative comments are related to 'cancel, order' (83 times). The comments on 'cancel, order' are also associated with comments on 'food, rider' (18 times), 'order, food' (16 times), and 'customer, service' (13 times). Social media is rampant with complaints about order cancellation, such as food riders canceling their order, and the customer service's refund on the canceled order has been ignored. Besides that, comments related to cultural appropriation and discrimination appear 30 times and 23 times in negative sentiments, respectively. Controversial topics like Food Panda's Bollywood-themed advertising video with an all-Malay cast were criticized online for copying Indian culture by not incorporating any artiste from Indian ethnicity. This issue has been the most talked about among social media users, and most have urged the Foodpanda company to remove the video to fight discrimination. In addition, issues related to rental bias have around 12 occurrences of negative sentiments on social media. Social media users complained about the hosts of Airbnb, and some of them were declining the booking based on the customers' ethnicity, race, and religion. Gig companies should be more alert to cultural, racial, and religious issues by updating their strategies to address discrimination and racism. Treating each other in the gig economy with dignity and respect is essential. Further, this paper proposed three models, Support Vector Machine, Logistic Regression, and Light Gradient Boosted Machine, to test the dataset and find the best model representing our data. Then, the performance of each model was assessed.

Based on Figure 7, the Support Vector Machine has a precision of 0.73, recall of 0.74, and F1 score of 0.74. On the other hand, a light Gradient Boosted Machine has a precision of 0.75, an F1 score of 0.70, a recall of 0.72, and Logistic

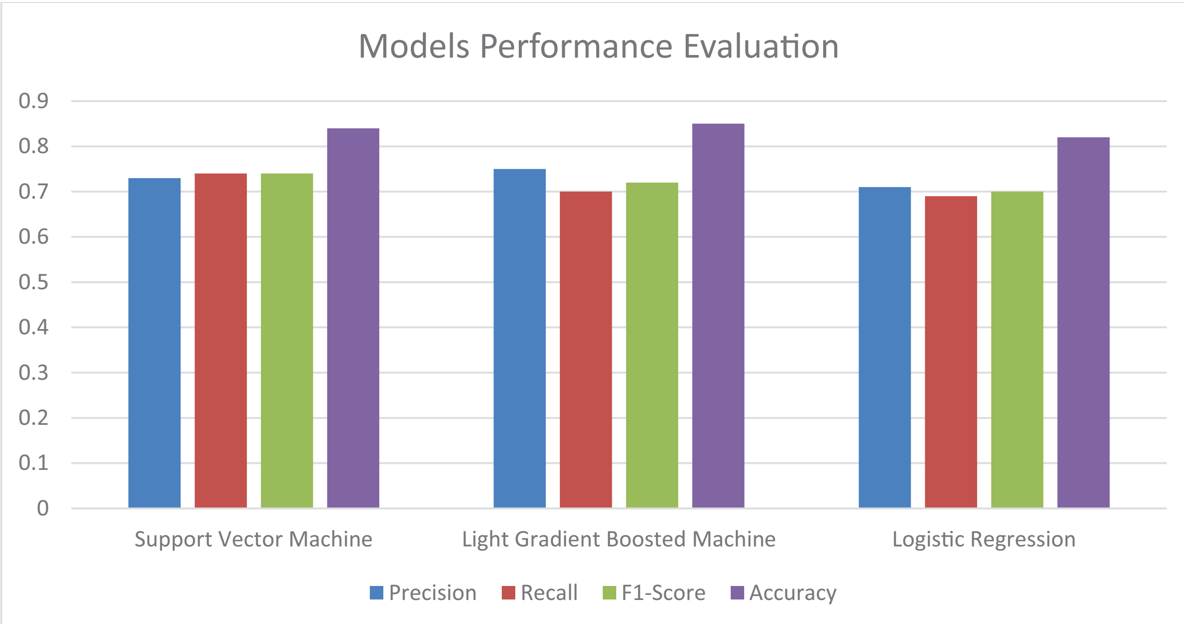

Figure 7: Graph of Models Performance Evaluation

Regression has a recall of 0.69, a precision of 0.71, and an F1 score of 0.70. Looking at these models' performance in Figure 7, LGBM is slightly higher than other models. It can be supported by the results of accuracy for these models where we can see that LGBM has the highest accuracy, which is 0.85 (85% accuracy) compared to other models, Support Vector Machine (84%), and Logistic Regression (82%). Hence, these results can enable further analysis of that algorithm. The value of the F1 score is 0.70, according to LGBM's categorization result. Therefore, 72% of comments can be appropriately classified by our proposed technique. LGBM worked well in dealing with high-dimensional and imbalanced data. Thus, LGBM shows better results with the datasets.

5. Conclusion

This study performs sentiment analysis by applying text mining for user data generated on Twitter, Reddit, and YouTube on the three-leading app-based gig companies in Malaysia, such as Grab, Food panda, and Airbnb. The data was collected from these social media for 16 years and then processed and annotated. The Vader lexicon and supervised learning algorithms (SVM, LGBM, and LR) have been adopted to analyze the sentiments of the users' comments. This study suggested that the gig companies' overall performance was toward users' positive sentiments. Results show the highest, 52% positive comments towards the gig economy, 29% neutral, and 19% negative comments. The frequency of unigrams and bigrams in the positive and negative polarity was found where the negative indicated the consumers' complaints, while the positive ones indicated the satisfaction features. We found that most people show satisfaction with punctuality, payment option, and company collaboration features. In contrast, people complain about the bad service, culture, discrimination, and biasedness features in social media. Furthermore, the best model was LGBM, with the highest accuracy of 85%. This model can serve as the baseline for future research, especially in the gig economy context.

The study offers a fundamental framework for utilizing text data from customer reviews in the context of the gig economy on social networks, laying the groundwork for future studies that will exploit big data in each sector of the gig economy to benefit both businesses and consumers. The study's findings also significantly contribute to the practical use of social network data mining in understanding users' demands and, as a result, informing wise business decisions and enterprise management. Finally, the findings also point authorities and policymakers in the direction of soliciting feedback from the public before drafting and implementing management strategies through social media.

The sectors of online food delivery, online transportation, and room-sharing will all have initiatives to improve their services and products to attract more clients. Additionally, the research will serve as the basis for data analysis tools, incorporating this approach to poll customer experience perceptions of all goods and services.

In the future, a system that can automatically update data should be installed. Before saving to the database, data will be automatically pulled from the website, and duplicate entries will be removed. Increase data collection from many sources and advance research into big data analysis. Businesses will find it more comfortable to examine reports and make wiser decisions when consumer opinion reports are applied to websites, particularly mobile ones.

The limitation of this study is that it primarily focuses on only three key players in the gig economy: Food panda, Grab, and Airbnb. For future research, including more companies in different industries will give more samples, provide a lot more information, and give better-generalized insights into consumer perceptions of the gig economy. We also recommend using more specific keywords to collect data. For example, using the product name instead of the company name, Panda Mart for Food panda and Grab car, Grab delivery, and Grab food for Grab may allow more precise data collection. In the context of the research method, only the VADER lexicon and three algorithms were applied in classifying the social media comments into sentiments. Extending this research with other lexicons and algorithms is the solution that can be applied to find universal models in classifying comments regarding the gig economy into sentiments.

References

Abdullah, M., & Hadzikadic, M. (2017). Sentiment analysis of Twitter data: Emotions revealed regarding Donald Trump during the 2015–16 primary debates. *29th International Conference on Tools with Artificial Intelligence (ICTAI), 2017,* Boston, MA, 6–8 Nov 2017 (pp. 760–764). Piscataway, NJ: IEEE Publications. https://doi.org/10.1109/ICTAI.2017.00120

Adam, M., Kee, D. M. H., Junaina, I., Fadhilah, N., Uwais, N., Al Rashidi, F., Al Shammari, H., Quttainah, M. A., Srivastava, A., & Pandey, R. (2020). The influence of customer satisfaction on grab services in Malaysia. *International Journal of Tourism and Hospitality in Asia Pacific*, *3*(2), 26–37. https://doi.org/10.32535/ijthap.v3i2.820

Agapitou, C., Liana, A., Folinas, D., & Konstantoglou, A. (2020). Airbnb is customers' choice: Empirical findings from a survey. *Sustainability*, *12*(15), 1–13. https://doi.org/10.3390/su12156136

Al-Adwan, A. S. (2019). Revealing the influential factors driving social commerce adoption. *Interdisciplinary Journal of Information, Knowledge, and Management*, *14*, 295–324. https://doi.org/10.28945/4438

Kandhro, I., Ameen Chhajro, M., Kumar, K., Lashari, H. N., & Khan, U. (2019). Student feedback sentiment analysis model using various machine learning schemes A review. *Indian Journal of Science and Technology*, *14*(12), 1–9. https://doi.org/10.17485/ijst/2019/v12i14/143243

Anastasia, S., & Budi, I. (2016). Twitter sentiment analysis of online transportation service providers. *International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2016, Malang, 15–16 Oct 2016* (pp. 359–365). Piscataway, NJ: IEEE Publications https://doi.org/10.1109/ICACSIS.2016.7872807

Anwar, M. A., & Graham, M. (2021). Between a rock and a hard place: Freedom, flexibility, precarity, and vulnerability in the gig economy in Africa. *Competition and Change*, *25*(2), 237–258. https://doi.org/10.1177/1024529420914473

Bae, S. Y., & Han, J. H. (2020). Considering cultural consonance in the trustworthiness of online hotel reviews among generation Y for sustainable tourism: An extended TAM model. *Sustainability*, *12*(7), 2942. https://doi.org/10.3390/su12072942

Baj-Rogowska, A. (2017). Sentiment analysis of Facebook posts: The uber case. *Eighth International Conference on Intelligent Computing and Information Systems (ICICIS), 2017*, Cairo, Egypt, 5–7 December 2017 (pp. 391–395). Piscataway, NJ: IEEE Publications. https://doi.org/10.1109/INTELCIS.2017.8260068

Belanche, D., Casaló, L. V., Flavián, C., & Pérez-Rueda, A. (2021). The role of customers in the gig economy: How perceptions of working conditions and service quality influence the use and recommendation of food delivery services. *Service Business*, *15*(1), 45–75. https://doi.org/10.1007/s11628-020-00432-7

Belk, R. (2014). You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research*, *67*(8), 1595–1600. https://doi.org/10.1016/j.jbusres.2013.10.001

Black, C. M. (2020). The future of work: The gig economy and pressures on the tax system. *Canadian Tax Journal/Revue Fiscale Canadienne*, *68*(1), 69–97. https://doi.org/10.32721/ctj.2020.68.1.sym.black

Borzaga, C., Salvatori, G., & Bodini, R. (2019). Social and solidarity economy and the future of work. *Journal of Entrepreneurship and Innovation in Emerging Economies*, *5*(1), 37–57. https://doi.org/10.1177/2393957518815300

Chakraborty, K., Bhatia, S., Bhattacharyya, S., Platos, J., Bag, R., & Hassanien, A. E. (2020). Sentiment analysis of COVID-19 tweets by deep learning classifiers: A study to show how popularity is affecting accuracy in social media. *Applied Soft Computing*, *97*, 106754. https://doi.org/10.1016/j.asoc.2020.106754

Chang, W. L. (2017). Using sentiment analysis to explore the degree of risk in sharing economy. *IEEE International Conference on Big Data (Big Data), 2017*, Boston, MA, USA, 11–14 December 2017 (pp. 2701–2709). Piscataway, NJ: IEEE Publications. https://doi.org/10.1109/BigData.2017.8258233

Collins, C., Hasan, S., & Ukkusuri, S. V. (2013). A novel transit rider satisfaction metric: Rider sentiments measured from online social media data. *Journal of Public Transportation*, *16*(2), 21–45. https://doi.org/10.5038/2375-0901.16.2.2

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

Fen, C. W., Ismail, M. A., Zayet, T. M. A., & Dewi, K. (2020). Sentiment analysis of users' perception towards public transportation using Twitter. *International Journal of Technology Management and Information System*, *2*(1), 92–101. https://myjms.mohe.gov.my/index.php/ijtmis/article/view/8907/991

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, *57*(3), 535–574. https://doi.org/10.1257/jel.20181020

Graham, M., Hjorth, I., & Lehdonvirta, V. (2017). Digital labor and development: Impacts of global digital labor platforms and the gig economy on worker livelihoods. *Transfer*, *23*(2), 135–162. https://doi.org/10.1177/1024258916687250

Healy, J., Nicholson, D., & Pekarek, A. (2017). Should we take the gig economy seriously? *Labour and Industry: A Journal of the Social and Economic Relations of Work*, *27*(3), 232–248. https://doi.org/10.1080/10301763.2017.1377048

Hootsuite. (2020). *Global digital insights.* https://datareportal.com/reports/digital-2020-global-digital-overview.

Hutto, C. J., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for. *Eighth International AAAI Conference on Weblogs and Social Media, Ann Arbor, Michigan, USA*, June 1–4, 2014 (pp. 216–225). Palo Alto, California, U.S: AAAI Press. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewPaper/8109.

Anvar Shathik J., & Krishna Prasad, K.. (2020). A literature review on the application of sentiment analysis using machine learning techniques. *International Journal of Applied Engineering and Management Letters (IJAEML)*, *4*(2), 41–77. https://doi.org/http. https://doi.org/10.5281/zenodo.3977576

Kim, S. H., & Yoo, B. K. (2021). Topics and sentiment analysis based on reviews of omnichannel retailing. *Journal of Distribution Science*, *19*(4), 25–35. https://doi.org/10.15722/jds.19.4.202104.25

Koncz, P., & Paralic, J. (2011). An approach to feature selection for sentiment analysis. *15th IEEE International Conference on Intelligent Engineering Systems, 2011,* Tyskland, 23–26 June 2011(pp. 357–362). https://doi.org/10.1109/INES.2011.5954773

Krismawati, D., Calvin, W., Mariel, F., Arsyi, F. A., & Pramana, S. (2022). Sentiment analysis on Indonesia's economic growth using deep learning neural network method. *Journal of Asian Finance, Economics, and Business*, *13*(6), 9–18. https://doi.org/http. https://doi.org/10.13106/jidb.2022.vol13.no6.9

Mandloi, L., & Patel, R. (2020). Twitter sentiments analysis using machine learning methods. *International Conference for Emerging Technology, INCET 2020,* Belgaum, India. June 5–7, 2020 · (pp. 1–5). Piscataway, NJ: IEEE. https://doi.org/10.1109/INCET49848.2020.9154183

MDEC. (2022). *Embracing the sharing economy utilizing digital platforms in the sharing of goods, assets, and services sharing economy contributes.* https://mdec.my/sharingeconomy/

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Mirani, T. B., & Sasi, S. (2016). Sentiment analysis of ISIS-related tweets using absolute location. *International Conference on Computational Science and Computational Intelligence*, Las Vegas, NV, USA, 15–17 December 2017 (pp. 1140–1145). Piscataway, NJ: IEEE. https://doi.org/10.1109/CSCI.2016.0216

Ministry of Finance of Malaysia. (MoF) (2020). *National economic regeneration plan together to generate the economy-short-term economic recovery plan*. https://www.mampu.gov.my/wp-content/uploads/2021/04/Teks-Perutusan-YAB-PM-PENJANA-05062020-1.pdf

Sham, N., & Mohamed, A. (2022). Climate change sentiment analysis using lexicon, machine learning, and hybrid approaches. *Sustainability*, *14*(8), 1–28. https://doi.org/10.3390/su14084723

Nguyen, D. D., & Pham, M. C. (2018). Search-based sentiment and stock market reactions: An empirical evidence in Vietnam. *Journal of Asian Finance, Economics, and Business*, *5*(4), 45–56. https://doi.org/10.13106/jafeb.2018.vol5.no4.45

Oyer, P. (2020). The gig economy. *IZA World of Labor, 17*, 471. https://doi.org/10.15185/izawol.471

Park, Y. E., & Javed, Y. (2020). Insights discovery through hidden sentiment in big data: Evidence from Saudi Arabia's financial sector. *Journal of Asian Finance, Economics, and Business*, *7*(6), 457–464. https://doi.org/10.13106/jafeb.2020.vol7.no6.457

Pournarakis, D. E., Sotiropoulos, D. N., & Giaglis, G. M. (2017). A computational model for mining consumer perceptions in social media. *Decision Support Systems*, *93*, 98–110. https://doi.org/10.1016/j.dss.2016.09.018

Ragini, J. R., Anand, P. M. R., & Bhaskar, V. (2018). Big data analytics for disaster response and recovery through sentiment analysis. *International Journal of Information Management*, *42*, 13–24. https://doi.org/10.1016/j.ijinfomgt.2018.05.004

Ray, S. (2011). The emerging trend of e-commerce in India: Some crucial issues, prospects, and challenges. *Computer Engineering and Intelligent Systems, 2*(5), 17–36. https://iiste.org/Journals/index.php/CEIS/article/view/503

Saragih, M. H., & Girsang, A. S. (2017). Sentiment analysis of customer engagement on social media in transport online. *International Conference on Sustainable Information Engineering and Technology (SIET)*, Batu City, Indonesia, November 24–25, 2017. Piscataway, NJ: IEEE. https://doi.org/10.1109/SIET.2017.8304103

Sargeant, M. (2017). The gig economy and the future of work. *Ejournal of International and Comparative Labour Studies*, *6*(2), 2–12. https://eprints.mdx.ac.uk/id/eprint/22247

Tripopsakul, S. (2018). Social media adoption as a business platform: An integrated tam-toe framework. *Polish Journal of Management Studies*, *18*(2), 350–362. https://doi.org/10.17512/pjms.2018.18.2.28

Trivedi, S. K., & Singh, A. (2021). Twitter sentiment analysis of app-based online food delivery companies. *Global Knowledge, Memory, and Communication*, *70*(8/9), 891–910. https://doi.org/10.1108/GKMC-04-2020-0056

Ubaidillah, N. Z., Yi, C. Y., Hassan, M. K. H., Ali, S. S. S., & Hwang, J. Y. T. (2019). The determinants of generation Z's intention to use the grab E-hailing services. *International Journal of Academic Research in Business and Social Sciences*, *9*(11), 483–495. https://doi.org/10.6007/IJARBSS/v9-i11/6570

Upwork & Freelancers Union. (2019). *Freelancing in America.* https://content-static.upwork.com/blog/uploads/sites/11/2020/04/06165223/Freelancing-in-America.pdf

Valdivia, A., Luzón, M. V., & Herrera, F. (2017). Sentiment analysis in TripAdvisor. *IEEE Intelligent Systems*, *32*(4), 72–77. https://doi.org/10.1109/MIS.2017.3121555

Veluchamy, A., Nguyen, H., Diop, M. L., & Iqbal, R. (2018). Comparative study of sentiment analysis with product reviews using machine Learning and lexicon-based approaches. *SMU Data Science Review*, *1*(4), 1–22.

Vidya, N. A., Fanany, M. I., & Budi, I. (2015). Twitter sentiment to analyze the net brand reputation of mobile phone providers. *Procedia Computer Science*, *72*, 519–526. https://doi.org/10.1016/j.procs.2015.12.159

Wang, Y. S., Li, H. T., Li, C. R., & Zhang, D. Z. (2016). Factors affecting hotels' adoption of mobile reservation systems: A technology-organization-environment framework. *Tourism Management*, *53*, 163–172. https://doi.org/10.1016/j.tourman.2015.09.021

Yan Chi, E. C., Goh, H. L., Wui, L. Y., Fong, L. S., & Meng, A. T. G. (2020). Online impressions on E-hailing services: A study on positive and negative sentiments on GRAB Malaysia and GO-JEK Indonesia on Twitter platform. *European Journal of Molecular and Clinical Medicine*, *7*(3), 842–859. https://doi.org/10.31580/apss.v6i2.1323

Zvarevashe, K., & Olugbara, O. O. (2018). A framework for sentiment analysis with opinion mining of hotel reviews. *Conference on Information Communications Technology and Society (ICTAS), 2018*, Durban, South Africa, March 8–9, *2018* (pp. 1–4). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICTAS.2018.8368746