Strategy For Increasing the Brand Reputation of The PLN Mobile Application Based on Social Media Sentiment Analysis Using Machine Learning

Novi Ainur Riza[1,2*], Bagus Jati Santoso[3]

ABSTRACT

Social media is currently a trendy medium for people in Indonesia to express opinions. Users can easily express their experiences with a product through social media, including the PLN Mobile application from PT PLN (Persero). The application becomes a digital platform to meet various customer needs related to electricity services. Twitter social media has provided data on all tweets and reviews that can be accessed with certain keywords publicly through the Twitter API (Application Programming Interface). This research will use the text mining method with the word cloud approach, network explorer, types of emotions, and sentiment analysis, which can be used to analyze user opinions. The tools used are Orange Data Mining, which applies text preprocessing, including transformation, tokenization, normalization, and filtering, which aims to analyze text. The method used for sentiment analysis classification of Twitter users' opinions is VADER: Lexicon- and Rule-Based Sentiment Analysis. This method analyzes and classifies sentiment on social media towards the PLN Mobile application in Indonesia. The resulting classification model can be used as an early warning about how belief occurs on social media towards the PLN Mobile application. In addition, several recommendations on managerial aspects were also produced referring to the research data.

KEYWORDS: PLN Mobile Application, API Twitter, Sentiment Analysis, VADER

[1]HSSE, PT PLN (persero), Gresik, Indonesia

[2]Business Analytics, Interdisciplinary School of Management and Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

[3]Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

*Corresponding author: novies13@gmail.com

1. INTRODUCTION

PLN is Indonesia's largest electric energy company, providing electricity services to millions of customers nationwide. In the current digital era, PLN has launched the PLN Mobile application to provide easy access and information to customers(Murdifi, 2020). This application allows users to carry out various activities related to electricity services, such as paying bills, monitoring energy consumption, and filing complaints.

The PLN Mobile application is a digital application created by PT PLN (Persero) to provide electricity services through mobile apps. The PLN Mobile application has several online services, including paying electricity bills, buying tokens, recording independent meter numbers, adding power, complaining about disturbances and complaints, monitoring token purchases, monitoring postpaid electricity usage, billing notifications, blackout notifications, information on the progress of troubleshooting, to the maintenance of the power grid.

Based on this background, this research will carry out a sentimental analysis of the opinions of PT PLN (Persero) customers regarding the PLN Mobile application using machine learning. Gathering information on text mining is done using the word cloud approach, network explorer, types of emotions, and sentiment analysis. The tools used are Orange Data Mining (Demšar et al., 2013a) by applying text preprocessing, which includes transformation, tokenization, normalization, and filtering, which aims to analyze text. The method used for sentiment analysis classification of Twitter users' opinions is VADER: Lexicon- and Rule-Based Sentiment Analysis (Abimanyu et al., 2022; Amalia et al., 2022; Marwa & Kristanto, 2022).

In addition, this research can also provide benefits for PLN in identifying opportunities to improve the features and functionality of the PLN Mobile application based on positive feedback from users. By analyzing user sentiment, PLN can gain insight into which aspects of the application are most liked and useful and identify areas that need improvement or improvement.

2. LITERATURE REVIEW

The PLN Mobile application is the first application issued by PT PLN (Persero), launched on 31 October 2016, coinciding with the 71st National Electricity Day. The PLN Mobile application is a synergy of PLN services to customers after utilizing Contact Center 123, Facebook, Twitter, Email, and the PLN Web. Through the PLN Mobile application, customers can find information ranging from electricity bills, token transactions, electricity usage history, payment locations through the nearest banking, and the status or progress of requests and complaints. In addition, this application also provides information to customers regarding the employment status of PT PLN (Persero).

Text mining is part of data mining. Data mining works extensively for extracting information from many types of data. Other types of data mining are relational, web, and big data mining. Text mining is a type of data mining that specifically analyzes text. Text

mining for research has been widely used to analyze various text forms, such as sentiment analysis.

Sentiment Analysis or sentiment analysis is an intellectual process to define the feelings and emotions of users (Amalia et al., 2022; Amelia et al., 2022; Fang & Zhan, 2015; Kumar & Meera, 2022; Rohmah et al., 2021). Sentiment analysis is a field of Natural Language Processing (NLP). Massive changes to Internet-based application development have led to the personalization of large volumes of user reviews for various information.

VADER (Valence Aware Dictionary for Sentiment Reasoning), including a type of sentiment analysis based on a lexicon (a library) of sentiment-related words (Abimanyu et al., 2022; Marwa & Kristanto, 2022). In this approach, each word in the lexicon is scored as positive or negative, and in most cases, how positive or negative the number is.

3. METHODS

In the early stages, the researcher begins the research process by understanding the context and research objectives and planning the steps to be taken. This stage is the first step in research. The next stage is to make the background and formulate the problems of this research. This stage involves an in-depth understanding of the background of the research topic. The researcher should delve into comprehensive related literature to identify relevant research areas. Based on this understanding, the researcher formulates specific and researchable research problems. After understanding the background of the research and formulating the problem, the research objectives are determined. At this stage, the researcher sets clear and specific goals to be achieved through research. Research objectives must be closely related to the problems that have been formulated and help direct research steps.

The next stage is data collection and literature review. This stage involves collecting data that is relevant to the research objectives. Data can be obtained through surveys, interviews, observation, experiments, or secondary data collection. The researcher also conducted a careful literature review to understand better the previous research conducted in the field. Literature studies can be carried out by studying secondary data, namely from scientific journals sourced from various journals, both national and international journals, books, and internet literacy, whose publication age is no more than five years. The secondary data source obtained is literature still related to this research. Previous research is reviewed in the form of journal reviews to make it easier to find similarities in topics, theories, and differences and to examine previous studies' results as learning materials in conducting this research.

The last stage in this research is the preparation of recommendations and research reports. At this stage, the researcher makes recommendations based on the results of the research that has been analyzed. These recommendations contain suggestions or actions that can be taken based on research findings to solve problems or achieve research objectives. In addition, the researcher also compiled a research report consisting of

sections such as introduction, methodology, results, analysis, conclusions, and recommendations.

4. RESULTS

The analysis carried out includes a discussion of text data that has been collected in such a way using crawling data techniques. The data is taken from social media Twitter. Then, a predictive analysis was carried out, which produced a classification model to predict public sentiment regarding the PLN Mobile application. This model will classify a text on social media as whether it tends to have a positive, negative, or neutral sentiment. The following is a workflow on the Orange Data Mining application from the research results shown in Figure 1.
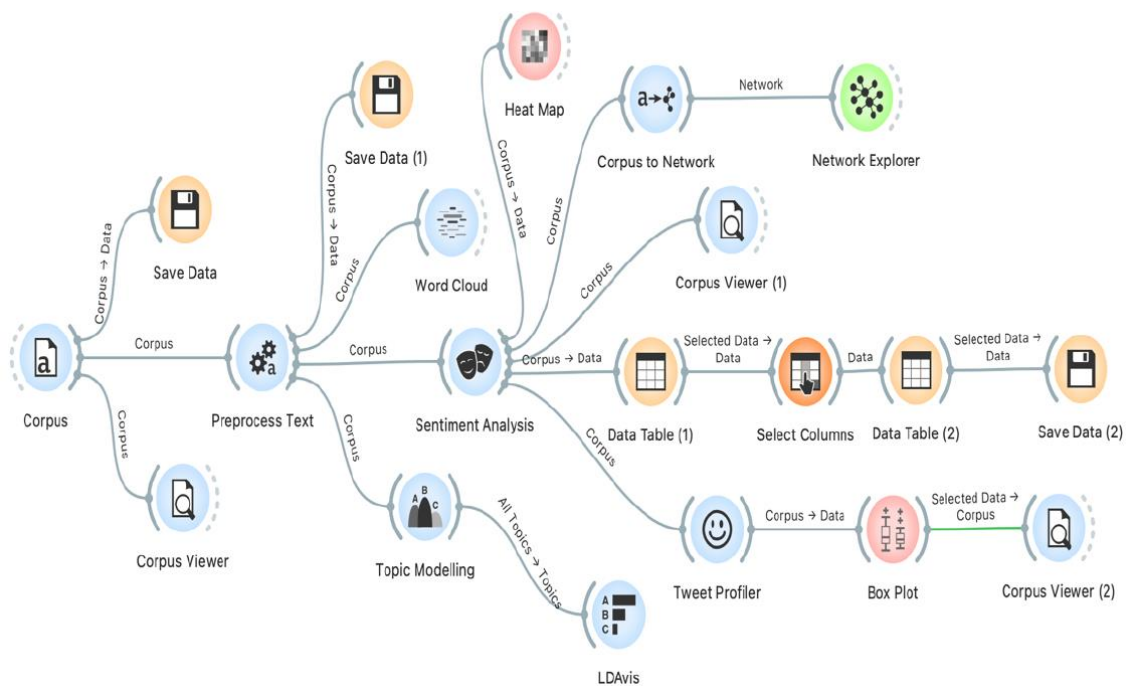
FIGURE 1. Workflow Design in Orange Data Mining(Demšar et al., 2013b)

In the early stages, Twitter text data was collected by adding a Twitter widget to the Orange Data Mining application worksheet. At Orange, we will use a widget system to do data mining. Each widget has its function and can receive input or output. The Twitter widget allows querying tweets via the Twitter API, which can make queries based on content, author, or both and aggregate the results if you want to build a larger data set.

Preprocessing

The preprocessing stage is the stage for preparing data before being processed to obtain a classification model. The first step of preprocessing text is transformation, namely changing the input data. It applies lowercase transformation by default. After transformation, the second step is tokenization, which breaks text into smaller components (words, sentences, bigrams). After tokenization, the third step is filtering, namely deleting or saving word choices. The last process is normalization, which applies

text stemming and lemmatization. Text separated word for word will become a text that stands alone in a sentence.(Indah Nurhafida & Sembiring, 2021; Wiguna & Rifai, 2021)

Sentiment Analysis

After collecting data from the Twitter API and preprocessing the data, the next step is to perform sentiment analysis(Fang & Zhan, 2015) using the VADER (Valence Aware Dictionary for Sentiment Reasoning) method with Orange Data Mining(Demšar et al., 2013b). To calculate the overall total score of tweets, what is taken is to use the formula:

$$P = \frac{f}{N} \, x \, 100 \tag{1}$$

Information:

P = Percentage

f = Frequency (Sentiment)

N = Total Data

TABLE 1. Sentiment Analysis Results

Sentimen	Count	Percentage	Round up
Posititive	560	27,2772 %	27 %
Negative	183	8,9138 %	9 %
Neutral	1310	63,8091 %	64 %
Total Data	2053		100%

After that, it is calculated using Microsoft Excel. In Table 4.2, the total number of tweets obtained from the analysis results is 2,053 with a positive score percentage of 27%, while the percentage for a negative score is 9%, and a neutral score percentage is 64%. The percentage results for neutral sentiment are greater than positive and negative sentiments. The result of positive sentiment is greater than negative sentiment.

Tweet Profiler

The process of knowing the mood of netizens is in the form of classifying classes of emotions. After conducting sentiment analysis, the sentiment analysis widget is then connected to the tweet profiler widget so that the sentiment analysis results can be used to determine the mood of netizens. The emotions used in the profiler's tweets are Paul Ekman's classification of emotions, with output consisting of six mood classes or moods The six sentiment moods are anger, contempt, disgust, fear, joy, sadness, and surprise (Ekman, 1984). The results of each mood class can be visualized with the box plot widget.

From the results of the visualization formed in Figure 2, it can be explained that the most dominant emotion from netizens about the PLN Mobile Application is joy. There are 1544 or 75% joy emotions contained in 2053 tweets. Another emotion is a surprise, 350 or 17% contained in 2,053 tweets. Next, there is fear, which amounts to 122, with a percentage of 6% contained in 2,053 tweets. Then there is sadness (sadness), which amounts to 1 or 6% contained in 2,053 tweets. Disgust, which amounted to 8 or 0.5%, was contained in 2,053 tweets. Finally, anger emotions were 7 or 0.5% contained in 2,053 tweets.

FIGURE 2. Results of Profiler Tweets with Paul Ekman

Wordcloud

Figure 3 shows the word cloud results with the highest number of words. This image shows that the more frequency of word occurrences, the larger the font size will be in Wordcloud. Wordcloud or word cloud is a variation to display the most word results from the text preprocessing stage(Oesper et al., 2011) . The colors in words in wordcloud give an attractive appearance and are easy to understand. It can be seen in the figure that the word that appears the most is "pln", which appears 1767 times, followed by the word "mobile" 1496 times. Then there is the word "application," which appears 901 times. The word "electricity" appears 763 times, the word "complaint" appears 685 times, the word "report" appears 641 times, and the word "menu" appears 604 times.

FIGURE 3. Wordcloud result in this study

Topic Modelling

Topic modeling is a process for viewing topic groups from the corpus. Topic modeling uses LDA (Latent Dirichlet Allocation), a generative method that represents a list of words from a data set that is similar/related to a particular topic group(Alghamdi & Alfalqi, 2015). A unique feature of the LDA model is that the topics don't have to be distinct, and words can appear in multiple topics. It is good enough to determine what topics are shared on Twitter about particular social events.

6

The ranking of each word in topic modeling with LDAvis can be seen in Figure 4. Sequentially, the first rank is the word complaint, followed by the word menu, get, wait, car, pln, sorry, report, mhn, info, application, address, complete, brother, confirmation, data, follow, follow up, look after, number.
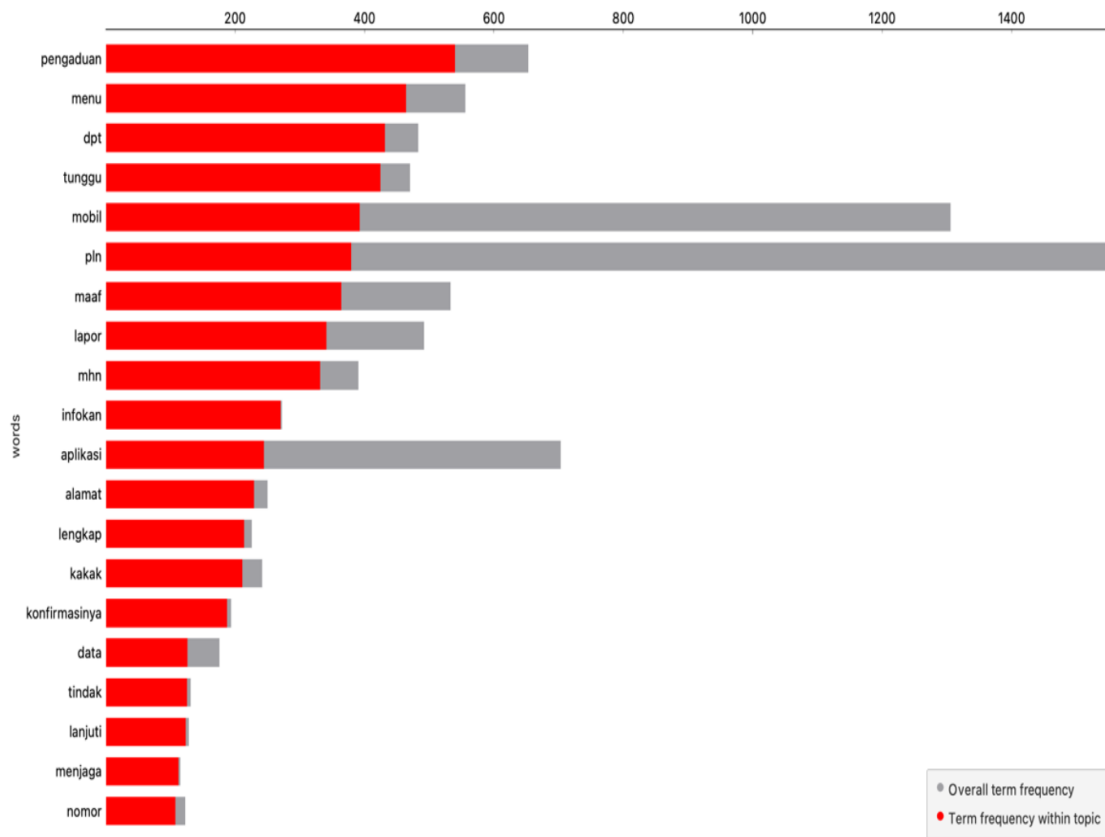

FIGURE 4. LDAvis Ranking Results Topic 1

Heat Map

The next step is to do a sentiment analysis of a collection of Twitter comments. The Lexicon-Vader method gives a score on a very negative to very positive scale with an allowance of 0 as a neutral comment. The compound score results represent the total sentiment score with a range of -0.6 to 1, as shown in the heat map in Figure 4.28. It shows that most customer clusters show very positive sentiments towards the existence of the PLN Mobile application. It will certainly be a concern for PT PLN (Persero) in the future to be able to increase promotions to customers who do not know the features of the PLN Mobile application.

Managerial Implications

Based on the research results, the analysis of opinions and comments on PT PLN (Persero) customers who use the PLN Mobile application used text mining analysis were obtained. The analysis research results show that many users already know about this PLN Mobile application, but the features used are not maximized. The feature that users often use is the complaint feature, even though there are many features in PLN Mobile, such as new installation, adding power, changing power, temporary connection, buying

tokens, electricity bills, recording independent meters, monitoring electricity usage monitor purchase of tokens, blackout information, and electric car charging station installation (SPKLU).

5. CONCLUSIONS

Several points can represent the data based on the sentiment analysis process on the results of crawling data from social media Twitter. By conducting sentiment analysis using VADER, it is known that the public/netizens have a greater positive response than negative responses, with a percentage of positive responses of 27% and positive responses. Negative by 9%. Another response is a neutral response of 64% to the PLN Mobile application.

By using the box plot widget, the mood results of netizens can be seen from the PLN Mobile application. The results of emotions/moods of netizens using tweet profiler with Paul Ekman's most dominant emotional classification are joy (happy) with a percentage of 75%, surprise (surprised) 17%, fear (fear) 6%, sadness (sad) 6%, disgust) 0.5%, anger (angry) 0.5%.

Based on the research results, it is known that many users already know about this PLN Mobile application, but the features used are not maximized. Features that users often use are the complaints feature, even though there are many features in PLN Mobile such as new installation, add power, change power, temporary connection, buy tokens, electricity bills, recording independent meters, monitoring electricity consumption, monitoring purchase of tokens, blackout information, and electric car charging station installation (SPKLU).

REFERENCES

Abimanyu, D., Budianita, E., Cynthia, E. P., Yanto, F., & Yusra, Y. (2022). Analisis Sentimen Akun Twitter Apex Legends Menggunakan VADER. Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI), 5(3), 423–431. https://doi.org/10.32672/jnkti.v5i3.4382

Alghamdi, R., & Alfalqi, K. (2015). A Survey of Topic Modeling in Text Mining. International Journal of Advanced Computer Science and Applications, 6(1). https://doi.org/10.14569/ijacsa.2015.060121

Amalia, Y., Jannah, N., & Prasetyo, R. B. (2022). Analisis Sentimen dan Emosi Publik pada Awal Pandemi COVID-19 Berdasarkan Data Twitter dengan Pendekatan Berbasis Leksikon (Analysis of Public Sentiment and Emotion at the Beginning of COVID-19 Pandemic Based on Twitter Data with Lexicon Based Approach). Seminar Nasional Official Statistics 2022, 597–607.

Amelia, R., Darmansah, D., Prastiwi, N. S., & Purbaya, M. E. (2022). Impementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Masyarakat Indonesia Mengenai Drama Korea Pada Twitter. JURIKOM (Jurnal Riset Komputer), 9(2), 338. https://doi.org/10.30865/jurikom.v9i2.3895

Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., Starič, A., Štajdohar, M., Umek, L., Žagar, L., Žbontar, J., Žitnik, M., & Zupan, B. (2013a). Orange: Data mining toolbox in python. Journal of Machine Learning Research, 14.

Ekman, P. (1984). Expression and the Nature of Emotion. In Approaches to Emotion.

Fang, X., & Zhan, J. (2015). Sentiment Analysis using Product Review Data. Journal of Big Data, 2(1). https://doi.org/10.1186/s40537-015-0015-2

Indah Nurhafida, S., & Sembiring, F. (2021). Analisis Text Clustering Masyarakat Di Twiter Mengenai Mcdonald'Sxbts Menggunakan Orange Data Mining. SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika), 28–35.

Kumar, A. V., & Meera, K. N. (2022). Sentiment Analysis Using K Means Clustering on Microblogging Data Focused on only the Important Sentiments. International Conference on Emerging Trends in Engineering and Technology, ICETET, 2022-April, 1–5. https://doi.org/10.1109/ICETET-SIP-2254415.2022.9791723

Marwa, E. A., & Kristanto, A. B. (2022). Analisis Sentimen Pengungkapan Informasi Manajemen: Text Mining Berbasis Metode VADER. Owner, 6(3), 2853–2864. https://doi.org/10.33395/owner.v6i3.895

Murdifi, A. (2020). Lebih Dekat Dengan Pelanggan, PLN Luncurkan Aplikasi New PLN Mobile. Press Release. https://web.pln.co.id/media/siaran-pers/2020/12/lebih-dekat-dengan-pelanggan-pln-luncurkan-aplikasi-new-pln-mobile

Oesper, L., Merico, D., Isserlin, R., & Bader, G. D. (2011). WordCloud: A Cytoscape plugin to create a visual semantic summary of networks. Source Code for Biology and Medicine, 6. https://doi.org/10.1186/1751-0473-6-7

Rohmah, A., Sembiring, F., & Erfina, A. (2021). Implementasi Algoritma K-Means Clustering Analysis untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: SMK Yaspim Gegerbitung). Seminar Nasional Sistem Informasi Dan Manajemen Informatika, 290–298. https://sismatik.nusaputra.ac.id/index.php/sismatik/article/view/32

Wiguna, R. A. raffaidy, & Rifai, A. I. (2021). Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law Menggunakan Orange Data Mining. Journal of Information Systems and Informatics, 3(1), 1–12. https://doi.org/10.33557/journalisi.v3i1.78