

The Application of Sentiment Analysis on Customers Responses in Social Media (Case Study: Indihome and First Media Internet Service Provider)

Feliks P. Sejahtera Surbakti*, Andi Arifita Ginting

Program Studi Teknik Industri, Fakultas Teknik, Universitas Katolik Indonesia Atma Jaya
Jalan Raya Cisauk-Lapan No. 10, Sampora, Cisauk, Tangerang, Banten 15345

Article Info

Abstract

Article history:

Received
26 November 2022

Accepted
28 Desember 2022

Keywords:

Big Data, Data Science,
Sentiment Analysis,
Naive Bayes Classifier,
Twitter

The development of technology in industry 4.0 era leads to a digitalization process which is marked by the emergence of the internet as one of the important enablers. Indihome and Firstmedia are the most popular Internet Service Providers (ISPs) in Indonesia in providing internet services. These two companies certainly have competition in winning the hearts of customers. One of the important factors in this business competition is customer response in the form of reviews which are usually done through social media tools such as Twitter. For this reason, it is important to study how the customer responds in the form of a given sentiment. Sentiment analysis can be a way to study customer responses to products and services. Sentiment Analysis is a method that is part of Big Data in Data Science which is useful for assessing sentiment in the form of positive or negative sentiments. The data collected was 1000 tweets for each ISP which will be processed with the help of Rstudio. The processing results obtained 177 positive tweets, 489 neutral, and 334 negative tweets for indihome while firstmedia obtained 620 positive tweets, 222 neutral tweets, and 158 negative tweets. The result is firstmedia ISPs get a better positive response and the number of complaints is less than indihome. Internet service providers (ISPs) can improve product and service quality by creating a single database system and also implementing the service design blueprint.

1. INTRODUCTION

The need for technology is an absolute thing that plays a big role in today's life, where technology is the answer to many human problems (Zaid, 2015). This important role is increasingly evident during the Covid-19 pandemic, where many routine activities must be disrupted. With the existence of technology today, it is easier for people to carry out their daily activities (Rohmah, 2021). One of the technologies whose existence plays an important role is the internet. The internet is a very popular term among young people, adults, and the elderly (Dewaweb, 2021). The important role of the internet, especially during the COVID 19 pandemic era, can be seen in the data released by the Indonesian Internet Service Providers Association (APJII) where there is a penetration of internet usage from 64.8% to 73.7% or from 196.71 million people to 266, 91 million people (APJII, 2020). According to the Minister of Communication and Information, there is a shift in internet use during the pandemic, which was previously centered on offices, campuses, schools, and public facilities to

become housing, settlements, and residences (Kominfo, 2020).

The increasing internet users is something that needs to be considered by various internet service providers (ISPs). One of them is the quality of service. Understanding the customer's desire for service quality can be an important thing for a company to win the competition (Lubis & Andayani, 2018). According to Prihantono (2019) the customer is a resource object for the company that must be measured, maintained, and maximized. With the advancement of technology, customers are now easier to evaluate each product and service. One of the things that customers usually do today before buying or using a product or service is to judge based on the opinions of other customers. The use of social media is an important stage in the decision-making process in the use or purchase of a product or service. Information that usually comes to the attention of customers is in the form of brand, price, quality, and so on where the information will be used as consideration for buying or not (Raheni, 2018). For this reason, according to Wirya (2019)

*Corresponding author. Feliks P. Sejahtera Surbakti
Email address: feliks.prasepta@atmajaya.ac.id

companies need to analyze what customers think about the products and services they offer.

In this study, sentiment analysis is used to find out customer opinions on the products or services they use. Sentiment analysis that will be discussed in this research is sentiment analysis of internet service products (ISP) Indihome and First Media on Twitter social media. Twitter is a microblogging platform that allows users to share and express opinions, thoughts, events, and so on (Wirya, 2019). Twitter is also an aspirational network that has massive users and is one of the popular media (BPPD, 2014). According to Hartanto (2017), Twitter is used as a medium to see opinions on a service or product that will be used as a rating material in the process of reviewing and improving products/services. This study discussed regarding customer opinions or responses to the Indihome and First Media, then which ISP is currently superior, and what things must be developed to improve the service quality of the two ISPs.

Our paper is structured as follows. In the next section, we report on the background and current related work regarding sentiment analysis study, especially the implementation of sentiment analysis on internet service providers area. We then thoroughly explain our method to undertake the sentiment analysis on internet service providers and discuss the outcomes of the study on results and discussion section. Finally, we conclude the paper with a summary of conclusion and suggestions for further research.

2. BACKGROUND AND RELATED WORK

The development of information technology certainly cannot be separated from the emerge of internet. The internet itself is a communication network that uses interconnected electronic media (Amaluddin & Ramdani, 2019). To be able to obtain an internet network, of course, people or organizations require the services of internet service providers, commonly referred to as Internet Service Providers (ISPs). In general, ISPs are companies or manufacturers that provide internet access services or online-based communication and information media with a wide network that allows users to connect to the global internet network. Some of the functions or roles in ISPs are as follows (Amaluddin & Ramdani, 2019): as an internet connection service provider, as an internet connection to the nearest internet gateway, as a modem device provider for dial-up internet connection, as a liaison media for internet service users with information services on the World Wide Web (WWW), as a place provider for the homepage and as a medium that helps service users to upload and download data from the internet.

Twitter is a microblogging platform that allows users to share and express opinions, thoughts, events, and so on (Wirya, 2019). Twitter was founded by Jack Dorsey, where he has a desire to create a service that can help many people with various information. In a tweet is limited to 140 characters to communicate various kinds of information ranging from important to information in the form of daily activities. Twitter is an aspirational network that has massive users and is one of the popular media (BPPD, 2014). According to Pak and Paraoubek (in Hartanto, 2017) Twitter is used as a medium to see opinions on a service or item that will be used as a rating material in the process of reviewing and improving products/services.

Sentiment analysis is a process of understanding and also grouping emotions (positive, negative, and neutral) in an article using text analysis techniques (Arviana, 2021). In the process of sentiment analysis, there are two approaches, namely Machine Learning Approach and Knowledge Based Approach. For the Machine Learning approach, there are Naive Bayes, Maximum entropy, SVM, and Neural Network methods. As for the Knowledge Based approach, the methods that can be used are Lexicon Based and Pointwise Mutual Information (PMI) (Isa, 2017). Sentiment analysis is a method of understanding, extracting sentiment data which will be grouped by polarity (positive, negative, and neutral) (Tineges et al., 2020).

Meanwhile, according to Liu (in Rofiqoh et al., 2017) sentiment analysis or opinion mining is a computational study of people's opinions, sentiments, and emotions through entities and attributes that are owned and expressed in text form. According to Paquette (in Wirya, 2019) Consumer Sentiment Marketing is a concept that refers to consumer feelings towards the market where this is an important factor that needs to be taken into account, especially in the current era of social media. According to Mullen (in Wirya, 2019) the ultimate goal of a sentiment analysis process is product or service development, evaluation of public opinion, finding new ideas, measuring reactions, and others.

In previous study, Wirya (2019) used sentiment analysis to research the preference of customer on online transportation service between Gojek and Grab application. This study uses a Naive Bayes approach and uses R Studio as a research tool and Fishbone Diagram to conduct the problem evaluation process. The results from this study are that Grab has a better positive response than Gojek, and there are several suggestions that can be improved to enhance services of online transportation in the future. Furthermore, the second study by Lestari, et al (2020) with the title "Naive

Bayes Classifier Method with Textblob for Sentiment Analysis for Indihome and First Media Services". This study uses a Naive Bayes approach and uses Python as a research tool. The result from this study was both vendors have more negative responses than positive responses.

Moreover, Tineges, et al (2020) with the title "Sentiment Analysis against Indihome Services Based on Twitter with the Support Vector Machine (SVM) Classification Method" conduct the sentiment analysis using different tools and techniques from previous study. This research focuses on company services in this case Indihome with SVM approach. The results of this study conclude that the negative response from service users is greater than the positive response. In addition, Utami (2020) conducted study on comparative analysis of the quality of service of wireless-based internet networks on Indihome and First Media Internet Service Providers (ISPs). This study discussed the comparison of ISPs with the QOS method approach. The results obtained are that both ISPs get a bad category with several recommendation were suggested to improve the services.

All of these studies provide several insights on how to conduct study to analyze customers complaints or opinions, especially on internet service providers companies. In this study, the authors will conduct the analysis of sentiment of people on Twitter's social media, case study: Internet Service Providers (ISPs) Indihome and First Media. This study will focus on the sentiments generated from Twitter social media conversations against Indihome and First Media. This study uses a Naive Bayes approach and uses R Studio as a research tool. Furthermore, the authors use fishbone diagram (Ishikawa diagram) to find the root causes of the problem as the basis for future service designs.

3. RESEARCH METHOD

In this study, the research design used a social survey study and a comparative design. Where researchers will conduct a survey based on Indihome and First Media data from social media twitter. Then a comparison will be made on the opinions of various customers towards the two ISPs. This research uses a combination of quantitative and qualitative approaches. Primary data will be collected by manual process using the Twitter Developer API where the data will be divided into two, namely training data and testing data. Secondary data will be obtained after processing the data in this study.

The approach used in this study is the Naive Bayes approach where the Naive Bayes approach is a probability or opportunity-based classification

method where this method calculates a set of opportunities by adding up the frequency and combination of values from the given dataset (Pramana et al., 2018). According to Boehmke (in Wirya, 2019) this approach is a simple probabilistic algorithm that uses Bayes' Theorem to predict text which is characterized by strong assumptions about independence. According to Brownie (in Wirya, 2019) this approach has advantages in terms of handling missing data, building models, and can also make predictions based on probability. In addition, the Naive Bayes approach has advantages, among others, it can be used for quantitative and qualitative data, does not require a lot of data, fast calculations, easy to understand, easy in the coding process and so on (Widianto, 2019).

In this study, the data collection process was carried out with a Quota Sampling collection technique where data had to be collected until the amount of data was determined by taking into account the characteristics of Indihome and First Media. The data that will be collected is up to 1000 tweets for each ISP which this figure is based on suggestions from several previous studies (Wirya, 2019), (Setyobudi et al., 2018), and (Bharata & Sulistyowati, 2018). The collection is done using the R Studio command by studying several packages that will be used in the process of collecting data from Twitter and installing them into the application first.

The data will be divided into two categories: training data and testing data. Training data will be collected with the Twitter Developer API. The Twitter Developer API will provide users with access to information from Twitter in the form of keys and tokens that will be used in R Studio Command. Training data is used to train algorithms in machine learning such as sentiment analysis. Testing Data will also be collected with the Twitter Developer API based on customer opinions of ISP Indihome and First Media. The API will provide access to collect tweets that will become data in real time. Data testing is carried out to assess the performance of the previously trained algorithm. The difference between training data and testing data is that data testing only includes input data without the expected value, while training data includes both. Following in Table 1 is the example of datasets.

Table 1.
Datasets example

| | | |
|------------------|---------|--|
| Training Data | Tweet 1 | @IndiHome knpa jaringan internet IndiHome hilang2 |
| | Tweet 2 | @IndiHome min knpa internet d rumah saya gbisa pake, masa bayar aja tp gbisa d pake hehe |
| | Tweet 3 | min knp ini gabisa wifinya ya tuluunngg @IndiHome etc |
| Testing Data | Tweet 1 | @firstmedia tolong lah, tugas lg numpuk internet jgn mati mulu dong?! |
| | Tweet 2 | @FirstMedia min belum bisa nih sampai sekarang, saya butuh banget loh |
| | Tweet 3 | @firstmedia Sangat mengecewakan jaringan sering gangguan iuran dinaikan terus tanpa pemberitahuan sangat tidak profesional. etc |

After obtaining the sentiment assessment data from a series of tweets, then predictions are made based on the current assessment results to estimate sentiment in the future. The classification uses Naive Bayes. At this stage, the data that has been assessed for sentiment will be separated into two, where 80% of tweets will be used as training data and the remaining 20% will be used as testing data to predict sentiment in the future. Then after the data is processed, the researcher will study and analyze to find a solution to the situation. Then the problem will be visualized in a service design tool called Fishbone Diagram. In addition, researchers will also make observations based on datasets that have been collected previously. Which will be analyzed next.

4. DISCUSSIONS AND RESULTS

After processing the data starting from the raw data in the form of tweet data for each ISP until it is processed in such a way as to be clean data which is finally visualized, a picture of the problem for each ISP can be seen. Broadly speaking, none of these ISPs has perfect service because they both have problems that need to be improved. To see a comparison of the two at this time, each of the problems will be described in the discussion of each of the existing factors.

Starting from the price factor where each ISP has problems in terms of excessive bills or unilateral tariff increases which are detrimental to the customers of these two ISPs. However, in terms of shortages, indihome customers are seen

complaining about large costs starting from the installation process to the monthly so, this is a concern from the side of the indihome ISP to fix it in the future.

Then from the people factor, both of them certainly have similar problems. Each ISP has problems with long repair times, but on firstmedia it tends to be longer with repair times that sometimes touch the number of 3-5 days of repair compared to the 50 hours needed by Indihome for repairs. However, the indihome ISP also has other problems, such as the slow response of the admin when problems occur and the slowness of field technicians visiting the location of customer problems.

Furthermore, from the process factor, both ISPs tend to have similarities in terms of slow response to problems for various reasons given. Meanwhile, in terms of the installation process, the problem is that in indihome the cable installation is chaotic, while in firstmedia the cable installation has to drill through the road so that it often interferes with the activities of the surrounding community.

Then from the place factor which is also not too different where there are many disturbances in various places so that it requires more service posts to respond to any existing problems. Then in terms of the availability of network services, especially indihome ISPs, which are sometimes not evenly distributed or require a certain number of customers to develop a network in certain areas, this is a problem for prospective customers. Then from the product factor, of course, a classic problem from every ISP which also occurs in indihome and firstmedia where the network is often unstable or even completely dead. In addition, there are also problems of difficulty accessing certain sites and difficulties opening other services such as movie watching services.

In general, these two ISPs have almost the same problems, so both of them certainly have almost the same appeal, but if they have to choose, prospective consumers need to consider several things which are facts about current services, such as complaints about the slow (slow) internet network where this is a matter that has a direct impact on customer satisfaction using ISP services. As for the current condition, based on the results of the study, it shows that indihome ISPs have more complaints about network slowness than firstmedia ISPs. In addition, customers also need to consider the price of the services offered where indihome tends to be more expensive, but the next thing that also needs to be considered is the service provider company where indihome is under the auspices of PT. Telekomunikasi Indonesia (Telkom) where this company certainly gets the attention of the state because the Telkom company is part of a State-

Owned Enterprise (BUMN) while firstmedia is under the auspices of the Lippo Group which is a private sector. So if you look at the resources owned, of course Indihome should also be considered an option for users considering its status as part of a BUMN. To make it easier to see the problems of the two ISPs, the following is a comparison between the two ISPs in tabular form to see the differences:

Table 2.
Comparison of Customer Complaints of Two ISPs

| Factor | Indihome | Firstmedia | |
|---------|---|--------------------------------------|---------------------------|
| Price | The bill does not match the application | The bill suddenly went up | |
| | Expensive installation fee | | |
| | Rate increase without confirmation | | |
| People | Long repair process | Long repair process | |
| | Technician came a long time | | |
| | Admin response is slow | | |
| Process | Old problem response | Unclear responsibility | |
| | Chaotic cable installation | Cabling is interrupting | |
| Place | There are disturbances in some areas | There are disturbances in some areas | |
| | Uneven network | | |
| Product | Difficult to access certain sites | Difficult to access certain sites | |
| | Unstable network | Unstable network | |
| | Troubled service pack | Less variety of services | Fiber optic cable problem |
| | | | |

In response to current global developments, an internet service is needed that can not only provide services but also be able to satisfy every customer with the best service they have to win the competition (Lubis & Andayani, 2018). So based on the table above, it is necessary to improve the quality of the two ISP services. Starting from indihome based on the ISP price factor, this requires aligning services with applications, reviewing financing policies, and increasing information to customers. Furthermore, from the people factor, it is necessary to improve SOPs in handling problems, evaluating the availability of technicians, and improving the quality of customer service. From the process factor, it is also necessary to improve SOPs and supervision in the network installation process.

The place factor evaluates regional service posts and improves services out of urban areas. Finally, the indihome product factor improves the evaluation of the site, improves controlling of network quality and improves customer services so that customers get the information they need when there is a service disruption.

On the other hand, firstmedia based on the price factor, the ISP needs to evaluate the customer payment administration process. From the people factor, improving the quality of equipment and increasing the repair time to be faster. Furthermore, from the process factor, it is necessary to provide clear SOPs for personnel for any problems that may occur and increase supervision in the product installation process. From the place factor, it is necessary to prepare more service posts so that the service process is more optimal. Finally, from the product factor, it is necessary to evaluate every customer problem on a regular basis, improve the quality-of-service support equipment, improve controlling processes, and provide opportunities for new service programs for customers such as service bundling packages.

The results of the comparison of ISPs in this research period resulted in recommendations for the use of ISP services, namely firstmedia by obtaining a better positive response which was shown in the results of data processing and fewer complaints compared to indihome with the following details: indihome received 177 positive tweets, 489 neutral tweets, and 334 negative tweets. Meanwhile, firstmedia received 620 positive tweets, 222 neutral tweets, and 158 negative tweets. The negative results received by indihome also occurred in several other studies that discussed sentiment analysis on indihome ISPs such as those conducted by (Ulkhag & Br. Barus, 2017) and (Rizkia et al., 2019). So, based on the results of this study, firstmedia ISPs can become the main choice for customers in determining internet service providers.

5. CONCLUSIONS

Based on the research process, it was concluded that from 1000 tweet data from each ISP, indihome obtained 177 positive tweets, 489 neutral tweets, and 334 negative tweets. Meanwhile, firstmedia received 620 positive tweets, 222 neutral tweets, and 158 negative tweets. This shows that firstmedia ISPs get a better positive response which is shown in the results of data processing and the number of complaints is less than indihome. The two ISPs need to improve the quality of each starting from indihome which needs to evaluate several complaints such as high cost, slow response to customer complaints, network installation process, quality of service (internet speed), to network availability. Meanwhile, Firstmedia also

needs to evaluate several complaints such as control of customer billing, the process of resolving network disturbances, improving product and service quality.

Based on the conclusions of the research results, suggestions can be given in the twofold. Firstly, internet service providers (ISPs) should improve the quality of products and services because customers demand for excellent service significantly increasing in modern world today. The course of action that can be taken by ISP such as creating a single database system and implementing the service design blueprint. Secondly, further researchers are expected to deepen the process of sentiment analysis and conduct advanced research, using classifiers, social media platforms, research subjects, and other processing tools.

REFERENCES

1. APJII. (2020). Buletin APJII. (<https://apjii.or.id/downfile/file/BULETINAPJIIEDISI74November2020.pdf>). Diakses tanggal 7 Juli 2022.
2. Bharata, H. K., & Sulistyowati, H. (2018). Analisa Sentimen tentang Reuni 212 pada Dokumen Twitter Berbahasa Indonesia menggunakan R Studio. *Jurnal Gerbang*, 8(2), 1–8.
3. BPPD. (2014). Jaring Aspirasi Lewat Twitter. (<https://bapenda.jatengprov.go.id/liputan/jaring-aspirasi-lewat-twitter/>). Diakses tanggal 9 Juli 2022.
4. Dewaweb. (2021). Pengertian Internet, Sejarah dan Perkembangannya. (<https://www.dewaweb.com/blog/pengertian-internet/>) Diakses tanggal 11 Juli 2022.
5. Hartanto. (2017). Text Mining Dan Sentimen Analisis Twitter Pada Gerakan Lgbt. *Intuisi : Jurnal Psikologi Ilmiah*, 9(1), 18–25.
6. Kominfo. (2020). *Kementerian Komunikasi dan Informatika*. (https://kominfo.go.id/content/detail/26060/terjadi-pergeseran-penggunaan-internet-selama-masa-pandemi/0/berita_satker) Diakses tanggal 11 Juli 2022.
7. Lubis, A. S., & Andayani, N. R. (2018). Pengaruh Kualitas Pelayanan (Service Quality) Terhadap Kepuasan Pelanggan Pt. Sucofindo Batam. *Journal of Applied Business Administration*, 1(2), 232–243. (<https://doi.org/10.30871/jaba.v1i2.619>).
8. Pramana, S., Yuniarto, B., Mariyah, S., Santoso, I., & Nooraeni, R. (2018). *Data Mining Dengan R Konsep Serta Implikasi*. In Media.
9. Prihantono, D. N. (2019). Komunikasi Bisnis Berorientasi Pelanggan Di Perusahaan Distribusi. *Jurnal Audience*, 1(2), 123–143. (<https://doi.org/10.33633/ja.v1i2.2687>).
10. Raheni, C. (2018). The Effect Of Social Media On Consumer Buying Intrerest Study Case Study. *Jurnal Sinar Manajemen*, 5(2), 82–85. (<http://jurnal.unismuhpalu.ac.id/index.php/JS M>).
11. Rizkia, S., Budi Setiawan, E., & Puspadari, D. (2019). Analisis Sentimen Kepuasan Pelanggan Terhadap Internet Provider Indihome di Twitter Menggunakan Metode Decision Tree dan Pembobotan TF-IDF. *E-Proceeding of Engineering*, 6(2), 9683–9693.
12. Rohmah, I. (2021). *Peran Teknologi di Masa Pandemi Covid-19*. (<https://radarbojonegoro.jawapos.com/read/2021/01/09/234513/peran-teknologi-di-masa-pandemi-covid-19>) Diakses tanggal 5 Agustus 2022.
13. Setyobudi, W., Alwi, A., & Astuti, I. P. (2018). Sentimen Analisis Twitter Terhadap Penyelenggaraan Gojek Traveloka Liga 1 Indonesia. *Komputek*, 2(1), 56. (<https://doi.org/10.24269/jkt.v2i1.68>).
14. Ulkhaq, M. M., & Br. Barus, M. P. (2017). Analisis Kepuasan Pelanggan dengan Menggunakan SERVQUAL: Studi Kasus Layanan IndiHome PT. Telekomunikasi Indonesia, Tbk, Regional 1 Sumatera. *Jurnal Sistem Dan Manajemen Industri*, 1(2), 61. (<https://doi.org/10.30656/jsmi.v1i2.365>).
15. Widiyanto, M. H. (2019). *Algoritma Naive Bayes*. (<https://binus.ac.id/bandung/2019/12/algoritma-naive-bayes>) Diakses tanggal 5 Agustus 2022.
16. Wirya, V. L. (2019). *Sentiment Analysis On Twitter Data For Indonesia Online Transportation Service (Case Study Of Gojek And Grab Application)*.
17. Zaid, M. (2015). *Pentingnya Teknologi dalam Kehidupan*. (<https://www.kompasiana.com/maszaid/5520c85fa333115e4946ce15/pentingnya-teknologi-dalam-kehidupan>). Diakses tanggal 5 Agustus 2022.