ISSN 2409-2665
Journal of Logistics, Informatics and Service Science
Vol. 10 (2023) No. 4, pp.189-205
DOI:10.33168/JLISS.2023.0413

Sentiment Analysis of E-Wallet Companies: Exploring Customer Ratings and Perceptions

Yoon-Teck Bau, Tee Eng Leong*, Chien-Le Goh

[1] Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia

kelvin.tee168@gmail.com (Corresponding author)

Abstract. This sentiment analysis research reports a systematic study of customer reviews of unstructured data for seven popular e-wallet companies, including Alipay, Google Wallet, Grab Superapp, PayPal, Samsung Wallet, Shopee MY, and Touch 'n Go eWallet. Previously, companies faced challenges in effectively utilizing customer reviews to understand and assess customer sentiment toward their products or services. However, with advancements in sentiment analysis techniques, companies can harness the power of customer reviews to gain valuable insights and improve their offerings. The purpose of this study is to explore the use of sentiment analysis in e-wallet companies, where understanding customer sentiment is crucial for enhancing user experiences and driving business success. The research methods employed in this study start by collecting customer reviews spanning four years, from 2019 to 2022. Next, four data pre-processing methods are applied to transform the raw review data into a suitable format for sentiment analysis: data standardization, tokenization, stop word removal, and lemmatization. Sentiment analysis methods were then used to classify reviews as positive, neutral, or negative for each e-wallet company. This study introduced a novel method using rating accuracy to evaluate the polarity sentiment classifications. The results of this study revealed that all e-wallet companies had high rating accuracies for positive sentiments, indicating positive customer sentiment towards their services. However, the rating accuracies for negative sentiments were lower, suggesting challenges in accurately predicting and classifying negative customer sentiment. For neutral sentiments, the rating accuracies were generally low, except for Alipay in 2019, which demonstrated the highest accuracy in capturing customer-neutral sentiment. The evaluation of the findings of this study has important implications for both theory and practice in the e-wallet industry. The practical implications of this study offer concrete guidance for e-wallet companies to enhance their services based on customer sentiments. In contrast, the theoretical implications highlight the need for ongoing research and innovation in sentiment analysis methods that consider other than customer ratings within the e-wallet industry. This study contributes to sentiment analysis by introducing rating accuracy as a measure for evaluating customer reviews accurately. The contribution will provide a more comprehensive understanding of customer sentiments. The methods employed in this study can be applied to enhance sentiment analysis in various customer relationship management domains.

Keywords: Sentiment analysis, E-wallet companies, Rating accuracy evaluation, Customer reviews

1. Introduction

Sentiment analysis has become a popular tool for understanding public opinion and perception of various products and services. In the financial industry, e-wallets have become increasingly popular as a convenient and secure way to conduct transactions. As a result, understanding customers' sentiments towards various e-wallets is crucial for companies in the e-wallet industry to make informed decisions and improve their services.

The existing literature on sentiment analysis in the context of e-wallet companies primarily focuses on analyzing customer sentiments using polarity sentiment analysis techniques. However, there needs to be more literature regarding the accuracy of sentiment classifications with customer-provided ratings. This research addresses this gap by conducting a rating accuracy evaluation for polarity sentiment analysis, specifically focusing on customer reviews of various e-wallet companies. The research problem can be defined as the need to assess the matching between customer-provided ratings and polarity sentiments to enhance the understanding of customer sentiments and improve sentiment classification accuracy in the e-wallet industry.

The research objectives are to implement two sentiment analyses: polarity and rating accuracy evaluation for each financial company. The polarity analysis recognizes the particular companies' positive, negative, or neutral comments. Google Play store is the source to collect data from customers. Data collected from the Google Play store provides better insights from the customer inputs for each e-wallet company so that the respective company app developers can continuously enhance their apps based on users' reviews.

The business benefits of leveraging data collected from the Google Play store and employing sentiment analysis are significant. Using these insights, e-wallet companies can enhance their apps based on real customer feedback, improving user experiences and overall satisfaction. By monitoring the sentiment of user reviews, companies can assess the effectiveness of their app enhancements and determine whether they positively impact user experiences. It also enables companies to compare their performance with competitors by evaluating sentiment scores and identifying areas where they can differentiate themselves. Moreover, positive app reviews and high user ratings can attract new users looking for reliable and user-friendly e-wallet solutions.

The remainder of this paper is arranged as follows. Section 2 presents a comprehensive literature review, highlighting the existing sentiment analysis techniques and the data sources commonly used in such studies. Section 3 provides details of the methodology employed in this research, specifically focusing on the polarity analysis and rating accuracy evaluation. The results obtained from the analysis and subsequent discussions are presented in Section 4. Finally, in the last section, the conclusion of this research is provided, summarizing the key findings and implications derived from the study.

2. Literature Review

This literature review focuses on papers that investigate the field of sentiment analysis. Sentiment analysis, also known as opinion mining, is the process of determining the emotional tone of the text. The papers in this list detail author names, published years, dataset sources, and other details such as sentiment analysis tools.

Abdul Hanan et al. (2023) conducted a study to investigate the sentiment of road users on Twitter and its implications for road infrastructure. Using text-mining techniques and machine learning classifiers, the researchers analyze tweets to gain insights into how road users perceive road conditions in Selangor. The findings reveal a prevalent negative sentiment, highlighting potential areas for improvement in the road infrastructure. The study contributes to the existing literature on sentiment analysis in the context of road infrastructure. It offers valuable insights for local authorities and urban planners in effectively managing and enhancing road conditions based on public sentiment. Future research should focus on expanding the scope of the study to include a broader range of cities and states to strengthen the generalizability of the findings.

Aljedaani et al. (2022) presented a method that integrated TextBlob to analyze sentiments on Twitter about six US airline companies. The authors used the TextBlob sentiment analysis tool. Results of the study suggested that models performed better when trained using the TextBlob assigned sentiments than the original annotated sentiments in the dataset. The study concluded that TextBlob annotated labels can be used as assistance to human annotators.

Almalis et al. (2022) collected their dataset from StockNewsApi.com. The datasets were named TechNews and AllTickers. TechNews is a collection of financial news articles about technology companies. AllTickers is a collection of general market news from different economic sectors. An emotion tag that might be positive, negative, or neutral is present in every news item. In addition, each news item's title, key textual information, source, publication date, derived sentiment (positive, negative, or neutral), and stock tickers mentioned in the news are all included in the data.

Hermansyah et al. (2020) compared the use of TextBlob for sentiment analysis of consumer reviews of Telkom's digital products on Twitter. The dataset for the sentiment analysis was collected from Twitter by utilizing the tweepy python library to gather real-time tweets containing specific product keywords. The keywords used are Indihome, UseeTV, and Wifi.id. The dataset was separated into two sentiment categories that determined whether the sentiment was positive or negative.

Ikoro et al. (2018) presented a study that aimed to analyze the sentiment expressed on Twitter by UK energy company consumers. The authors collected more than 60,000 tweets from nine energy companies. The paper's main conclusion is that using a domain-specific lexicon can improve the accuracy of sentiment analysis.

Iwendi et al. (2022) presented a study that aimed to perform sentiment analysis on COVID-19-related fake news. In this study, 39 features were extracted from multimedia texts and used to detect fake news related to COVID-19. The true news dataset was collected from Harvard Health Publishing, World Health Organization (WHO), the Centers for Disease Control and Prevention (CDC), and The New York Times. The fake news dataset was collected from Facebook posts and other medical sites. The dataset consists of over 1,100 news articles and social media posts about COVID-19, where 586 were true news and 578 were fake news.

Jagadishwari et al. (2022) used sentiment analysis to determine the polarity of text about the COVID-19 vaccines distributed by India and the USA on Facebook and Twitter. A method from TextBlob was used to analyze the sentiments, which are shown as wordcloud, polarity score, and subjectivity score.

Khan et al. (2021) proposed a sentiment analysis of US-based tweets related to the COVID-19 pandemic using machine learning and a lexicon analysis approach. The dataset collected between January 2020 to May 2020 contained 11858 tweets, and tweet sentiments were labeled using TextBlob as positive, negative, or neutral.

Kristiyanti et al. (2020) proposed an e-wallet sentiment analysis model of two e-wallet companies in Indonesia, OVO, and DANA, by analyzing customer reviews on the Google Play Store. The dataset size is 2000 and filtered from March 2019 to January 2020. The stages of the study included data collection, initial data processing, and modeling with positive and negative sentiment reviews. The results indicated that OVO is a more widely used e-wallet application in Indonesia.

Kusrini & Mashuri (2019) presented a method for sentiment analysis on Twitter data that combined lexicon-based and polarity multiplication techniques. The authors described using a combination of these techniques to analyze sentiments on Twitter by evaluating the effectiveness of the lexicon-based and polarity multiplication methods separately and then comparing their performance. The study provided 250 English data, and 112 adjectives were used as lexicon.

Leelawat et al. (2022) collected their dataset from Twitter. From July 1 to December 31, 2020, tweets in English on trips to Thailand were collected using the Twitter Application Programming Interface (API), accessible using the Tweepy library. The tweets are required to include Phuket, Chiang Mai, or Bangkok to be considered valid.

Maindola, Singhal, & Dubey (2018) presented a method for sentiment analysis applied to digital wallets and Unified Payments Interface (UPI) systems in India post-demonetization. The authors analyzed users' sentiments using IBM Watson software tools in various social networks, considering different payment systems. The study analyzed documents, forums, tweets, and other sources where users shared their feedback for the wallets from November 8, 2016, to November 7, 2017. The authors aimed to identify users' sentiments towards digital wallets and UPI systems, post-demonetization.

Mansouri et al. (2022) proposed a model to classify tweets about COVID-19 vaccines into three classes such as negative, neutral, and positive. A large dataset of 174,490 tweets was collected from Twitter. After data collection, the data was analyzed using TextBlob to categorize the sentiments.

Nair et al. (2021) conducted a comparative study on Twitter sentiment on COVID-19 using the VADER technique. This research was useful for government and health officials to understand public opinion and make decisions based on the sentiment results. The VADER technique will only classify tweets as positive, negative, or neutral. VADER technique performance results are reduced because it did not look at the aspects of the sentences.

Poornima & Priya (2020) collected their dataset from Twitter. The tweets were compiled into a CSV file with the following information: id, datetime, sentiment, and feeling about a good or event. A collection of words, emoticons, symbols, URLs, and references to specific persons were included in the dataset.

Punetha & Jain (2023) collected their dataset from the Amazon.com website. The authors proposed a sentiment tagger using a combination of a review's star rating and textual feedback to classify binary class sentiment into positive or negative sentiment. There were 1,000 reviews, comments, and ratings in each dataset.

Rahmanti et al. (2022) analyzed the people's sentiment toward the COVID-19 vaccination on Twitter. Between October 15, 2020, and April 12, 2021, the authors collected 555,892 tweets from Indonesians about the COVID-19 vaccine. Additionally, the official platform of the Ministry of Health and the database of KawalCOVID19 also provided the authors with information on the daily trend of COVID-19 vaccine coverage, case growth, and case fatality. Before the vaccination rollout, it was observed that negative sentiments (65.50%) outnumbered positive sentiments. Positive sentiments (62.20%) exceeded negative sentiments after the vaccination deployment as trust increased.

Sangeetha & Kumaran (2023) collected their dataset from Kaggle.com about Amazon reviews. A 5-star rating system is used for Amazon products. Twenty-five thousand reviews made up the dataset, of which 19,500 were used for training and 5,500 for testing.

Smetanin & Komarov (2019) collected their dataset from a major e-commerce site in Russia. The "Women's Clothes & Accessories" category was the only one from which the dataset was collected. The authors collected a dataset of 821k that comes with labeled reviews.

Srivats Athindran et al. (2018) presented a study using sentiment analysis to understand customer perceptions of two leading smartphone brands, Vivo and Oppo, in India. The authors collected tweets about the Vivo Nex and Oppo FindX smartphones as customer feedback. They then compared the overall sentiment of the two smartphones and also the sentiment of individual features in the two smartphones, which can provide valuable feedback for the companies to improve their products.

Suddle & Bashir (2022) collected a dataset from various sources such as IMDB, Twitter, hotels, and Amazon reviews. The IMDB dataset comprises 50,000 reviews and their binary labels for sentiment analysis of movie reviews on IMDB. The dataset was highly balanced, including 25,000 positive and 25,000 negative reviews, with a vocabulary size of 124,253. The Twitter dataset included the emotional coloring of tweets, with a total of 100,000 comments that have binary labels. The vocabulary size for this dataset is 106,613, and it is slightly imbalanced, with 56,457 positive reviews and 43,543 negative reviews. The third dataset included reviews of musical instruments on Amazon. It is relatively small, with only 10,261 reviews, and highly imbalanced, with 6,986 reviews with a five-star rating. The London-based hotels' dataset included reviews of the ten most and least expensive hotels in London. It was relatively large, with 27,328 reviews, and imbalanced, with 18,326 reviews with a five-star rating.

Talahaturuson et al. (2022) presented a study that aimed to analyze the attitude of Indonesian citizens toward investment patterns by analyzing three months of tweets. The tweets were collected using Twint, an open-source Python library, and TextBlob was used to process and analyze the data. The results showed that 92% of the tweets had favorable feelings, and 42% were positive sentiments toward investments in Indonesia. However, the accuracy of the results could have been improved by the data pre-processing and data labeling methods used in the study.

Wang, Niu, & Yu (2020) presented a method for sentiment analysis on Twitter data that incorporated text information and sentiment diffusion patterns. The authors can collect Twitter tweets and retweet data through their study collaborations with Beijing Intelligent Starshine Information Technology Corporation, one of China's top big data gathering and mining service providers. The authors collected the dataset comprising more than 100,000 labeled tweets and retweets.

Wongkar & Angdresey (2019) collected a dataset related to the Republic of Indonesia presidential candidates in 2019 from Twitter. The study conducted several steps: collecting data using Python libraries and text processing. The authors used a crawler to collect tweets from Twitter social media. The tweets were obtained from January to May 2019.

Yang, Li, Wang, & Sherratt (2020) presented a method for sentiment analysis on Chinese e-commerce product reviews that combined sentiment lexicon and deep learning. The book reviews from Dangdang were crawled on the web pages and were used as the dataset. This dataset contained a rating column that ranged from one to five scores. The authors classified the scores into two groups, of which 1 to 2 scores were considered negative reviews, and 3 to 5 were considered positive reviews. One hundred thousand reviews were included in the dataset, of which 50,000 were positive, and 50,000 were negative.

3. Research Methodology

The research methodology employed in this study focuses on dataset collection, data pre-processing, and investigation of sentiment analysis and evaluation of findings. Fig. 1 illustrates the flowchart of the sentiment analysis methodology. The research methodology for this study follows a cyclic flow starting with data collection, followed by data pre-processing, sentiment analysis, evaluation of findings, and then looping back to data pre-processing and sentiment analysis for the next e-wallet company. In the data collection phase, customer reviews of e-wallet companies were collected over a four-year period. The collected data then underwent data pre-processing, which involved techniques such as data standardization, tokenization, stop word removal, and lemmatization to ensure the data was in a suitable format for sentiment analysis. The pre-processed data was then subjected to sentiment analysis using sentiment analysis tools, which classified the reviews as positive, neutral, or negative. The findings of the sentiment analysis were then evaluated to identify patterns and trends in customer sentiment towards e-wallet companies. Based on the evaluation, further data pre-processing and sentiment analysis iterations may be performed to refine the analysis and gain insights for each e-wallet company. This cyclic flow allows for an iterative and comprehensive analysis of customer sentiment in the e-wallet industry, enabling continuous improvement and a more nuanced understanding of customer perceptions. The details of each step in the research methodology will be explained by the following sections.

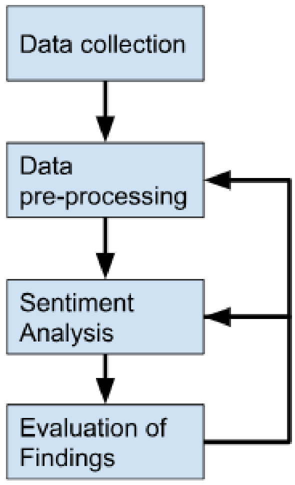

Fig. 1: Sentiment analysis methodology

3.1. Dataset collection

The first step in the methodology is dataset collection. Google Play Store reviews on each e-wallet company were extracted to form the datasets. The total of 10 features were recorded in every e-wallet company's datasets and the sources of customer ratings are from the feature called "score". In this study, seven e-wallet companies were selected for sentiment analysis, namely Alipay, Google Wallet, Grab Superapp, PayPal, Samsung Wallet, Shopee MY, and Touch 'n Go eWallet. The e-wallet companies were selected based on their market presence, popularity, geographic representation, variation in features, substantial user base and usage frequency. Google Play Scraper is the library to help retrieve user reviews from Google Play Store. The four-year reviews are filtered by year from 2019 to 2022 using APIs from the Google Play Scraper library to retrieve from the Google Play Store app data. The four-year reviews are collected from 2019 to 2022 for each e-wallet company, as follows:

- Alipay Play Store Review Dataset: This dataset contains 2,800 customer reviews for Alipay services.
- Google Wallet Play Store Review Dataset: This dataset contains 74,000 customer reviews for Google Wallet services.
- Grab Superapp Play Store Review Dataset: This dataset contains 296,000 customer reviews for Grab Superapp services.
- PayPal Play Store Review Dataset: This dataset contains 400,000 customer reviews for PayPal services.
- Samsung Wallet Play Store Review Dataset: This dataset contains 75,000 customer reviews for Samsung Wallet services.
- Shopee MY Play Store Review Dataset: This dataset contains 400,000 customer reviews for Shopee MY services.
- Touch 'n Go eWallet Play Store Review Dataset: This dataset contains 36,000 customer reviews for Touch 'n Go eWallet services.

3.2. Data Pre-processing

Data pre-processing is a crucial step in the sentiment analysis process to ensure that the collected data from the Google Play store is in a format that can be understood by the computers. The input data needs to be standardized and prepared to enhance the performance of the sentiment analysis model. There are four data pre-processing techniques applied in this study.

Firstly, non-alphabetic characters are removed from the data. This step helps eliminate symbols, numbers, or special characters that may not contribute to sentiment analysis. Additionally, converting all characters to lowercase is beneficial for consistency, ensuring that words with different capitalizations are treated the same way.

Secondly, tokenization is another important pre-processing step utilized in this research. It involves breaking down the text into individual words or tokens. Tokenization enables the sentiment analysis model to analyze and understand the sentiment associated with each word or token separately. By treating each word as a separate unit, the model can more effectively capture the users' sentiments.

Thirdly, removing stop words is performed as part of data pre-processing techniques. Stop words are common words such as "and", "the", and "but" that do not carry significant meaning or contribute much to sentiment analysis. The data size is reduced by removing stop words from the dataset, and the sentiment analysis model can focus on more informative words, enhancing its performance.

Fourth and lastly, lemmatization technique is applied to reduce the dimensionality of the dataset. Lemmatization, on the other hand, maps words to their dictionary base form. For example, the similar words such as running, runs, and ran are all mapped to the base "run". By applying this technique, variations of words are normalized, reducing redundancy and simplifying the sentiment analysis process.

3.3. Sentiment Analysis using TextBlob

TextBlob is the standard Python programming library that provides a convenient way to perform common natural language processing tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, and more. TextBlob performs sentiment analysis by utilizing a pre-trained sentiment analysis model based on a dictionary of words and their associated sentiment scores. When given a text input, such as a customer review or a piece of text, TextBlob follows a series of four steps to analyze the sentiment expressed within the text.

Firstly, the input text is tokenized into individual words or tokens. This tokenization process allows TextBlob to analyze the sentiment of each word or token separately. Each word is then assigned a polarity score based on its presence in the sentiment dictionary. The polarity score indicates the sentiment polarity associated with the word, whether positive, negative, or neutral.

Secondly, the polarity scores of all the words or tokens in the text are aggregated to compute an overall sentiment score for the entire text. This aggregation process combines the sentiment scores of individual words to determine the overall sentiment expressed in the text. The sentiment score represents the degree or intensity of the sentiment within the text.

Thirdly, TextBlob classifies the text based on the computed sentiment score into one of three categories: positive, neutral, or negative. This classification clearly indicates the sentiment expressed in the input text. For example, suppose the sentiment score indicates predominantly positive words. In that case, TextBlob will classify the text as having a positive sentiment.

Fourth and lastly, rating accuracy is a measure used to evaluate the alignment between the rating sentiment provided by customers and the polarity sentiment determined through sentiment analysis. In this work, the concept of rating sentiment evaluation as a novel approach is introduced to assess the match between these two sentiments. The evaluation involves comparing the two columns, rating sentiment, and polarity sentiment, and determining whether they are the same or different. To determine the rating accuracy, each sentiment in the rating sentiment column is compared to its corresponding sentiment in the polarity sentiment column. If the sentiments in both columns are the same, indicating a match, the result is labeled as "True." Conversely, if the sentiments differ, indicating a mismatch, the result is labeled as "False."

Implementing polarity analysis and rating accuracy evaluation aligns with the research objectives for several reasons. Firstly, polarity analysis allows for the assessment of customer sentiments by categorizing them as positive, negative, or neutral. This analysis provides valuable insights into the overall sentiment distribution and helps understand customer perceptions and opinions towards financial companies. Secondly, rating accuracy evaluation specifically focuses on evaluating the matching between the rating-based sentiment analysis provided by customers and the polarity-based sentiment analysis. By implementing both polarity analysis and rating accuracy evaluation, the research can comprehensively analyze and evaluate customer sentiments towards financial companies, shedding light on their satisfaction levels and providing valuable information for improving customer experiences.

3.3.1. Sentiment Analysis: Polarity

In sentiment analysis, "polarity" describes whether a text is positive, negative, or neutral. Positively polarized text is thought to have positive sentiment, the negatively polarized text is believed to have a negative sentiment. In contrast, the neutrally polarized text is considered to have a neutral sentiment. Polarity in sentiment analysis is frequently established by looking for particular words or phrases in the text. For instance, a text including the words "good", "happy", or "love" is likely to be considered positively. However, a text containing the terms "bad", "sad", or "hate" is probably going to be viewed negatively. Neutral sentiment refers to facts or knowledge. For example, the sentence "There is a book on the table." The neutral terms in this sentence are "book" and "table".

Polarity can be determined using various methods. Rule-based approaches use a set of predetermined rules to identify the sentiment of a text, and sentiment analysis tools employ algorithms to learn to recognize sentiment from training data. Polarity is a crucial component of sentiment analysis since it is frequently used to classify texts as positive, negative, or neutral. It is helpful for projects like opinion mining or consumer sentiment analysis. The function f shows the piecewise equation of polarity in Eq. 1.

$$f\,(\text{polarity}) \begin{cases} \text{positive} & \text{if polarity} > 0; \\ \text{neutral} & \text{if polarity} = 0; \\ \text{negative} & \text{if polarity} < 0. \end{cases} \qquad \text{(Eq. 1)}$$

The piecewise function above defines the metric used to quantify sentiment polarity in sentiment analysis, namely polarity score, which measures the sentiment or emotion expressed in a given text. The polarity score typically ranges from -1 to 1, with negative values indicating a negative sentiment, zero indicating a neutral sentiment, and positive values indicating a positive sentiment.

The polarity score in sentiment analysis measures the sentiment or emotion expressed in a chunk of text. The polarity score is calculated by analyzing the words in the text and comparing them to a lexicon or a list of words pre-labeled as positive, negative, or neutral. Each word in the text is assigned a polarity score based on its corresponding label in the lexicon. These scores can be based on various factors, such as the semantic orientation of the word or its association with specific sentiments. The average polarity score of the text is then calculated by summing up the individual polarity scores and dividing by the number of all the words in the text. For example, if a text contains predominantly positive words with high polarity scores, the overall polarity score of the text would be positive, indicating a positive sentiment. Conversely, if the text contains mostly negative words with low polarity scores, the overall polarity score would be negative, indicating a negative sentiment.

The data frame of customer reviews with polarity sentiment was presented in Fig. 2. Upon analyzing Fig. 2, the presence of the word "awful" classifies the review as negative sentiment. Conversely, words such as "good", "quick", "wonderful", "easy", and other positive words contribute to the classification of reviews as positive sentiment.

content	Polarity	Analysis
job	0	Neutral
awful customer service give u run around	-1	Negative
good quick service	0.516667	Positive
think pay pal wonderful always helpful let know straight away anything wrong thank pay pal wonderful service nice	0.46	Positive
easy fast nice	0.411111	Positive

Fig. 2: Customer reviews with polarity score and polarity analysis

3.3.2. Sentiment Analysis: Rating Accuracy

Sentiment analysis is crucial in understanding customer opinions and sentiments toward various products and services. In this report, a specific aspect of sentiment analysis, known as rating accuracy evaluation, is explored to evaluate the matching between the rating sentiment provided by customers and the polarity sentiment determined. This work is the first to utilize rating accuracy to assess the polarity sentiments. By comparing the two columns, rating sentiment, and polarity sentiment, a match is determined based on their sameness. If the sentiments in both columns are the same, the result is considered a match and labeled "True." Conversely, if the sentiments differ, the result is labeled "False."

An evaluation is conducted for each sentiment category from 2019 to 2022 to assess the rating accuracy, encompassing the entire datasets of e-wallet companies. The rating accuracy evaluation compares the matching between the customer-provided ratings and the polarity sentiment classifications.

The sentiment results are obtained by applying the rating accuracy function, as shown below, to predict customer sentiments based on customer inputs accurately. These findings explain the matching between customer perceptions and the polarity sentiment analysis outcomes. These findings also measure the model's effectiveness in capturing and classifying customer sentiments accurately. The function f shows the piecewise equation of rating accuracy in Eq. 2.

$$f\,(\text{rating}) = \begin{cases} \text{positive} & \text{if rating} >= 4; \\ \text{neutral} & \text{if rating} = 3; \\ \text{negative} & \text{if rating} < 3. \end{cases} \qquad \text{(Eq. 2)}$$

The piecewise function above defines the process of how the customer ratings were processed and transformed into sentiment categories through a predefined mapping based on the rating scores. The aim was to convert the numerical ratings provided by customers into sentiment-related labels that align with the sentiment analysis criteria. This process allowed for a standardized representation of the customer ratings in terms of positive, neutral, and negative sentiment-related labels.

Specifically, if the customer rating was greater than or equal to 4, it was assigned the sentiment score or category of "positive". This indicated that the customer had a positive perception or experience with the e-wallet company. If the rating was exactly 3, it was assigned the sentiment score or category of "neutral". This denoted that the customer had a neutral stance or perception, neither strongly positive nor negative, towards the e-wallet company. On the other hand, if the rating was less than 3, it was assigned the sentiment score or category of "negative". This signified that the customer had a negative sentiment or experience with the e-wallet company.

By mapping the customer ratings to these sentiment categories, the ratings were transformed into sentiment-related labels that facilitated a more intuitive and meaningful analysis of customer sentiments. This transformation allowed for a direct comparison between the customer ratings and the polarity sentiments obtained through sentiment analysis. By comparing the rating sentiment and polarity sentiment, this research can assess the rating accuracy. It provides insights between customer input ratings and polarity sentiment analysis results. These findings can help e-wallet companies identify areas for improvement in understanding and interpreting customer sentiments.

The data frame of customer reviews with rating score, rating sentiment and polarity sentiment is presented in Fig. 3. In Fig. 3, the "Actual" column represents the rating sentiment provided by the customers based on the "score" column. In contrast, the "Analysis" column represents the polarity sentiment determined through sentiment analysis by polarity score. For instance, the rating sentiment is positive in the first review, but the polarity sentiment is classified as neutral. This mismatch between the rating and polarity sentiments would result in a false outcome for the rating accuracy evaluation, as the two sentiments do not align.

content	score	Actual	Analysis
job	5	Positive	Neutral
awful customer service give u run around	1	Negative	Negative
good quick service	5	Positive	Positive
think pay pal wonderful always helpful let know straight away anything wrong thank pay pal wonderful service nice	5	Positive	Positive
easy fast nice	5	Positive	Positive

Fig. 3: Customer reviews with rating score, rating sentiment and polarity sentiment

The algorithm below is a rating accuracy analysis function that takes a review text input and aims to compare the rating sentiment assigned to the review text with the polarity sentiment. This analysis helps determine whether this sentiment analysis accurately predicts the sentiment based on the customer-provided rating.

Algorithm: Rating Accuracy Analysis

---

function getRatingAccuracyAnalysis(text)
1. Get rating sentiment of the review text based on the rating score within the range [1, 5]
2. If rating score < 3 then rating sentiment = negative sentiment
3. Else If rating score == 3 then rating sentiment = neutral sentiment
4. Else rating sentiment = positive sentiment
5. If (rating sentiment == polarity sentiment) then return True
6. Else return False

end function

---

4. Results and Discussion

The sentiment analysis results for all e-wallet companies are presented in the form of bar charts, providing a visual representation of the negative, neutral, and positive sentiment percentages and the corresponding rating accuracy percentages from 2019 to 2022.

The analysis of negative sentiment and rating accuracy percentages for all e-wallet companies from the year 2019 to 2022 are presented in both Fig. 4 and Fig. 5. As shown in Fig. 4 and Fig. 5, it is observed that Alipay's sentiment and rating accuracy percentages for negative sentiments remained relatively stable over the years. In additional, the negative sentiment percentages as shown in Fig. 4 for other e-wallet companies which are Google Wallet, Grab Superapp, PayPal, Samsung Wallet, and Shopee MY showed an increasing trend, indicating that customers expressed more negative opinions towards these e-wallet companies, while Touch 'n Go, experienced and decreasing trend. Notably, all rating accuracy percentages as shown in Fig. 5 are low, less than 50%, indicating that all e-wallet companies were less affected by the negative sentiments.

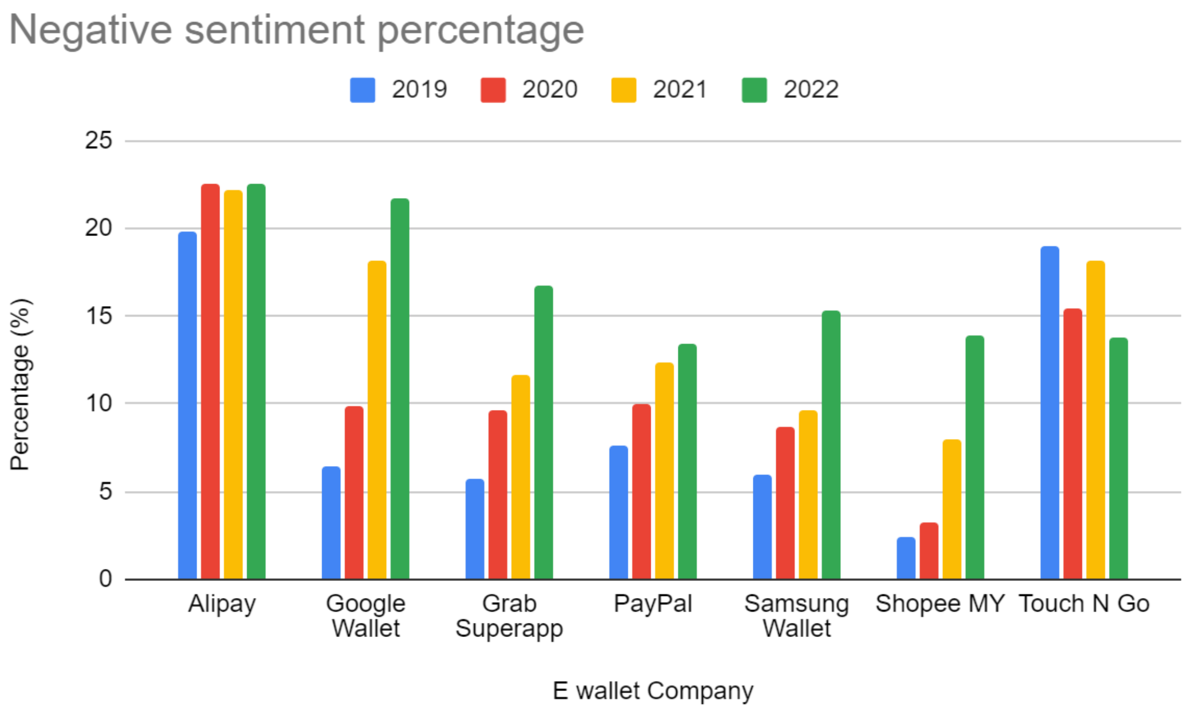

Fig. 4: Negative sentiment percentage

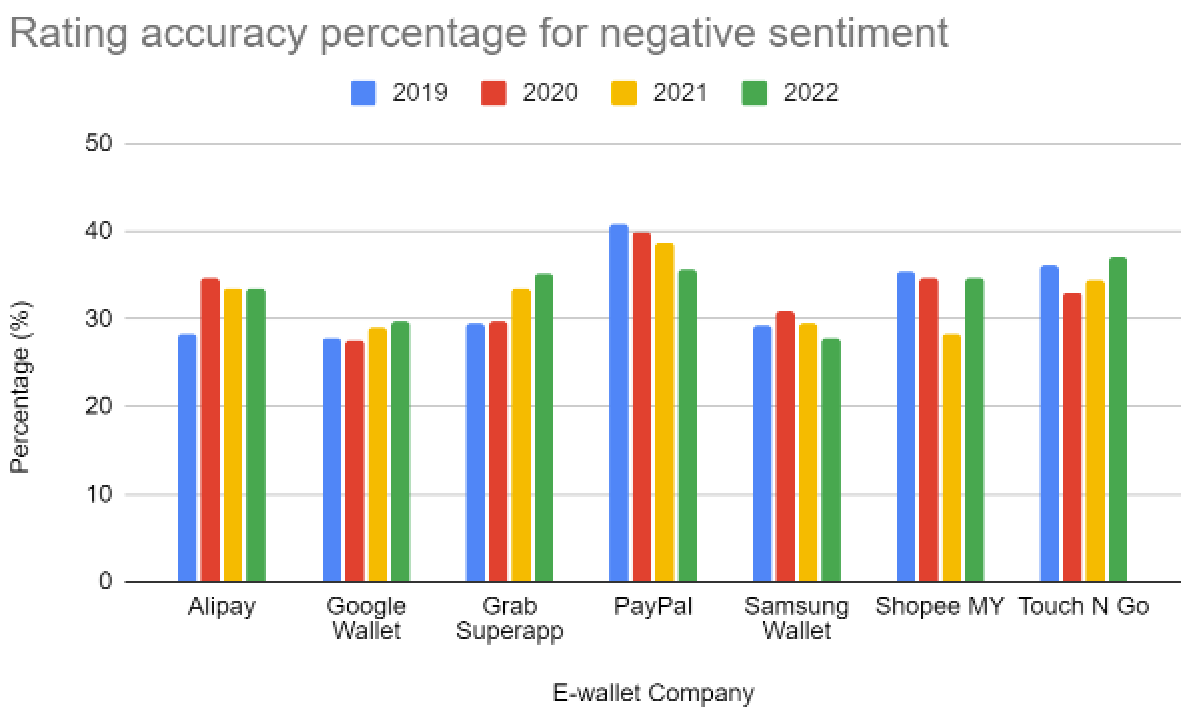

Fig. 5: Rating accuracy percentage for negative sentiment

Moving on to the analysis of neutral sentiment and rating accuracy percentages for all e-wallets from the year 2019 to 2022 are presented both in Fig. 6 and Fig. 7. As shown in Fig. 6 and Fig. 7, it is interesting to note that Alipay in 2019 showed relatively high neutral sentiment percentages and rating accuracy. It suggests that customers often expressed mixed or ambivalent opinions towards services provided by this e-wallet company. However, as shown in Fig. 6, the neutral sentiment percentages for Alipay decreased over the years, while the other e-wallet companies, except Touch 'n Go, experienced an increasing trend. From Fig. 7, the rating accuracy for neutral sentiments was generally low across all companies, indicating that customers are more inclined to give either negative or positive sentiments. Interestingly, Alipay demonstrated the highest rating accuracy for neutral sentiment in 2019, indicating that the customer neutral sentiment in that year was accurately captured.

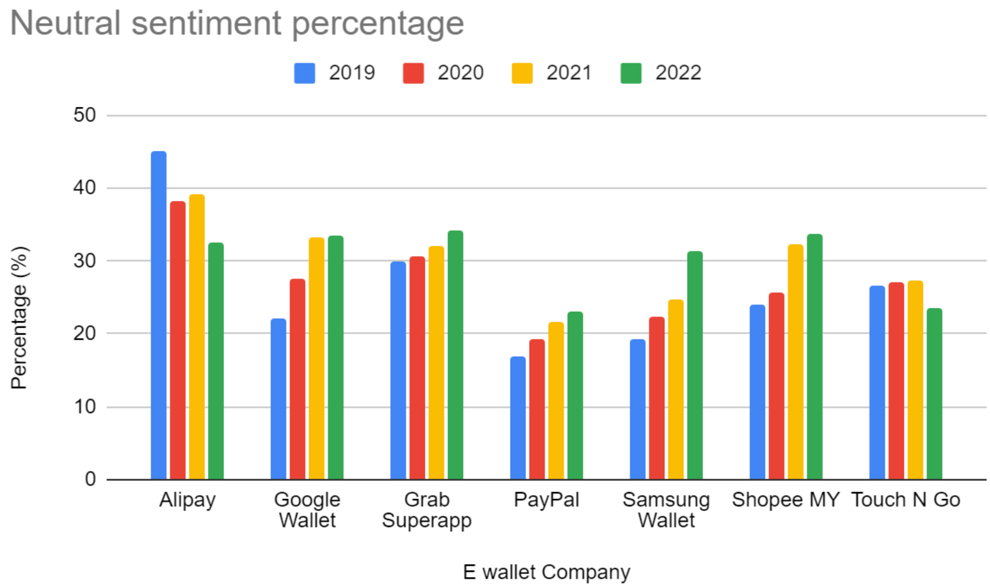

Fig. 6: Neutral sentiment percentage

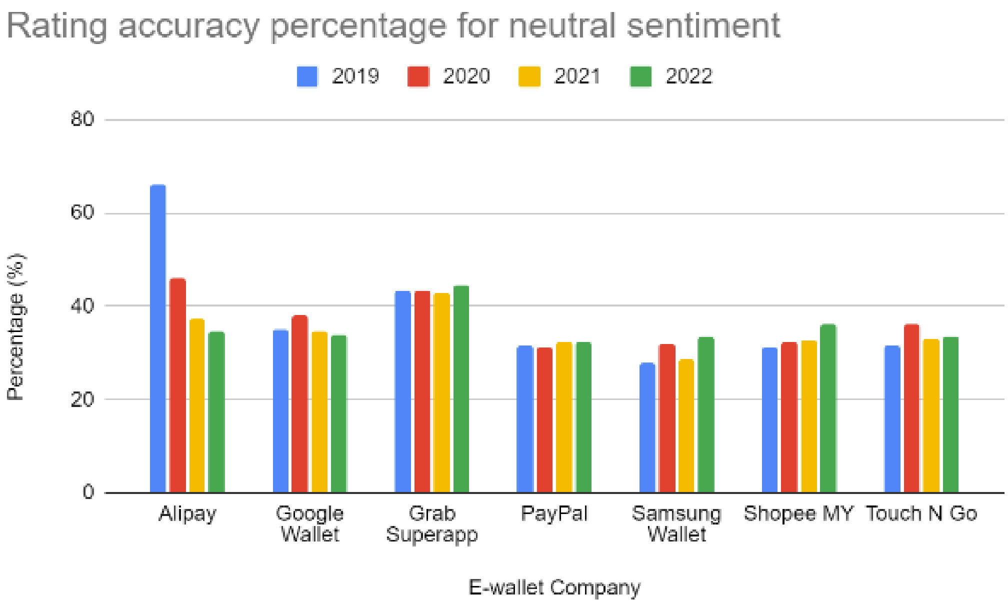

Fig. 7: Rating accuracy percentage for neutral sentiment

Lastly, the analysis of positive sentiment and rating accuracy percentage for all e-wallets from the year 2019 to 2022 are presented both in Fig. 8 and Fig. 9. As shown in Fig. 8, it is important to note that the positive sentiment percentages of some e-wallet companies which are Google Wallet, Grab

Superapp, PayPal, Samsung Wallet, and Shopee MY, showed a decreasing trend over the years, while Alipay and Touch 'n Go exhibited an increasing trend. Additionally, the rating accuracy percentages as shown in Fig. 9 for positive sentiment were generally high for all e-wallet companies, suggesting that the sentiment analysis model accurately predicted the positive sentiment.

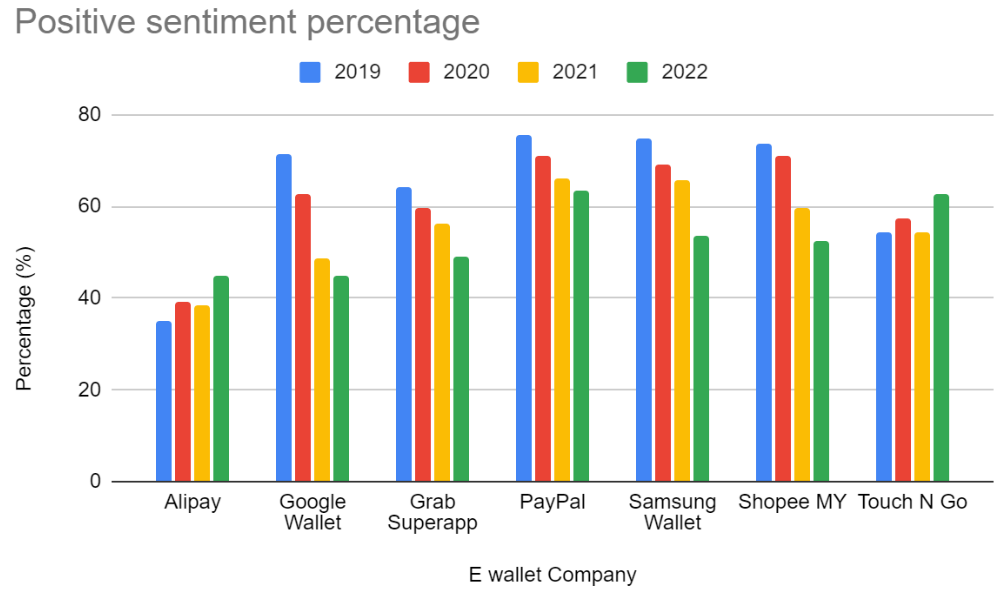

Fig. 8: Positive sentiment percentage

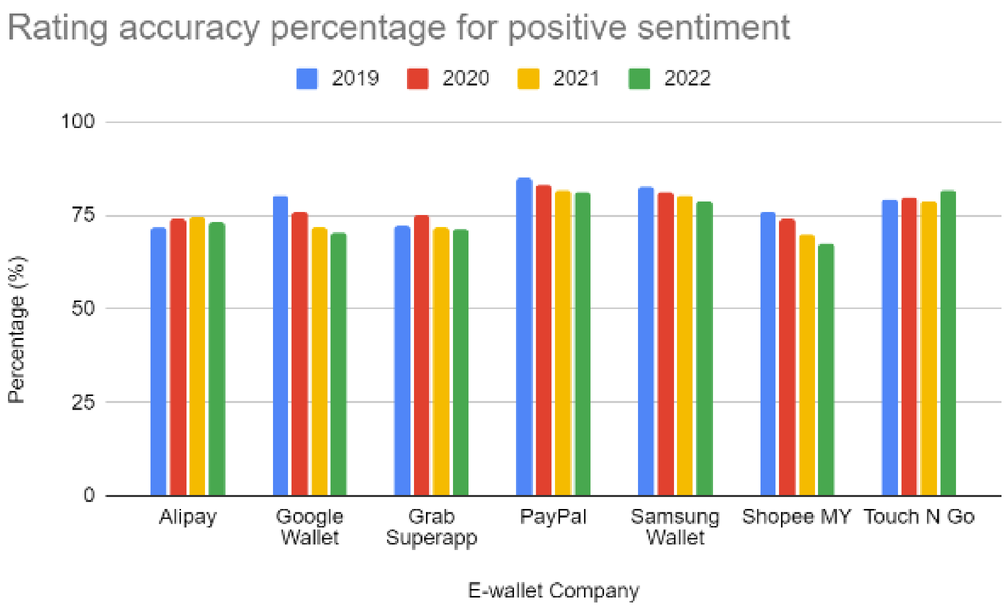

Fig. 9: Rating accuracy percentage for positive sentiment

The findings of this research align with existing literature on customer sentiment analysis in the context of e-wallet companies. The study reveals a prevalence of negative sentiments or an increasing negative sentiment trend expressed by customers towards certain e-wallet companies, highlighting the need for improvements in the identified areas. One novel aspect of this study is that the analysis of rating accuracy for different sentiment categories provides unique insights into the alignment between customer-provided ratings and polarity sentiment classifications, indicating room for improvement in accurately capturing the actual sentiments. Overall, this research contributes to the understanding of customer sentiment dynamics underscore the e-wallet industry needs for continuous monitoring and improvements.

These findings hold important implications for theory and practice in the e-wallet industry. Theoretically, this study establishes that all the implemented concepts, including data pre-processing methods, are necessary for achieving accurate sentiment analysis results. Removing any of these concepts would lead to a decrease in sentiment classification accuracy. Therefore, it is evident that a comprehensive approach utilizing all or more the relevant concepts is essential for obtaining reliable and meaningful sentiment analysis outcomes in the context of e-wallet companies. This theory highlights the need for ongoing research and innovation in sentiment analysis methods that consider other than customer ratings within the e-wallet industry. From a practical perspective, the findings of this research provide actionable insights for e-wallet companies. The comprehensive understanding of customer sentiment dynamics over time allows these companies to make informed decisions and take appropriate actions to enhance their services. By addressing the sentiment patterns identified through this research, e-wallet companies can improve user experiences, increase customer satisfaction, and ultimately gain a competitive edge.

To improve customer satisfaction, e-wallet companies can consider implementing the strategies identified by the study conducted by Davidavičienė et al. (2022). Firstly, enhancing convenience and user experience by streamlining the mobile app interface and ensuring smooth transactions will contribute to a positive customer experience. Secondly, improving communication channels, such as responding promptly to user messages and sharing interesting and relevant information, can foster better engagement and satisfaction. Encouraging consumer reactions and engagement through shareable content and incentivizing positive recommendations can generate positive sentiments. Additionally, providing value-added services beyond basic transactions, such as personalized financial advice or exclusive discounts, can enhance customer satisfaction. Lastly, continuously monitoring customer sentiment and promptly addressing any concerns or issues through feedback channels will demonstrate a commitment to customer satisfaction. E-wallet companies can improve customer satisfaction, address further sentiment issues, and strengthen customer loyalty by adopting these strategies.

5. Conclusion

In conclusion, the main findings of this study indicate that e-wallet companies generally receive positive sentiments from customers, as evidenced by their high rating accuracies for positive sentiments. However, the predicted negative sentiments are not accurate, as reflected in the lower rating accuracies for negative sentiments across all e-wallet companies. The rating accuracies for neutral sentiments were also generally low, except for Alipay in 2019, which demonstrated the highest rating accuracy in capturing customer-neutral sentiments. These findings hold important implications for theory and practice in the e-wallet industry. Theoretical implications highlight the necessity of employing comprehensive data pre-processing techniques in sentiment analysis for accurate classification results. It establishes the importance of considering multiple aspects such as data standardization, tokenization, stop word removal, and lemmatization to achieve accurate sentiment analysis outcomes. From a practical perspective, the findings provide valuable insights for e-wallet companies to enhance their services based on customer sentiments. By addressing the sentiment patterns, e-wallet companies can improve user experiences, increase customer satisfaction, and gain a competitive advantage in the market. The study highlights the significance of understanding and leveraging customer sentiment dynamics to drive business success continuously in the e-wallet industry.

The research objectives are to conduct both polarity analysis and rating accuracy evaluation, providing insights into sentiment polarity and assessing the sentiment analysis model's performance based on customer inputs. By achieving these research objectives, the understanding of customer sentiment can be enhanced and e-wallet services can be continuously improved based on user reviews collected from the Google Play store. This study achieved three key contributions in sentiment analysis for e-wallet companies based on customer inputs. Firstly, it seeks to collect an extensive and diverse dataset encompassing 28 datasets from 7 e-wallet companies over the last four years, enabling a comprehensive understanding of how customers' sentiment towards e-wallets has evolved. Secondly, it intends to implement four crucial data pre-processing techniques, including data standardization, tokenization, stop word removal, and lemmatization, to ensure meaningful analysis, and to improve sentiment classification accuracy. These techniques are pivotal in transforming raw customer review data into a suitable format for sentiment analysis. Thirdly, the study conducted both polarity analysis and rating accuracy analysis, providing insights into polarity analysis and evaluating the performance of the sentiment analysis model based on customer inputs.

The limitation of this study is that the sentiment analysis performed in this study relied on the TextBlob library, which utilizes pre-trained models and lexicons. While TextBlob is a widely used and effective tool, it may not capture customer sentiments in-depth. Based on this identified limitation, a suggest avenue for future research could involve exploring the use of deep learning-based sentiment analysis models for in-depth customer sentiments. Another suggest avenue for future research, the dataset can be expanded by covering a larger period, including more older years. It would provide a more comprehensive understanding of the evolving sentiment towards e-wallets over time. Additionally, analyzing sentiments across a wider range of e-wallet companies and conducting a comparative analysis based on their payment methods would further enhance the insights gained from this research. These future plans will provide valuable insights for researchers, industry practitioners, and policymakers in helping them to understand the sentiment towards e-wallets better and make more informed financial decisions.

References

Abdul Hanan, F., Mutalib, S., Mohd Yunus, A., Abdul Rashid, M. F., Kamarudin, S. N., & Abdul Rahman, S. (2023). A Study on Social Media Responses on Road Infrastructure using Sentiment Analysis. Journal of Logistics, Informatics and Service Science, 10(2), 1-14. doi:10.33168/JLISS.2023.0201

Aljedaani, W., Rustam, F., Mkaouer, M. W., Ghallab, A., Rupapara, V., Washington, P. B., … Ashraf, I. (2022). Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry. Knowledge-Based Systems, 255, 109780. doi:10.1016/j.knosys.2022.109780

Almalis, I., Kouloumpris, E., & Vlahavas, I. (2022). Sector-level sentiment analysis with deep learning. Knowledge-Based Systems, 258, 109954. doi:10.1016/j.knosys.2022.109954

Davidavičienė, V., & Davidavičius, S. (2022). Impact of Activities in Social Networks on Customer Loyalty. Journal of Logistics, Informatics and Service Science, 9(2), 183-195. doi:10.33168/LISS.2022.0211

Hermansyah, R., & Sarno, R. (2020). Sentiment Analysis about Product and Service Evaluation of PT Telekomunikasi Indonesia Tbk from Tweets Using TextBlob, Naive Bayes & K-NN Method. 2020 International Seminar on Application for Technology of Information and Communication (ISemantic), 511–516. doi:10.1109/iSemantic50169.2020.9234238

Ikoro, V., Sharmina, M., Malik, K., & Batista-Navarro, R. (2018). Analyzing Sentiments Expressed on Twitter by UK Energy Company Consumers. 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), 95–98. doi:10.1109/SNAMS.2018.8554619

Iwendi, C., Mohan, S., Khan, S., Ibeke, E., Ahmadian, A., & Ciano, T. (2022). Covid-19 fake news sentiment analysis. Computers and Electrical Engineering, 101, 107967. doi:10.1016/j.compeleceng.2022.107967

Jagadishwari, V., & Shobha, N. (2022). Sentiment analysis of Covid 19 Vaccines using Twitter Data. 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), 1121–1125. doi:10.1109/ICAIS53314.2022.9742995

Khan, R., Rustam, F., Kanwal, K., Mehmood, A., & Choi, G. S. (2021). US Based COVID-19 Tweets Sentiment Analysis Using TextBlob and Supervised Machine Learning Algorithms. 2021 International Conference on Artificial Intelligence (ICAI), 1–8. doi:10.1109/ICAI52203.2021.9445207

Kristiyanti, D. A., Putri, D. A., Indrayuni, E., Nurhadi, A., & Umam, A. H. (2020). E-Wallet Sentiment Analysis Using Naïve Bayes and Support Vector Machine Algorithm. Journal of Physics: Conference Series, 1641(1), 012079. doi:10.1088/1742-6596/1641/1/012079

Kusrini, & Mashuri, M. (2019, March). Sentiment Analysis In Twitter Using Lexicon Based and Polarity Multiplication. 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT), 365–368. doi:10.1109/ICAIIT.2019.8834477

Leelawat, N., Jariyapongpaiboon, S., Promjun, A., Boonyarak, S., Saengtabtim, K., Laosunthara, A., … Tang, J. (2022). Twitter data sentiment analysis of tourism in Thailand during the COVID-19 pandemic using machine learning. Heliyon, 8(10). doi:10.1016/j.heliyon.2022.e10894

Maindola, P., Singhal, N., & Dubey, A. D. (2018, January). Sentiment Analysis of Digital Wallets and UPI Systems in India Post Demonetization Using IBM Watson. 2018 International Conference on Computer Communication and Informatics (ICCCI), 1–6. doi:10.1109/ICCCI.2018.8441441

Mansouri, N., Soui, M., Alhassan, I., & Abed, M. (2022). TextBlob and BiLSTM for Sentiment analysis toward COVID-19 vaccines. 2022 7th International Conference on Data Science and Machine Learning Applications (CDMA), 73–78. doi:10.1109/CDMA54072.2022.00017

Nair, A. J., G, V., & Vinayak, A. (2021). Comparative study of Twitter Sentiment On COVID - 19 Tweets. 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 1773–1778. doi:10.1109/ICCMC51019.2021.9418320

Poornima, A., & Priya, K. S. (2020). A Comparative Sentiment Analysis Of Sentence Embedding Using Machine Learning Techniques. 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 493–496. doi:10.1109/ICACCS48705.2020.9074312

Punetha, N., & Jain, G. (2023). Bayesian game model based unsupervised sentiment analysis of product reviews. Expert Systems with Applications, 214, 119128. doi:10.1016/j.eswa.2022.119128

Rahmanti, A. R., Chien, C.-H., Nursetyo, A. A., Husnayain, A., Wiratama, B. S., Fuad, A., … Li, Y.-C. J. (2022). Social media sentiment analysis to monitor the performance of vaccination coverage during the early phase of the national COVID-19 vaccine rollout. Computer Methods and Programs in Biomedicine, 221, 106838. doi:10.1016/j.cmpb.2022.106838

Sangeetha, J., & Kumaran, U. (2023). A hybrid optimization algorithm using BiLSTM structure for sentiment analysis. Measurement: Sensors, 25, 100619. doi:10.1016/j.measen.2022.100619

Smetanin, S., & Komarov, M. (2019). Sentiment Analysis of Product Reviews in Russian using Convolutional Neural Networks. 2019 IEEE 21st Conference on Business Informatics (CBI), 01, 482–486. doi:10.1109/CBI.2019.00062

Srivats Athindran, N., Manikandaraj, S., & Kamaleshwar, R. (2018). Comparative Analysis of Customer Sentiments on Competing Brands using Hybrid Model Approach. 2018 3rd International Conference on Inventive Computation Technologies (ICICT), 348–353. doi:10.1109/ICICT43934.2018.9034283

Suddle, M. K., & Bashir, M. (2022). Metaheuristics based long short term memory optimization for sentiment analysis. Applied Soft Computing, 131, 109794. doi:10.1016/j.asoc.2022.109794

Talahaturuson, E., Gumelar, A. B., Gabriel Sooai, A., Sueb, S., Suprihatien, S., Altway, H. A., … Sulistiyani, S. (2022). Exploring Indonesian Netizen's Emotional Behavior Through Investment Sentiment Analysis Using TextBlob-NLTK (Natural Language Toolkit). 2022 International Seminar on Application for Technology of Information and Communication (ISemantic), 244–248. doi:10.1109/iSemantic55962.2022.9920431

Wang, L., Niu, J., & Yu, S. (2020). SentiDiff: Combining Textual Information and Sentiment Diffusion Patterns for Twitter Sentiment Analysis. IEEE Transactions on Knowledge and Data Engineering, 32(10), 2026–2039. doi:10.1109/TKDE.2019.2913641

Wongkar, M., & Angdresey, A. (2019, October). Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter. 2019 Fourth International Conference on Informatics and Computing (ICIC), 1–5. doi:10.1109/ICIC47613.2019.8985884

Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning. IEEE Access, 8, 23522–23530. doi:10.1109/ACCESS.2020.2969854