Article

Early Detection Method for Money Fraudulent Activities on E-commerce Platform via Sentiment Analysis

Journal of Entrepreneurship and Business
E-ISSN: 2289-8298

Vol. 9, Issue 2, pp. 121-142. December 2021

Masnita Misiran (Corresponding Author)
Centre for Testing, Measurement & Appraisal,
Universiti Utara Malaysia
and
Department of Mathematics and Statistics,
School of Quantitative Sciences, Universiti Utara Malaysia
Email: masnita@uum.edu.my

Shi Er Tan
Department of Mathematics and Statistics,
School of Quantitative Sciences, Universiti Utara Malaysia
Email: tan_shi_er@sqs.uum.edu.my

Pheng Hong Augustus Saw
Department of Mathematics and Statistics,
School of Quantitative Sciences, Universiti Utara Malaysia
Email: a_saw_pheng@sqs.uum.edu.my

Nur Azuin Mohd Subri
Department of Mathematics and Statistics,
School of Quantitative Sciences, Universiti Utara Malaysia
Email: nur_azuin_mohd@sqs.uum.edu.my

Nur Syazana Ahmad Darus
Department of Mathematics and Statistics,
School of Quantitative Sciences, Universiti Utara Malaysia
Email: n_syazana_ahmad@sqs.uum.edu.my

Zahayu Md Yusof
Department of Mathematics and Statistics,
School of Quantitative Sciences, Universiti Utara Malaysia
and
Institute of Strategic Industrial Decision Modelling
Universiti Utara Malaysia
Email: zahayu@uum.edu.my

Nazihah Ahmad
Department of Mathematics and Statistics,
School of Quantitative Sciences, Universiti Utara Malaysia
Email: nazihah@uum.edu.my

Faculty of Entrepreneurship and Business, Universiti Malaysia Kelantan
Locked Bag 36, 16100 Pengkalan Chepa
Kota Bharu, Kelantan, Malaysia
http://fkp.umk.edu.my/journal/index.html

Date Received: 05 December 2021
Date Accepted: 30 December 2021

DOI: 10.17687/jeb.v9i2.804

Abstract Shopee is one of the prominent online shopping platforms in Malaysia. Nonetheless, countless scam cases were reported on the platform, particularly on money fraudulent activities. This study constructed a quantitative model through linear programming that assessed sentiments based on customers'

reviews. Reviews from three selected Shopee products ('*M3 Smart Health Watch*', '*Sony Headset Wired Gaming Headphone*', and '*20000mAh Pineng 100% Original Powerbank*') were analysed using the proposed model. The data were converted into measurable metrics to enable quantitative fraud detection. The model enabled the early detection of possible money fraudulent activities on Shopee products based on customers' reviews. Resultantly, '*M3 Smart Health Watch*' is an authentic Shopee product. In contrast, '*Sony Headset Wired Gaming Headphone*' and '*20000mAh Pineng 100% Original Powerbank*' are money fraud products sold by scammers. The proposed model utilises free and readily available software, thus extending its usability to other small business owners.
Keywords: Risk Analysis, E-commerce, Shopee, Money Fraud, Sentiment Analysis.

1. Introduction

The digital era has made lives exceptionally efficient and convenient. The working methods is predominantly streamlined with easy-to-use technological support. Similarly, digitalisation influenced how we design our lives in the personal and domestic areas. Day-to-day domestic activities previously viewed as time-consuming can now be performed more effortlessly. Therefore, the online shopping platforms commonly termed the electronic commerce (e-commerce) industry have garnered public attention. E-commerce is defined as follows:

"*Electronic Commerce is defined as doing business electronically. This includes the sharing of structured or unstructured business information by any means (such as electronic mail or messaging, World Wide Web technology, electronic bulletin boards, smart card, electronic funds transfer, and electronic data interchange) among suppliers, customers, government bodies, and other partners to conduct and execute transactions in business, administrative and consumer activities.*"

United Nations Centre for Facilitation of Procedures and Practices for Acquisition, Commerce and Transport, 1997

Ultimately, e-commerce is not without its struggles. Online fraudulent activities due to insufficient and vulnerable cyber security is an ongoing battle commonly shared by online users, including buyers, sellers, and platform providers. Khrestina, Dorofeev, Kachurina, Usubaliev and Dobrotvorskiy (2017) argued the adaptation of emerging technology such as big data, blockchain, internet of things, cryptocurrency, and machine learning is a significant issue to be addressed in transitioning to the digital economy. As an online platform with a substantial user base, the Shopee platform in Malaysia is also affected by this threat, particularly during the Covid-19 Movement Control Order (MCO). During this MCO, Cyber Security Malaysia reported a rise in cybersecurity cases by 82.5% (838 cases) from March 18 to April 7, 2020 compared to 2019. From this statistic, 18% (152) cases included local enterprises, while the remaining 82% involved home users and others.

Notably, Caldeira, Brandao and Pereira (2014) reported a significant increase in internet fraud cases worldwide, thus resulting in billions of dollars losses annually. Smart agents were introduced to monitor financial activities in the e-commerce platform to efficiently track and detect suspicious transactions and report such activities for countermeasure (Maureen, 2015). Furthermore, previous studies have proposed methods to strengthen the e-commerce platform security. Firstly, Weng, Li, Ji, Chu, Lu, Du and He

(2018) developed the Anti-Fraud (ATF) system for large-scale e-commerce sites to detect e-commerce fraud. In addition, Nascimento, Barbosa, Perez, Caires, Hirama, Ramos and Louzada (2019) developed an acoustic analysis to identify fraud activities through the voice activation process for e-commerce transactions.

Presently, Shopee and Lazada denote two primary online shopping platforms in Malaysia. Despite Lazada being a more established e-commerce platform, recent years revealed the gaining momentum and shifting of customers' loyalty towards Shopee. This positive trend results in Shopee becoming the lead e-commerce market in Malaysia in 2020 while claiming dominance in Southeast Asia. In the first quarter of 2020 alone, iPrice disclosed that Shopee successfully attracted 81.82 million visits (including desktops and cell phones), while Lazada recorded 36.96 million visits. Additionally, Shopee gained 27.27 million visits per month than Lazada's at 12.32 million.

With these extensive business offerings, Shopee is vulnerable to scammers' threats. The modus operandi employed by scammers include purchase scamming where the scammer typically poses as buyers abroad. Once the transaction was finalised, they informed the seller that they unintentionally banked in extra money to the bank account and urgently needed a refund. Furthermore, they use threats to cause a fuss or legal issues if their demands are not fulfilled. Similarly, buyers are also vulnerable to scamming products offered on this platform.

Notably, detecting fraudulent purchases/products is challenging and must be actively rectified. Despite many initiatives conducted by Shopee to improve the system's security, scammers also actively identify methods to bypass the established security protocols on the platform. Thus, there is a growing need to identify dynamic and various strategies to overcome this issue. Chowdhury et al. (2018) proposed a specific method to avoid fraud on Facebook. Perceivably, this activity could be halted once people recognise fraud pages. Albeit theoretically sound, the time required to understand its technicality and come out with the analysis is significant, thus discouraging attempts to adopt this method, specifically in the small business setting. Rao, Kannan and Nihanth (2019) used sentimental analysis and data mining to build a framework for detecting fraud applications before installation and found that this method has the potential to be used as means to solve the issue with fraud application. In addition, Gao, Li, Yiyin, Nan, Jian and Tang (2019) increased precision in detecting misleading reviews by using sentiment analysis.

Furthermore, some weaknesses of Shopee are that their business model allows sellers to operate without background checks and Companies Commission of Malaysia (SSM) registration is not compulsory. This business model enables individuals to conduct buying and selling activities more freely. Unfortunately, it also encourages more fraudulent activities. Despite the various initiatives deployed in these e-commerce platforms, ongoing improvements are necessary to increase the security of the business transactions. Ultimately, aligning the opportunities and threats are challenges that most e-commerce platforms face.

In addition, limited studies have attempted to analyse the fraud threats in the Shopee platform. Previous research mainly concentrated on the quality of products offered (Thwe, Tungkasthan & Jongsawat, 2021), the feasibility of this platform as an e-commerce system (Suroso et al., 2020; Erpurini, 2021), and the investigation of customer's perception and behaviour towards this platform (Hidayah & Putri, 2021; Savero & Usman, 2021; Hidayah & Rejeki, 2021). Some exceptions include Km and Tamalanrea (2021) and Sufriadi (2021), whereby they explored fraud threats from the law perspective on buying and selling between buyer and seller via the online platform. Consequently, this study proposed a potential model that can be utilised to detect money fraudulent activities on Shopee based on customers' historical reviews.

2. Methodology

The data were collected based on three selected Shopee products: M3 Smart Health Watch, Sony Headset Wired Gaming Headphone, and 20000mAh Pineng 100% Original Powerbank. This is one of the products with high demand in Shopee. The data were pre-processed to remove all unnecessary contents, such as invalid values, non-text characters, and stop words. Subsequently, the pre-processed data were evaluated for sentiment analysis using lexicon-based analysis. Finally, the sentiment analysis determined whether the data requires further fraud detection. An outcome was concluded based on the fraud detection report. The fraud detection model using sentiment analysis is presented in Figure 1.

Figure 1. Proposed fraud detection model using sentiment analysis

The proposed fraud detection model is beneficial as a fraud Shopee product sold can be detected from the customers' reviews that illustrate their sentiments. A good review implies a good product, while a bad review may signal money fraud.

2.1 Data Collection

The customers' reviews data on the three selected Shopee products were collected from the Shopee official website. These reviews included the customers' product ratings, comments, and reviews with media content. The product ratings are the summary of feelings that exhibits sentiments in the form of star ratings, on a scale of one to five stars. Customers' comments allow for emotion prediction based on the text polarity, whether positive or negative. The reviews with media content detect whether the product is authentic.

2.2 Data Preprocessing

Customers' comments were pre-processed to facilitate subsequent analysis according to the following:

- Invalid values were removed. For example, the system provided the default value as "positive opinion" in the raw data. These values are meaningless for analysis, so they were filtered out.
- Non-text characters were removed. Special characters, punctuation marks, and numbers may conflict with the interpretation of data.
- Stop words were removed. Words such as "a", "the", "is", "are", "and", "with", and so forth were defined as stop words. Such words hinder the understanding of the content of the comments.

2.3 Sentiment Analysis

Sentiment analysis, also recognised as opinion mining, is a content mining method that classifies human emotions in different classes and assists consumers in comprehending the social slant of the product when observing product reviews. Sentiment analysis is possible for text, audio, image, or video. The analysis can be done using different algorithms, such as the naïve bays, *k* means, decision tree, or lexicon-based analysis, which is noted as the bag of words. Sentiment analysis is the most widely recognised content grouping device that determines whether consumers' emotion on a particular topic or a product is positive or negative.

This sentiment analysis can be classified into document, sentence, and phrase levels. Firstly, the document level reviews a single document of a single topic. Nevertheless, some content can be irrelevant to the topic. Secondly, the sentence level checks on every sentence in the document and is more applicable to any document type. This level of sentiment analysis defines every sentence as positive or negative depending on its polarity. Lastly, the phrase level considers the opinion or sentiment-related words. The phrase level is a pinpointed approach for sentiment analysis. In some cases, the exact opinions on a specific topic can be extracted from the contents.

Among these three levels of sentiment analysis, the sentence level was applied in this study using lexicon-based analysis. A custom algorithm was deployed. Customers' reviews of the Shopee products selected were analysed to determine the polarity. Initially, customers' product ratings were converted into a score for a review. A product rated five stars would get two positive scores. Subsequently, a product rated four stars would get one positive score. In contrast, a product would not get any score if it was rated three stars. Alternately, a product with a rating of two stars would get one negative score, while two negative scores were given for a product rated as one star.

After the comments were pre-processed, the whole sentences of the comments were analysed. An overall positive sentence would get one positive score. Else, it would get one negative score. Finally, reviews with media content were also studied. Based on the media content, a proven good product would receive one positive score. Otherwise, one negative score was applied.

The positive score for a review was denoted as R_P and a negative score as R_N. Hence, the positive and negative scores of the review previously converted were used to calculate review polarity by following the equation below,

$$\text{Review polarity,} \qquad R = R_P - R_s. \tag{1}$$

If 'R' is greater than zero, the review is positive. Else, the review is negative. Each review received one point (either positive or negative). For the subsequent review, Equation (1) was applied and points for positive or negative reviews increased based on the number of reviews and their polarity. Lastly, the review polarity was determined.

$$\text{The sentiment of the product,} \quad P_{sentiment} = P_S - N_s, \tag{2}$$

where P_S is the total number of positive reviews and N_S is the total number of negative reviews. A positive product sentiment may indicate that it is a good product. In contrast, a negative product sentiment may imply that it involves money fraud.

The sentiment analysis graph was completed to decide whether the data needed to be analysed for fraud detection. The selected Shopee products should have at least 70% positive reviews. Therefore, the product with more than 30% negative reviews was sent for fraud detection analysis. Linear programming was used to determine the feasible region of the model with decision variables being P_S and N_S. The model constraints could be computed as follows:

$$N_S \leq 0.30T_R$$

$$N_S \leq 0.30(P_S + N_S)$$

$$N_S \leq 0.30P_S + 0.30N_S$$

$$-0.30P_S + 0.70N_S \leq 0 \tag{3}$$

$$P_S + N_S \leq T_R \tag{4}$$

$$P_S \geq 0 \tag{5}$$

$$N_S \geq 0 \tag{6}$$

where T_R is the total number of reviews. The values of P_S and N_S are the *x* and *y* coordinates in the graph, respectively. A product that lies in the feasible region demonstrates its authenticity and does not require further fraud detection analysis. In comparison, the product outside the feasible region must be analysed for fraud detection.

2.4 Fraud Detection

After running the sentiment analysis, the data set was analysed for fraud detection. All customers' comments on the selected Shopee products were included for the fraud detection analysis. The output was divided into 'Fraud' or 'Not Fraud'. A fraud-related positive and negative words library was generated. This library was used to detect a comment considering a product as fraud. Table 1 exhibits the example of fraud-related positive and negative words library.

Table 1. Fraud-related positive and negative words library

Fraud Words	Positive Words	Negative Words
Cheat, cheating, fraud, scammer, scam, fake, blackmail, deceit, extortion, graft, hoax, barratry, con, duplicity, fraudulence, and similar.	Good, great, joy, nice, beautiful, sweet, thanks, marvellous, best, happy, fun, enjoy, thankful, and similar.	Adverse, alarming, angry, annoy, bad, banal, barbed, fail, faulty, fear, feeble, fight, malicious, mean, menacing, messy, and similar.

If positive and negative words appeared in the comments, they received one point on each category for each fraud. Therefore, each word grouped by each library was scored separately. The equation for this analysis is as written below,

The score of the comment, $$C_S = P_C - F_C - N_C, \tag{7}$$

where P_C is the score of positive words, F_C is the score of fraud words, and N_C is the score of negative words. If the comment score is greater than 0, the comment is positive (PC_S), or it will be a negative comment (NC_S) which will be considered to be fraud.

Subsequently, another graph was drawn to determine whether the Shopee products selected involved money fraud. If the selected Shopee product has more than 35% negative comments, it was considered money fraud. Linear programming was used to determine the feasible region of the model. The decision variables of this model are PC_S and NC_S. The constraints of the model can be written as,

$$NC_S \leq 0.35T_C$$

$$NC_S \leq 0.35(PC_S + NC_S)$$

$$NC_S \leq 0.35PC_S + 0.35NC_S$$

$$-0.35PC_S + 0.65NC_S \leq 0 \tag{8}$$

$$PC_S + NC_S \leq T_C \tag{9}$$

$$PC_S \geq 0 \quad (10)$$

$$NC_S \geq 0 \quad (11)$$

where, T_C is the total number of comments. The PC_S and NC_S are the x and y coordinates in the graph, respectively. If the point lies in the feasible region, it demonstrates no money fraud occurring as most comments are positive. Nonetheless, the selected Shopee product can be considered a money fraud if the point lies outside the feasible region.

3. Results and Discussion

Three Shopee products were included in this study to showcase how the analysis was conducted. The first Shopee product was M3 Smart Health Watch sold by Seller A as portrayed in Figure 2.

Figure 2. M3 Smart Health Watch

This product had 59 reviews which were included in the sentiment analysis. The scores converted from the M3 Smart Health Watch reviews are exhibited in Table 2.

Table 2. The Review Polarity of M3 Smart Health Watch

Product Rating	Customer's Comment	Media Content	Review Polarity	Positive Review	Negative Review
2	1	1	4	1	0
2	1	1	4	1	0
2	1	1	4	1	0
2	1	1	4	1	0
2	1	1	4	1	0
0	-1	0	-1	0	1
2	1	0	3	1	0
2	1	1	4	1	0
2	1	0	3	1	0
1	1	0	2	1	0
2	1	1	4	1	0
2	1	1	4	1	0
2	1	1	4	1	0

2	1	1	4	1	0
0	-1	0	-1	0	1
-2	-1	0	-3	0	1
2	1	1	4	1	0
2	1	1	4	1	0
2	0	1	3	1	0
2	1	1	4	1	0
2	0	0	2	1	0
2	1	0	3	1	0
1	1	0	2	1	0
2	1	0	3	1	0
0	1	0	1	1	0
2	0	0	2	1	0
2	0	0	2	1	0
2	0	0	2	1	0
0	0	0	0	0	1
2	1	0	3	1	0
0	0	0	0	0	1
2	1	0	3	1	0
2	1	0	3	1	0
2	0	0	2	1	0
2	1	0	3	1	0
2	1	0	3	1	0
2	0	0	2	1	0
2	1	0	3	1	0
2	0	0	2	1	0
2	0	0	2	1	0
2	0	0	2	1	0
2	0	0	2	1	0
2	1	0	3	1	0
2	0	0	2	1	0
2	1	0	3	1	0
-1	0	0	-1	0	1
2	0	0	2	1	0
2	1	0	3	1	0
2	1	0	3	1	0
1	1	0	2	1	0
2	1	0	3	1	0
-1	-1	0	-2	0	1
2	0	0	2	1	0
-2	0	0	-2	0	1

2	0	0	2	1	0
2	1	0	3	1	0
2	0	0	2	1	0
2	1	0	3	1	0
2	0	0	2	1	0

The review polarity of the M3 Smart Health Watch was calculated by summing up the scores converted based on the customers' product ratings, comments, and media content. If the review polarity is greater than zero, it receives one point in the positive review column. If the review polarity is less than zero, it receives one point in the negative review column. Consequently, the total number of positive and negative reviews obtained was calculated. The M3 Smart Health Watch product had 51 positive reviews and eight negative reviews. Furthermore, the sentiment of M3 Smart Health Watch was computed by subtracting the number of positive and negative reviews, which resulted in 43, denoting a positive value. The positive value indicates that the M3 Smart Health Watch is a good Shopee product. Nevertheless, the sentiment analysis graph was completed to determine whether the M3 Smart Health Watch required a fraud detection analysis as presented in Figure 3.

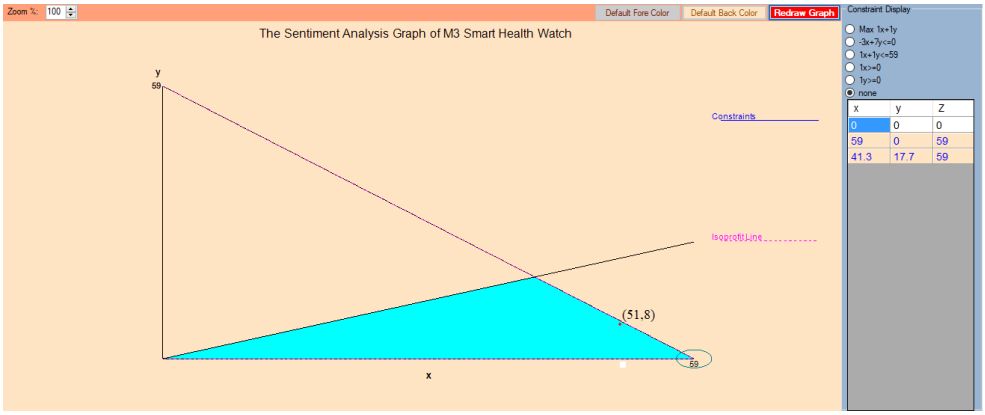

Figure 3. The Sentiment Analysis Graph of M3 Smart Health Watch

In Figure 3, point (51,8) lies in the feasible region of the graph. Therefore, the M3 Smart Health Watch did not require a fraud detection analysis as most reviews are positive. The sentiment analysis also revealed that the M3 Smart Health Watch is a good product. Thus, users can buy the watch without any worries.

Furthermore, the second Shopee product selected was the Sony Headset Wired Gaming Headphone sold by Seller B, as exhibited in Figure 4.

Figure 4. Sony Headset Wired Gaming Headphone

Similarly, this product had 59 reviews included in the sentiment analysis. The scores converted from the Sony Headset Wired Gaming Headphone reviews are portrayed in Table 3.

Table 3. The Review Polarity of Sony Headset Wired Gaming Headphone

Product Rating	Customer's Comment	Media Content	Review Polarity	Positive Review	Negative Review
2	1	1	4	1	0
-1	-1	0	-2	0	1
-2	-1	-1	-4	0	1
-2	-1	0	-3	0	1
2	1	1	4	1	0
1	-1	-1	-1	0	1
-2	-1	-1	-4	0	1
2	1	1	4	1	0
2	1	1	4	1	0
-1	-1	0	-2	0	1
-2	-1	0	-3	0	1
2	1	1	4	1	0
2	1	0	3	1	0
2	1	0	3	1	0
2	1	1	4	1	0
-2	-1	0	-3	0	1
0	-1	0	-1	0	1
0	-1	0	-1	0	1
2	1	1	4	1	0
2	1	0	3	1	0
2	1	0	3	1	0
2	1	1	4	1	0
-2	-1	-1	-4	0	1
2	1	0	3	1	0
2	1	0	3	1	0
2	1	0	3	1	0

-2	0	0	-2	0	1
2	0	0	2	1	0
2	0	0	2	1	0
2	1	0	3	1	0
2	0	0	2	1	0
1	1	0	2	1	0
2	1	0	3	1	0
2	1	0	3	1	0
0	1	0	1	1	0
2	0	0	2	1	0
1	1	0	2	1	0
2	1	0	3	1	0
-2	0	0	-2	0	1
1	0	0	1	1	0
2	0	0	2	1	0
2	0	0	2	1	0
0	0	0	0	0	1
2	1	0	3	1	0
2	1	0	3	1	0
2	0	0	2	1	0
0	0	0	0	0	1
2	0	0	2	1	0
-1	0	0	-1	0	1
2	1	0	3	1	0
2	1	0	3	1	0
2	1	0	3	1	0
-1	0	0	-1	0	1
2	1	0	3	1	0
-2	-1	0	-3	0	1
2	1	0	3	1	0
-2	-1	0	-3	0	1
2	0	0	2	1	0
2	1	0	3	1	0

The Sony Headset Wired Gaming Headphone had 40 positive reviews and 19 negative reviews. The sentiment of the Sony Headset Wired Gaming Headphone was 21, illustrating that it is a good Shopee product. In addition, the sentiment analysis graph was done to check whether the Sony Headset Wired Gaming Headphone required a fraud detection analysis, as demonstrated in Figure 5.

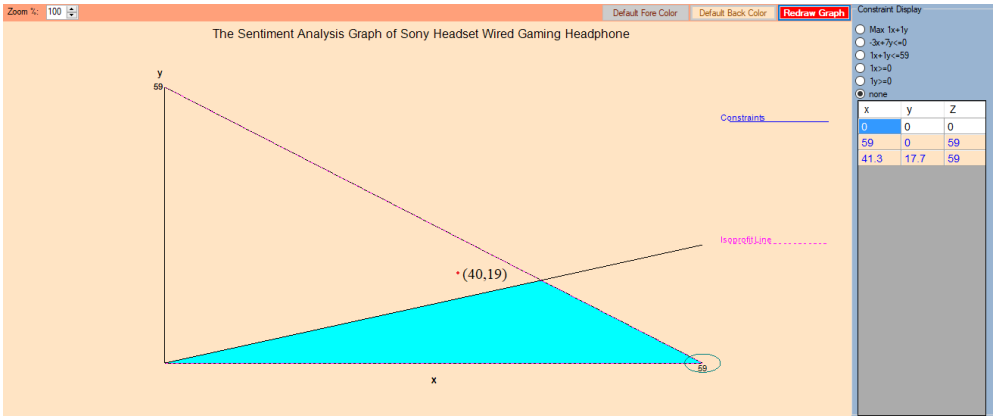

Figure 5. The Sentiment Analysis Graph of Sony Headset Wired Gaming Headphone

Based on Figure 5, point $(40,19)$ lies outside the feasible region, revealing that fraud detection analysis was required for this data set. Furthermore, the Sony Headset Wired Gaming Headphone product had less than 70% positive reviews, proving that it could be a money fraud product.

Therefore, the customers' comments on the Sony Headset Wired Gaming Headphone were reviewed. Every word in the comments was analysed based on the word library presented in Table 1 previously. The scores for each word grouped by each library are demonstrated in Table 4.

Table 4. The Comment Score of Sony Headset Wired Gaming Headphone

Positive word	Negative/ Fraud word	Comment Score	Positive Comment	Negative Comment
2	1	1	1	0
0	1	-1	0	1
0	1	-1	0	1
0	3	-3	0	1
6	0	6	1	0
2	3	-1	0	1
0	3	-3	0	1
2	0	2	1	0
5	1	4	1	0
0	1	-1	0	1
0	2	-2	0	1
3	1	2	1	0
1	0	1	1	0
1	0	1	1	0
1	0	1	1	0
0	3	-3	0	1
0	1	-1	0	1

0	1	-1	0	1
1	0	1	1	0
2	0	2	1	0
1	0	1	1	0
1	0	1	1	0
0	1	-1	0	1
1	0	1	1	0
1	0	1	1	0
1	0	1	1	0
3	0	3	1	0
1	0	1	1	0
2	0	2	1	0
2	0	2	1	0
1	0	1	1	0
1	0	1	1	0
3	0	3	1	0
2	0	2	1	0
2	0	2	1	0
1	0	1	1	0
1	0	1	1	0
1	0	1	1	0
1	0	1	1	0
0	1	-1	0	1
1	0	1	1	0
0	1	-1	0	1
3	0	3	1	0

Based on Table 4, the Sony Headset Wired Gaming Headphone had 43 comments, with 30 positive and 13 negative comments. The fraud detection graph was completed to determine whether the Sony Headset Wired Gaming Headphone had more than 35% negative comments, as portrayed in Figure 6.

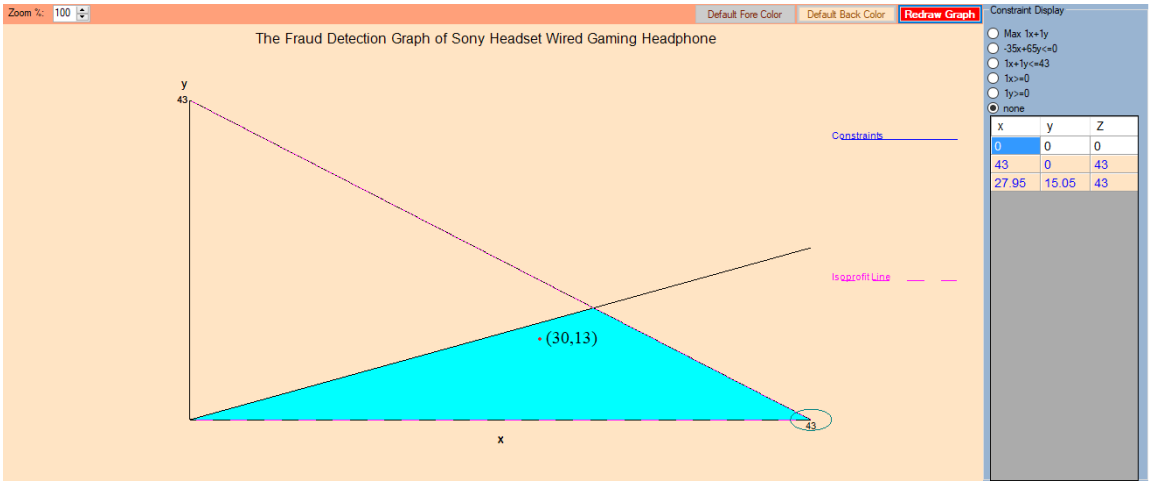

Figure 6. The Fraud Detection Graph of Sony Headset Wired Gaming Headphone

Figure 6 above suggests no money fraud occurring as point $(30,13)$ lies in the feasible region, with most comments being positive. Despite having less than 70% positive reviews, this graph proved that the Sony Headset Wired Gaming Headphone is a good Shopee product as the total number of positive comments was more than 65%.

The last Shopee product selected was the 20000mAh Pineng 100% Original Powerbank which, similarly sold by Seller B, as presented in Figure 7.

Figure 7. 20000mAh Pineng 100% Original Powerbank

Equivalently, this third product underwent the same sentiment analysis performed on the two earlier products. All 78 reviews of 20000mAh Pineng 100% Original Powerbank included in the analysis were converted into scores and demonstrated in Table 5.

Table 5. The Review Polarity of 20000mAh Pineng 100% Original Powerbank

Product Rating	Customer's Comment	Media Content	Review Polarity	Positive Review	Negative Review
-2	-1	-1	-4	0	1
2	1	1	4	1	0

-2	-1	-1	-4	0	1
2	1	0	3	1	0
2	-1	-1	0	0	1
2	1	1	4	1	0
0	1	1	2	1	0
-2	-1	-1	-4	0	1
-2	-1	-1	-4	0	1
0	-1	-1	-2	0	1
-1	-1	0	-2	0	1
-2	-1	0	-3	0	1
1	1	0	2	1	0
-2	-1	0	-3	0	1
2	-1	-1	0	0	1
2	1	1	4	1	0
-2	-1	-1	-4	0	1
0	-1	0	-1	0	1
-2	-1	-1	-4	0	1
0	-1	0	-1	0	1
2	1	1	4	1	0
2	1	0	3	1	0
2	1	1	4	1	0
-2	-1	-1	-4	0	1
1	-1	0	0	0	1
2	1	0	3	1	0
2	-1	0	1	1	0
1	1	0	2	1	0
2	1	1	4	1	0
2	1	0	3	1	0
-2	-1	-1	-4	0	1
2	1	0	3	1	0
2	1	1	4	1	0
2	1	1	4	1	0
2	1	1	4	1	0
2	1	0	3	1	0
2	1	1	4	1	0
2	-1	-1	0	0	1
2	1	1	4	1	0
2	1	0	3	1	0
1	1	0	2	1	0
2	1	0	3	1	0
2	0	0	2	1	0

2	1	0	3	1	0
2	1	0	3	1	0
2	1	0	3	1	0
2	0	0	2	1	0
0	0	0	0	0	1
2	0	0	2	1	0
2	1	0	3	1	0
2	0	0	2	1	0
2	0	0	2	1	0
2	0	0	2	1	0
2	0	0	2	1	0
2	0	0	2	1	0
2	1	0	3	1	0
2	1	0	3	1	0
-2	-1	0	-3	0	1
2	0	0	2	1	0
1	1	0	2	1	0
1	1	0	2	1	0
2	1	0	3	1	0
-2	0	0	-2	0	1
2	1	0	3	1	0
2	0	0	2	1	0
2	0	0	2	1	0
2	0	0	2	1	0
0	0	0	0	0	1
2	0	0	2	1	0
2	1	0	3	1	0
-2	0	0	-2	0	1
2	0	0	2	1	0
-2	0	0	-2	0	1
2	1	0	3	1	0
2	1	0	3	1	0
2	0	0	2	1	0
2	0	0	2	1	0
-1	-1	0	-2	0	1

The 20000mAh Pineng 100% Original Powerbank received 53 positive reviews and 25 negative reviews. The sentiment of 20000mAh Pineng 100% Original Powerbank was calculated as 28, suggesting a good Shopee product. Furthermore, the sentiment analysis graph was completed as presented in Figure 8.

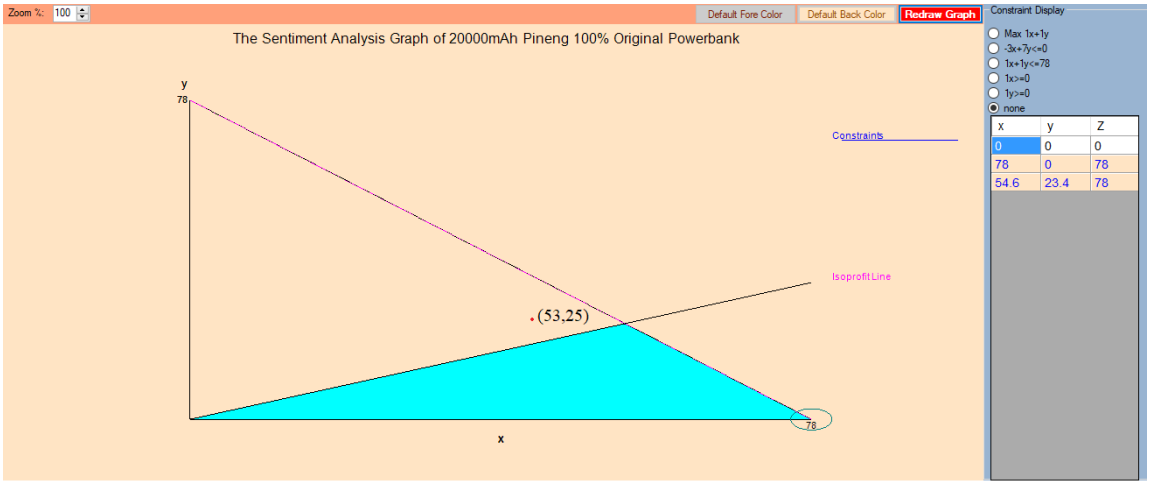

Figure 8. The Sentiment Analysis Graph of the 20000mAh Pineng 100% Original Powerbank

Resultantly, point $(53,25)$ lies outside the feasible region in Figure 8, revealing that the 20000mAh Pineng 100% Original Powerbank had more than 30% negative reviews. Consequently, fraud detection analysis was required. By considering only the comments of 20000mAh Pineng 100% Original Powerbank, the comment scores were calculated and presented in Table 6.

Table 6. The Comment Score of 20000mAh Pineng 100% Original Powerbank

Positive word	Negative/ Fraud word	Comment Score	Positive Comment	Negative Comment
0	2	-2	0	1
2	0	2	1	0
0	6	-6	0	1
1	0	1	1	0
0	2	-2	0	1
7	0	7	1	0
4	1	3	1	0
1	3	-2	0	1
0	2	-2	0	1
0	5	-5	0	1
0	3	-3	0	1
1	0	1	1	0
2	1	1	1	0
0	3	-3	0	1
0	1	-1	0	1
6	0	6	1	0
1	4	-3	0	1
0	1	-1	0	1

0	6	-6	0	1
0	1	-1	0	1
2	0	2	1	0
2	0	2	1	0
1	0	1	1	0
0	1	-1	0	1
0	1	-1	0	1
4	0	4	1	0
0	3	-3	0	1
2	0	2	1	0
1	0	1	1	0
1	1	0	0	1
0	1	-1	0	1
3	0	3	1	0
4	0	4	1	0
3	0	3	1	0
2	0	2	1	0
2	0	2	1	0
2	0	2	1	0
2	1	1	1	0
2	0	2	1	0
1	0	1	1	0
1	0	1	1	0
2	0	2	1	0
4	0	4	1	0
2	0	2	1	0
3	0	3	1	0
4	0	4	1	0
1	0	1	1	0
2	0	2	1	0
0	2	-2	0	1
3	0	3	1	0
1	0	1	1	0
1	0	1	1	0
2	0	2	1	0
1	0	1	1	0
1	0	1	1	0
1	0	1	1	0
0	1	-1	0	1

Based on Table 6, the 20000mAh Pineng 100% Original Powerbank product had 37 positive comments and 21 negative comments overall. The fraud detection graph presented in Figure 9 was completed to determine whether it is a money fraud product.

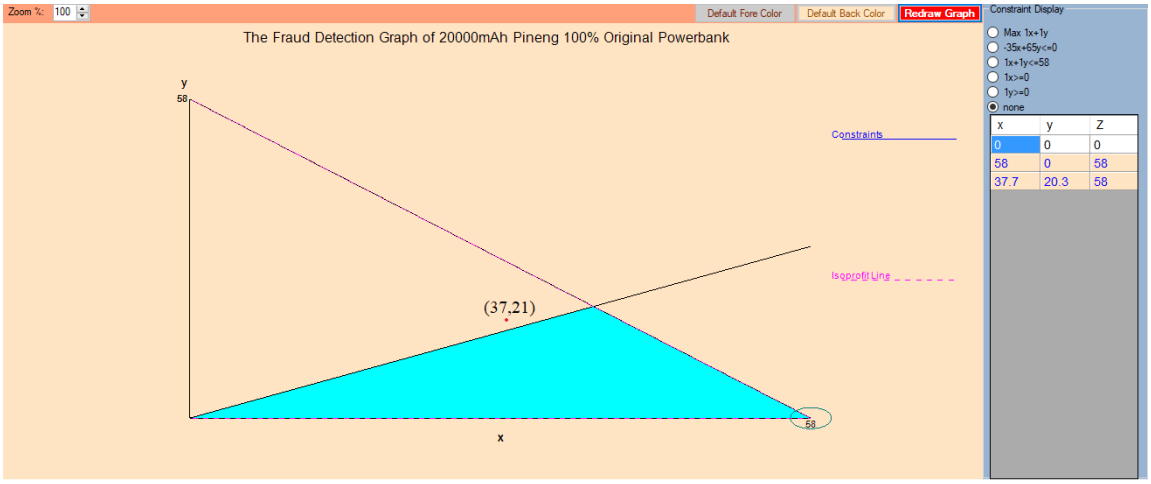

Figure 9. The Fraud Detection Graph of 20000mAh Pineng 100% Original Powerbank

In Figure 9, point (37,21) lies outside the feasible region, thus proving that the 20000mAh Pineng 100% Original Powerbank is a money fraud product. Consequently, buying this specific Shopee product is not advisable.

Notably, Sony Headset Wired Gaming Headphone and 20000mAh Pineng 100% Original Powerbank products were sold by the same seller B. The latter was revealed as a money fraud product. Therefore, there is a significant possibility that the Sony Headset Wired Gaming Headphone is also a money fraud product.

It is vital to note that assessing the authenticity of Shopee products should be based on its reviews and the seller. Having a strong trust in the seller is paramount when buying a particular product on Shopee. Hence, it is not recommended to purchase from a seller that has been identified as a potential scammer. E-commerce platforms can benefit from this early detection mechanism by conducting regular screenings on their product offerings to eliminate scammers from the platforms. This exercise will improve acceptance and trust among users (buyers and genuine sellers alike), thus sustaining their business model in the future.

This study explored the quantitative measurement on fraud detection constructed using linear programming analysis towards otherwise qualitative assessment (sentiment analysis) obtained from the reviews observation on the Shopee platform. Furthermore, the analysis was backed by visualisation to improve understanding of the feasible fraudulent activity for each product. The ability to provide measures for this sentiment study using simple and easy-to-use free software should help other entities use this method more effectively while being cost-effective.

4. Conclusion

This study presented the mechanism of early detection for money fraudulent activities on Shopee products by a model supported by sentiment analysis. The polarity of customers' reviews on selected Shopee products was analysed.

Based on the findings, Seller A's M3 Smart Health Watch was concluded as a good product. In contrast, the Sony Headset Wired Gaming Headphone and 20000mAh Pineng 100% Original Powerbank by Seller B were identified as money fraud products. Notably, the fraud detection graph for Sony Headset Wired Gaming Headphone suggested no indicator for money fraud activity. Nevertheless, buyers are recommended to be cautious when buying this product from Seller B as this seller included other money fraud products in their shop, such as the 20000mAh Pineng 100% Original Powerbank.

The online marketplace has a plethora of opportunities for many individuals. Despite that, money fraud cases are highly damaging and can make users lose trust in the online marketplace. Therefore, both sellers and customers will benefit from a proper use of the online marketplace. Taking necessary steps to detect any fraudulent activities and secure every user's safety are imperative. It has become extremely critical to take the required measures to detect and enforce fraud under strict laws.

Resultantly, this study developed an effective strategy to detect money fraud activities on Shopee. Future research can be conducted by implementing the same method to identify fraudulent activities on other online marketplaces. This detection mechanism can be improved by transforming its formulation into an easy-to-use system that is quicker and more accessible for e-commerce platforms. Furthermore, future research can use this system as an in-house cleaning exercise conducted at regular intervals. Once this exercise is in place, a safer environment on the e-commerce platform can be fostered, with only genuine sellers and buyers taking part in a sustainable and trustworthy environment. Automation of this in-house cleaning exercise can significantly improve the perception and acceptance of the e-commerce platform among individuals, attracting more global participation.

Acknowledgement
The authors would like to thank Universiti Utara Malaysia for its support in completing this research. We would also like to thank the Research and Innovation Management Centre for facilitating the management of this research work.

References

Caldeira, E., Brandao, G., & Pereira, A. C. (2014). Fraud Analysis and Prevention in e-Commerce Transactions. 2014 9th Latin American Web Congress.

Chowdhury, S. M. Mazharul Hoque & Ghosh, Priyanka & Hossain, Syed & Abujar, Sheikh & Afrin, Most. (2017). Analyzing Sentimental State of Any Text Data Using Lexical Analysis and Naive Bayes Algorithm Based on Twitter Data. *1st International Conference on Emerging Technologies in Data Mining and Information Security*.

Chowdhury, S. M. Mazharul Hoque & Nasrin, Samia & Ghosh, Priyanka & Abujar, Sheikh & Hossain, Syed. (2018). Fraud detection of Facebook business page based on sentiment analysis. *8th International Conference on Advances in Computing & Communications, ICACC-2018.*

Erpurini, W. (2021). E-Commerce Used Analysis with SWOT Method in Online Shop Shopee. *The 1st International Conference on Regional Economic and Development.*

Gao, X., Li, S., Zhu, Y., Nan, Y., Jian, Z., & Tang, H. (2019). Identification of deceptive reviews by sentimental analysis and characteristics of reviewers. *Journal of Engineering Science and Technology Review, 12(1),* 196-202.

Hidayah, S. R., & Putri, C. E. (2021). Factors affecting interest in utilization and use of online shop (study on shopee customers). *Journal of Soft Computing Exploration, 2(2),* 107-114.

Hidayah, D. W., & Rejeki, A. (2021, December). Consumer Trust on Buying Interest In Online Shopping: E-Commerce Shopee. *Journal Universitas Muhammadiyah Gresik Engineering, Social Science, and Health International Conference, 1(2),* 617-631.

Khrestina, M. P., Dorofeev, D. I., Kachurina, P. A., Usubaliev, T. R., & Dobrotvorskiy, A. S. (2017). Development of algorithms for searching, analyzing and detecting fraudulent activities in the financial sphere.

Km, J. P. K., & Tamalanrea, M. (2021). Online Selling and Buying Fraud: The Law Of Electronic Transaction Perspective Penipuan Berbasis Jual Beli Online: Perspektif Hukum Transaksi Elektronik. *Jurnal Cendekia Hukum*, 6(2).

Maureen, A. (2015). A Survey of Ecommerce Transaction Fraud Prevention Models. A Survey of Ecommerce Transaction Fraud Prevention Models, 7.

Nascimento, D. C., Barbosa, B., Perez, A. M., Caires, D. O., Hirama, E., Ramos, P. L., & Louzada, F. (2019). Risk Management in E-Commerce—A Fraud Study Case Using Acoustic Analysis through Its Complexity. *Entropy*, 21(11).

Rao, M. R., Kannan, N., & Nihanth, C. H. V. S. (2019). Detecting fraud apps using sentiment research. *International Journal of Recent Technology and Engineering, 8(2),* 580-584.

Savero, R., & Usman, O. (2021). The Effect of Consumer's Perception of Security, Easy, And Consumer Trust on Buying Interest in Online Shopee Selling Site for Uni College Students. The Effect of Security Perceptions, Ease of Transaction, and Consumer Trust on Purchase Intention at Shopee.

Sufriadi, Y. (2021). Prevention Efforts Against E-Commerce Fraud Based on Indonesian Cyber Law. *In 2021 9th International Conference on Cyber and IT Service Management (CITSM), 1-6. IEEE.*

Suroso, J. S., Kaburuan, E. R., Lee, D., Tama, N. S., & Tee, K. (2020, December). Analysis of E-Commerce System in Marketplace (Shopee Indonesia). *In 2020 8th International Conference on Orange Technology (ICOT),* 1-4. IEEE.

Thwe, Y., Tungkasthan, A., & Jongsawat, N. (2021, November). Quality Analysis of Shopee Seller Portal by Using Category Recommendation System Approach. *In 2021 19th International Conference on ICT and Knowledge Engineering (ICT&KE),*1-4. IEEE.

Weng, H., Li, Z., Ji, S., Chu, C., Lu, H., Du, T., & He, Q. (2018). Online E-Commerce Fraud: A Large-Scale Detection and Analysis. *2018 IEEE 34th International Conference on Data Engineering (ICDE).*