Available online at
https://jcrinn.com/
https://crinn.conferencehunter.com/

Journal of Computing Research and Innovation 9(1) 2024

Journal of Computing Research and Innovation

Review of the Lazada application on Google Play Store: Sentiment Analysis

Amyra Nur Ain Zulkifli[1], Nur Amalina Shafie[2*]

[1,2] *College of Computing, Informatics and Mathematics, Universiti Teknologi MARA Negeri Sembilan, Seremban Campus, Negeri Sembilan, Malaysia*

ARTICLE INFO

Article history:
Received: 28 January 2024
Revised: 1 February 2024
Accepted: 9 February 2024
Online first: 1 March 2024
Published 1 March 2024

Keywords:
Sentiment Analysis
LAZADA
Google Playstore
Natural Language Processing

DOI:
10.24191/jcrinn.v9i1.412

ABSTRACT

Sentiment analysis is a technique for gaining meaningful insight from unstructured and unorganized textual content from multiple platforms. It is a Natural Language Processing (NLP) method that may categorize data or reviews as positive, negative, or neutral. Analyzing reviews on the internet could yield helpful, actionable insights that could be economically beneficial to vendors or other interested parties. There are various online shopping platforms due to customer demand and rely on reviews and ratings while selecting choices. However, it could be difficult to tell whether the reviews are positive, negative, or both. The objective of this research is to classify the reviews on Lazada which is one of the online shopping platforms as positive, neutral or negative sentiments and to examine the words used most frequently in Lazada users' reviews on the Google Play Store. This research used data from 7267 reviews that were extracted from the Google Play Store between 2019 and 2022 using the Google-play-scraper Python script. The reviews have been analyzed using the Valence Aware Dictionary and Sentiment Reasoner (VADER)'s to determine whether they are positive, neutral, or negative. The results indicate that 4229 reviews are positive. There are about 2857 negative sentiments and 181 neutral sentiments. It demonstrates that more people are happy using the Lazada app between 2019 and 2022. The results also demonstrate that sentiment analysis is an effective tool for categorising and evaluating other people's reviews and feedback.

1. INTRODUCTION

Sentiment analysis, often known as opinion mining, is a study field that examines people's feelings or views regarding subjects, events, personalities, situations, services, goods, organizations, and their features (Naseem et al., 2021). It is a technique of gaining meaningful insight from unstructured and unorganized textual contents from various social platforms and online sources, such as chats on social platforms such as Twitter, WhatsApp, and Facebook, as well as online blogs and comments (Hossain & Rahman, 2022). Sentiment analysis is a method of Natural Language Processing (NLP) that categorizes data or reviews as positive, negative, or neutral. It can also be used to determine the text's emotional tone. It classifies

[2*] Corresponding author. *E-mail address*: amalina@uitm.edu.my

particular feelings like joy, annoyance, shock, anger, and sadness. Many businesses use it to analyze brand and product reviews and to better understand customer feedback (Naseem et al., 2021). It will classify text into predefined sentiments using a variety of algorithms. If the positive incidence outnumbers the bad, the data will contain positive feelings, and vice versa (Dudhankar et al., 2022).

Sentiment analysis is frequently used in business processes to enhance decision-making and customer satisfaction (Afifah et al., 2021). Additionally, posting reviews online has become an increasingly common way for consumers to share their feelings and thoughts about the things they have purchased or the services they have received. Analyzing the huge worth of online reviews could result in useful actionable knowledge that may be economically beneficial to vendors and other interested parties (Dudhankar et al., 2022).

The review is one example of unstructured and disorganized data that can be analyzed by using sentiment analysis. User reviews have rich information when a user is using the apps. Usually, user reviews have two essential parts: rating and opinion. The rating indicates the overall evaluation of the user experience using a numerical scale. While opinion is a deeper story about the user's experience when using the apps (Nugroho et al., 2021). It can figure out user satisfaction from this information including the ratings they gave. User satisfaction plays an important role in online applications because it can affect both existing users and potential new users. Satisfaction can be one of the reasons for the user to continue or discontinue using the application. Previous empirical studies, it has shown that application store reviews include information that is useful to analysts and application designers, such as user requirements, bug reports, features requests and documentation of user experiences with specific application features. All of this feedback can represent a "voice of the users" and be able to drive the development effort and improve forthcoming releases (Naseem et al., 2021).

In the app's development process, the usability factor is the most crucial aspect because it can reflect the level of ease apps when used. Usability evaluation always involves users, one of them is user reviews (Nugroho et al., 2021). Google Play Store and Apple Store both have sections for user reviews and ratings that can help to maintain the applications. Therefore, the management as well as new users can refer to that reviews section to discover more about the capabilities of the applications from the reviews submitted by the users. Furthermore, it can assist new users in deciding whether they should download the applications or not. Also, management can use user feedback to address any issues raised by users to make improvements. Lazada is among Malaysia's most-used e-commerce sites in 2022. The app is available for download from the Google Play Store or the Apple App Store and requires Android 4.4 or higher for Android devices and iOS 11.0 or higher for Apple devices. Additionally, it can perform a simple Google search and log into an account. According to information in Google Play Store, Lazada was launched on June 8, 2013, and has been downloaded more than 100,000,000 times worldwide. Lazada includes electronic devices and accessories, TVs and household appliances, health and cosmetic products, groceries and pets, clothing and accessories for adults, children, and even automobiles and motorcycles. The most recent version, 7.12.0, was updated on October 23, 2022, which shows how regularly the developers are upgrading the apps and how widely people are used.

Ranjan and Mishra (2020) stated that due to many demands from customers, there are many online shopping platforms that have been developed. There are numerous platforms available, including Lazada, Shopee, Shein, and even the brand's own website. However, because it is an online platform, users' feedback and ratings must be relied upon while making decisions. It can be difficult and complicated to tell whether user evaluations or comments are positive or negative. And sometimes, the feedback includes both good and poor aspects. Lazada is a mobile app that facilitates online buying and selling and is simple to access. This enables users to conduct online shopping activities that are accessible from anywhere at any time. Before designing a mobile application, developers should constantly consider numerous criteria to ensure that the application is fit for the purpose. Personalization, data entry models, and navigation

https://doi.org/10.24191/jcrinn.v9i1

flexibility should all be considered in order to provide users with happiness. Furthermore, apps that deal well with information overload should be examined, as it is an essential element that considerably bothers online customers nowadays. According to users, usability simply refers to the application's ease of use, efficiency, effectiveness, and user pleasure (Ninyikiriza et al., 2020).

Therefore, to assist possible new users and management, the study tends to see the sentiment of the community, particularly Lazada users, and there needs to be an analysis approach that can summarize the reviews as positive, negative, or neutral. The difference between this study and earlier studies is that past studies used data from Twitter, whereas this study uses reviews from the Google Play Store. Hence, this research is classifying the reviews on Lazada to positive, neutral or negative sentiments and examine the words used most frequently in Lazada users' reviews on the Google Play Store.

2. METHODOLOGY

The data for this study was derived from reviews posted by users on the Lazada app in the Google Play store. It primarily used text reviews from 2019 until 2022. This research used a web scraping approach to collect data, and a software tool called Google-play-scraper package in Python. The extracted Lazada user reviews on the Google Play Store are in the form of sentences or paragraphs. Data processing or sub-processing is a necessary step before the data can be used for sentiment analysis, thus it must be done. Tokenizations, stop-word removal, stemming, nominal to text, transformed cases and generating n-grams are all a part of data processing. Furthermore, this procedure can be assisted by reducing the number of words, having correctly matched stems, and saving time and memory space. After this stage has been completed, sentiment analysis can take place in RapidMiner. Fig. 1 shows the flowchart to better understand the process of sentiment analysis conducted.

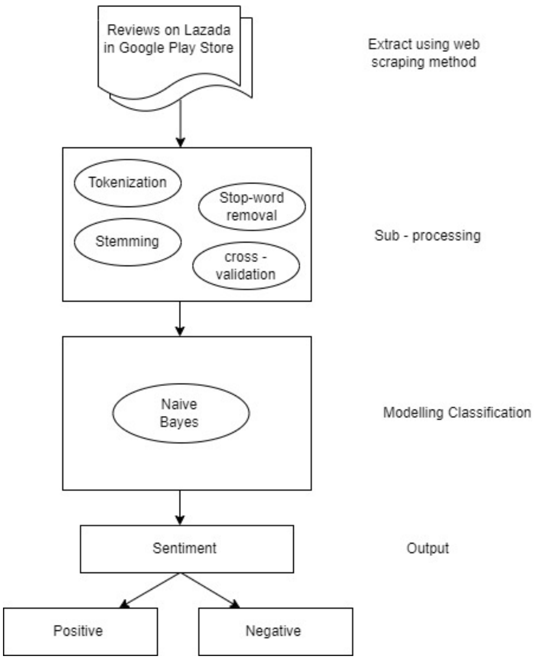

Fig. 1. Flowchart of Sentiment Analysis Technique

https://doi.org/10.24191/jcrinn.v9i1

2.1 Data Processing

2.1.1 Web scraping

Web data extraction, also referred to as harvesting or web scraping, is used to extract information from websites. Unstructured data is converted into structured data in this way. Web scraping, according to previous research, is a method for automatically extracting data from diverse web documents. Additionally, it might be saved in the local system in an organized format for future study. Web scraping tools can now be used to convert web data into structured information such as spreadsheets.

This research included the web scraping process for the website of Play Store to get reviews from Lazada users. Google-play-scraper would be used to scrape the web. The Python package Google-Play-scraper is useful for web scraping data from the Google Play Store, such as app reviews, ratings, and metadata.

The package can be installed in Python using pip or a similar package manager. Then it used the import statement to import the package into the Python script. After finishing, the search query would be specified for the data that needs to be scrapped. This could be the categories, search terms, or other interesting data. The data from the Google Play Store was then scraped using the package. The search query is entered into the function along with additional parameters like the number of reviews to scrape or the results' sorting order. The data can also be exported into other forms, including CSV or Excel.

2.1.2 Review Sub-Processing

Tokenization is the first sub-processing that the data must go through. Tokenization is the process of splitting up or segmenting a paragraph into sentences and words. The strings are separated into discrete words by removing spaces from each word and added as a single array. Individual words or sentences can be broken down into smaller units inside a text document or paragraph. Tokens are the name of each of these objects. "I love using this app", for example, is broken down into little units like ["I", "love", "using", "this", "app"]. Once the review has completed tokenization, stop-word removal must be performed to get rid of any terms that were either overlooked or did not add anything to the analysis. The idea is to remove the most common words across all the documents. Stopwords can be articles, prepositions, pronouns, or conjunctions. The, a, an, with, and other words are examples of words that can be eliminated.

Stemming is a method for determining the root or stem of a word. It is a procedure that removes all affixes from a word, including those that appear before and after the term. It removed all affixes from each word, reducing it to its root word. The word "connect" is the root of several other terms, such as connect, connected, connecting, and connections. The Nominal Text operator transformed all nominal characteristics into string attributes. Each nominal value is simply utilized as the new attribute's string value. The new value would likewise be missing if the nominal attribute's value is missing. The transform cases operator changed all the characters in a document to either lowercase letters or capital letters, as appropriate. The Generate n-Grams (Terms) operator created the term n-Grams of tokens in a document. The term n-Gram is defined as a series of consecutive tokens of length n. The term n-Grams generated by this operator consists of all series of consecutive tokens of length n.

2.2 Sentiment Analysis

After the review's sub-processing is completed, sentiment analysis can be performed. The machine learning classifier is performed in this stage to classify the sentiment of reviews as positive, negative, or neutral.

https://doi.org/10.24191/jcrinn.v9i1

This study's research tool was RapidMiner. RapidMiner is a programme used to carry out data mining workflows for a variety of activities, including distinct data mining applications and optimisation strategies. RapidMiner's outstanding ability to programme the execution of complicated workflows within a visual user interface, without the need for traditional programming skills, is one of its key features. The data is ready to be sentiment after sub-processing is completed. The "Extract Sentiment" operator toolbox in RapidMiner would be used with the "Vader" model. The text is scored using the Valence Aware Dictionary and Sentiment Reasoner (VADER) lexicon and rule-based sentiment. VADER generates scores based on a dictionary of words and is especially sensitive to the attitudes expressed in social media.

VADER can be applied directly to unlabelled text data. Its sentimental analysis is based on a lexicon that converts lexical features into sentiment ratings, which measure the intensity of emotion. The polarity-based method is used to classify the text as either positive or negative. It ranges from -1 to +1, with -1 indicating severely negative sentiment and +1 indicating extremely positive sentiment. The scoring string is then added together to produce the final score. If the score is more than zero, the specific review is considered a positive sentiment. The review is a negative sentiment if the overall score is less than 0. The review would be categorized as neutral if it received a score of 0. The number of used and unused tokens, the total of the positive and negative components, and a nominal attribute with all words contributing to scoring are all available if tick the advanced output option. Lastly, data visualization involved putting information into a visual context, like a graph or map, to help others understand it and draw conclusions. Data visualization's main objective is to make it simpler to spot patterns, trends, and outliers in a large dataset. One of the most important processes in the data analysis process is data visualization, which requires that after the data has been gathered, processed, and modelled, it would be visualized to conclude. In this study, the graph and word cloud were produced using Tableau data visualization software.

3. RESULT AND DISCUSSION

3.1 Web Scraping

The technique to use to obtained the data was web data scraping. The Python script can run because the Google-play-scraper on the Python package has already been installed. The whole coding is included in the appendices. 10,000 reviews of Lazada apps in the Google Play store were scraped with variables User, Text, Date and Score. The score is the star or rating that users assign to reviews. Some data cleaning is required before importing the data into RapidMiner for analysis. Fig. 2 depicts the summary in the flowchart to obtain the final data. The reviews are filtered out only from 2019 to 2022. After filtering by year, 72.67%, or 7267 reviews, can be used as data for analysis. Then, remove the variables user and date because they are not relevant to the analysis.

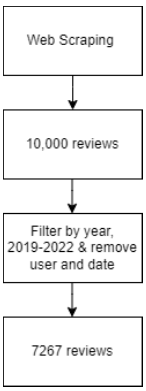

Fig. 2. Flowchart to obtain Final Data

https://doi.org/10.24191/jcrinn.v9i1

Table 1. 3 out of 7267 Final Data

Text	Star
My fourth time using this service. Very convenient but the app is kind of confusing: I have to add a product to my wishlist before I can proceed to check out. I hope the app can be more user friendly.	5
when it comes to paying with a debit card, the otp is not appearing, which makes the transaction unsuccessful. i tried again for three times because i thought it was a problem with the internet connection but its really the apps problem. please fix. VERY INCONVENIENT!	1
It is hard to navigate in the app there are too many redirection. There are so many clickbait products, you should take action on them. It is very confusing to use. This app still needs a lot of improvements.	1

Based on Table 1, 3 out of 7267 reviews and the star rating given by Lazada users were shown. The first review received a 5-star rating, while the other two received a 1-star rating from the reviewer itself.

3.2 Sentiment Analysis

The Valence Aware Dictionary and Sentiment Reasoner (VADER) is a vocabulary and rule-based sentiment that has been used to find the sentiment in the reviews. The extract sentiment operator parameter in RapidMiner must be set to model Vader. A word dictionary is used by VADER to generate the scores. The scores were produced automatically by the VADER using a scale of -1 to +1. Table 2 and Table 3 displays the top 5 positive and negative words with their generated scores for this analysis.

Table 2. Top 5 Positive Words and their generated score

Words	Score
Good	0.49
Easi	0.49
Great	0.79
Free	0.59
thank	0.49

Based on Table 2, VADER had assigned an artificially generated score to each of the words. The top five positive words and their generated scores are shown above. The values for good, easy, and thank were all 0.49. Meanwhile, the word great has a score of 0.79 while the word free has a score of 0.59.

Table 3. Top 5 Negative Words and their generated score

Words	Score
Annoi	-0.44
Cancel	-0.26
Problem	-0.44
Bad	-0.64

https://doi.org/10.24191/jcrinn.v9i1

worst	-0.79

Based on Table 3, VADER had assigned an artificially generated score to each of the words. The top five negative words and their generated scores are shown above. The annoi and issue values were both -0.44. Meanwhile, the term cancel has a -0.26 score, bad has a -0.64 score, and worst has a -0.79 score.

Table 4. Result of Sentiment Analysis for Three Reviews

No	Score	Scoring String	Negativity	Positivity	Sentiment
1	1.4	Kind (0.62) Confusing (-0.23) Hope (0.49) Friendly (0.56)	0.2	1.7	Positive
2	-1.5	Unsuccessful (-0.38) Problem (-0.44) Please (0.08) Inconvenient (-0.36)	1.6	0.1	Negative
3	0.0	Hard (-0.10) Confusing (-0.23) Improvements (0.33)	0.3	0.3	Normal

Table 4 demonstrates the reviews score, scoring string, negativity, positivity, and sentiment for each review for 3 statements reviews out of 7267 in total. The first review was labelled as positive sentiment because the sum of the scoring string, 0.62 + (-0.23) + 0.49 + 0.56 = 1.4, indicates a positive value, therefore, it is a positive sentiment. The positivity of the review is equal to 1.7 while the negativity is equal to 0.2. The second review, on the other hand, has a sum of scoring string of -1.5 as the negativity equals 1.6 and the positive equals 0.1. As a result, it is classed as a negative sentiment. Finally, the third review was classed as neutral when the sum of -0.01, -0.23, and 0.33 equalled 0. The values for negativity and positivity are both 0.3.

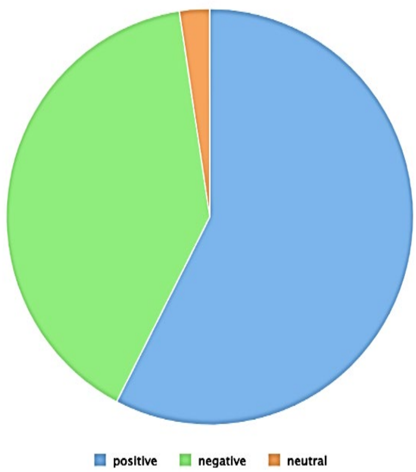

Fig. 3. Pie Chart of Polarity of Sentiment

Fig. 3 displays that the positive (blue) ratings outnumber negative and neutral ones by 58.2%. Negative (green) sentiments accounted for 39.3%, while neutral (orange) sentiments accounted for 2.5%.

https://doi.org/10.24191/jcrinn.v9i1

Table 5. Statistics of Sentiment of Reviews

Nominal value	Absolute count	Fraction
Positive	4229	0.582
Negative	2857	0.393
Neutral	181	0.025

According to Table 5, positive reviews account for 4229 of the 7267 reviews on the Google Play Store, for a fraction of 0.582. There are 2857 negative reviews, whereas just 181 (or 0.025) of the reviews are classified as neutral. It demonstrates that more people are happy using the Lazada app.

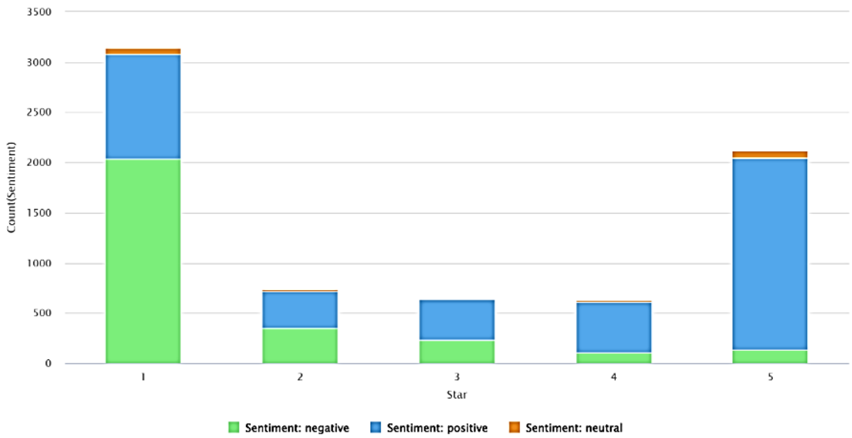

Fig. 4. Stacked Bar Chart of Sentiment by Star

Fig. 4 reveals that there are 2037 negative reviews, 1038 positive reviews, and 65 neutral reviews in 1-star. Meanwhile, 1915 of 2115 5-star reviews are positive, 132 are negative, and 68 are neutral. This finding demonstrates that the previous research by (Sadiq et al., 2021) was correct about the high star rating does not always correlate with positive reviews.

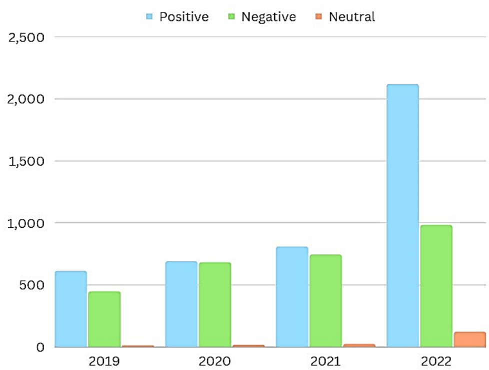

Fig. 5. Bar Chart of Sentiment by Year

Fig. 5 shows the sentiment bar chart by year. The graph illustrated the increasing trend in all three sentiments year after year. Positive sentiment increased from 612 to 690 to 808, then to 2119. The negatives

https://doi.org/10.24191/jcrinn.v9i1

are 448, 681, 744, and 984. The neutral reviews are 14, 17, 26 and 124. These trends can provide significant information to the Lazada application's developers and operators, assisting them in understanding changing user opinions and experiences and identifying areas for development.

3.3 Frequent Word and Word Cloud

The research was continued by identifying the most frequently occurring words in the reviews and visualizing them in a word cloud. The top 100 most used words are displayed in the table below. Frequent words are separated into positive and negative sentiments in accordance.

Table 6. Frequency of the Positive Words

Word	In document	Total
good	1062	1292
easi	742	842
great	489	559
free	427	536
thank	489	526
pleas	467	520
help	356	409
love	354	405
want	344	392
hope	362	389

According to Table 6, people used the word "good" 1292 times in 1062 reviews, making it the word most frequently associated with positive opinions. 842 times in 742 reviews helped "easi" move up to second place. There were 559, 536, and 526 mentions of great, free, and thank, respectively.

Fig. 6. Word Cloud of Positive Words

Fig. 6 displays the word cloud of positive sentiments with the word "good" as the bigger word in the word cloud. This indicated that "good" was the frequent word mentioned by the users.

https://doi.org/10.24191/jcrinn.v9i1

Table 7. Frequency of the Negative Words

Word	In document	Total
annoi	482	588
cancel	292	392
problem	313	378
bad	327	369
worst	255	289
wrong	225	261
fail	176	221
disappoint	181	195
stop	166	191
scam	152	180

Table 7 shows that the word "annoying" is the one most commonly relating to negative feelings, appearing 588 times in 482 evaluations. With 392 mentioned, "cancel" climbed into the second position. Problems, bad and worst were mentioned 378, 369, and 289 times, respectively.

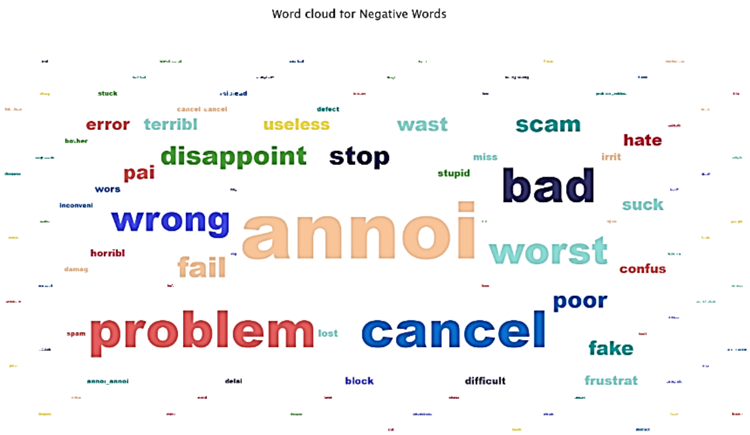

Fig. 7. Word Cloud of Negative Words

Fig. 7 displayed the word cloud of negative sentiments with the word "annoi" being the bigger word in the word cloud. This indicated that "annoi" was the frequent word mentioned by the users.

In summary, 7267 reviews from Lazada between 2019 and 2022 have been categorized as neutral, negative, or positive in the Google Play Store. The results showed that 4229 reviews were good, more than there were negative or indifferent evaluations. There were 2857 negative reviews and 181 neutral reviews. These numbers show that there are more positive evaluations of Lazada than negative ones. The word cloud shows that certain words appear in the evaluations more frequently when they are higher in size. The findings indicate that Lazada users expressed the terms "good" and "annoying" with increasing frequency.

4. CONCLUSION

Sentiment analysis is a technique for gaining meaningful insight from unstructured and unorganized textual content from multiple platforms. Sentiment analysis is an NLP method that may categorize data or reviews

https://doi.org/10.24191/jcrinn.v9i1

as positive, negative, or neutral. Analyzing reviews on the internet could yield helpful actionable insights that could be economically beneficial to vendors or other interested parties. However, it might be challenging to determine whether the reviews are positive, negative, or both. The primary objective of this study was to classify the reviews on Lazada in the Google Play Store. Based on sentiment analysis using VADER reveals that 4229 of the 7267 reviews were positive, 2857 were negative, and 181 were neutral. The opinions expressed in each rating (star) demonstrated that they do not always offer honest reviews, which was supported by a previous study. Users have given a 1-star rating of 1038 positive reviews and a 5-star rating of 132 negative reviews. Thus, 4229, 2857, and 181 of the reviews, which were good, negative, and neutral correspondingly, achieved the first objective. Based on the outcome of objective 2, the word cloud demonstrates that “good” and “annoying” are the most frequently used words in comments for positive and negative sentiments, respectively. As a result, the words “good” and “annoying” are larger to show how frequently users comment on these terms. 1288 times the word “good” and 138 times the word “annoying” were commonly used. The outcomes of this research can guide app developers in spotting and realizing crises that they haven’t noticed. The results of the sentiment analysis can assist them in discovering exactly why users are unhappy with it so that the issues can be fixed and specifically acknowledged. Information from this research can be beneficial for app developers and users. It may also provide new users with guidance about how to download Lazada and assist in providing a general perspective of how applicable applications look. As a side benefit, eCommerce developers could also provide customers with better degrees of satisfaction and improve their job performance.

5. ACKNOWLEDGEMENTS

We acknowledge the grant provided by Universiti Teknologi MARA 600-RMC/GPM LPHD 5/3 (136/2021).

6. CONFLICT OF INTEREST STATEMENT

The authors agree that this research was conducted in the absence of any self-benefits, commercial or financial conflicts and declare the absence of conflicting interests with the funders.

7. REFERENCES

Afifah, K., Yulita, I. N., & Sarathan, I. (2021). Sentiment analysis on telemedicine app reviews using xgboost classifier. *2021 International Conference on Artificial Intelligence and Big Data Analytics* (pp. 22-27). doi: 10.1109/ICAIBDA53487.2021.9689762

Annie, A. (2021). *The state of mobile 2021*. Retrieved from https://www.appannie.com/en/go/state-of-mobile-2021/

Apptentive. (2019). Mobile app benchmarks: The average ratings, reviews, and retention rates. *Apptentive Blog*. Retrieved 29 March, 2023, from https://www.apptentive.com/blog/2019/03/11/mobile-app-benchmarks-the-average-ratings-reviews-and-retention-rates/

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer New York.

Bush, T. (2020, Jun). Descriptive analysis: How-to, types, examples. *Pestle Analysis*. Retrieved from https://pestleanalysis.com/descriptive-analysis/

Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems*, *45*(4), 1007–1016.

https://doi.org/10.24191/jcrinn.v9i1

Dudhankar, V., Sen, N., Langde, A., & Kupade, V. (2022, June). Google playstore review sentiment analysis. *International Research Journal of Modernization in Engineering Technology and Science*, *4*(06), 3901-3911. https://www.irjmets.com/uploadedfiles/paper/issue_6_june_2022/26972/final/fin_irjmets1656231688.pdf

Hossain, M. S., & Rahman, M. F. (2022, Jan). Sentiment analysis and review rating prediction of the users of Bangladeshi shopping apps. *IGI Global*. Retrieved from https://www.igi-global.com/chapter/sentiment-analysis-and-review-rating-prediction-of-the-users-of-bangladeshi-shopping-apps/306094.

Hu, P., Li, Q., & Ye, Y. (2019). Customer review analysis using natural language processing techniques: a case study of e-commerce platforms. *Sustainability*, *11*(8), 2234.

iPrice Group. (2021). *iPrice insights: State of ecommerce in Southeast Asia 2021*. Retrieved from https://iprice.my/insights/mapofecommerce/en/

javaTpoint. (2021). *Machine learning decision tree classification algorithm - javatpoint*. Retrieved from https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm

Kanaris, I., Stamatatos, E., & Fakotakis, N. (2020). Tf-idf vs word2vec vs glove: An overview. *arXiv preprint*. arXiv:2010.02545

Kim, M.-G., & Yoon, Y.-J. (2020). Negative consequences of text classification: A critical review and practical remedies. *Journal of the Association for Information Science and Technology*, *71*(8), 936–947.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2015). Efficient estimation of word representations in vector space. *arXiv preprint*. arXiv:1301.3781.

Naseem, S., Mahmood, T., Asif, M., Rashid, J., Umair, M., & Shah, M. (2021). Survey on sentiment analysis of user reviews. *2021 International Conference on Innovative Computing (ICIC)* (pp.1-6). https://doi: 10.1109/ICIC53490.2021.9693029.

Ninyikiriza, D. L., Sourav, A. I., & Setyohadi, D. B. (2020). Increasing user satisfaction of mobile commerce using usability. *International Journal of Advanced Computer Science and Applications*, *11*(8).

Nugroho, K. S., Sukmadewa, A. Y., Wuswilahaken DW, H., Bachtiar, F. A., & Yudistira, N. (2021). Bert fine-tuning for sentiment analysis on Indonesian mobile apps reviews. In *6th International Conference on Sustainable Information Engineering and Technology 2021* (pp.258–264). https://doi.org/10.1145/3479645.3479679

Raj, A. (2020, Nov). Perfect recipe for classification using logistic regression. *Towards Data Science*. Retrieved from https://towardsdatascience.com/the-perfect-recipe-for-classification-using-logistic-regression-f8648e267592#:~:text=Logistic%20regression%20is%20a%20classification

Rajendran, S. (2021). Improving the performance of global courier & delivery services industry by analyzing the voice of customers and employees using text analytics. *International Journal of Logistics Research and Applications*, *24*(5), 473-493. https://doi.org/10.1080/13675567.2020.1769042 doi:10.1080/13675567.2020.1769042

Ranjan, S., & Mishra, S. (2020). Comparative sentiment analysis of app reviews. *In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT* (pp. 1-7). https://doi: 10.1109/ICCCNT49239.2020.9225348

https://doi.org/10.24191/jcrinn.v9i1

Similarweb. (2023, Jul). Top websites ranking most visited ecommerce shopping websites in Malaysia. *Similarweb LTD*. Retrieved from https://www.similarweb.com/top-websites/malaysia/e-commerce-andshopping/#:~:text=shopee.com.my%20ranked%20number,eCommerce%20%26%20Shopping%20websites%20in%20Malaysia.

Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, *28*(1), 11–21.

Sunil, R. (2019, Mar). *Understanding support vector machine algorithm from examples (along with code)*. Retrieved from https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/

Turney, P. D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Association for Computational Linguistics*. arXiv:cs/0212032

Wei, H. (2019, Mar). Nlp pipeline 101 with basic code example—feature extraction. *Voice Tech Podcast*. Retrieved from https://medium.com/voice-tech-podcast/nlp-pipeline-101-with-basic-code-example-feature-extraction-ea9894ed8daf#:~:text=Feature%20extraction%20step%20means%20to

Yee, A., Kee, D., Xing, C., Qian, P., Siew, M., & Dehrab, A. (2019, 10). Lazada group. *Journal of International Conference Proceedings*, *2*(No 2 (2019)), 19-29. https://doi: 10.32535/jicp.v2i2.599.