Analysis of Twitter Social Media on the Supporting Criteria of Largest Marketplace Customer Satisfaction in Indonesia

Muhammad Iqbal
Department of Technology Management
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
Iqbalelfaita@gmail.com

Mohammad Isa Irawan
Department of Mathematics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
mii@its.ac.id

Rita Ambarwati Sukmono
Department of Management
Universitas Muhammadiyah Sidoarjo
Sidoarjo, Indonesia
ritaambarwati@umsida.ac.id

Abstract

Marketplace use during the COVID-19 pandemic in Indonesia increased by 400%. To reach customers, marketplace use social media to connect directly with customers. Criticisms and messages from customers conveyed through social media are data that can be analyzed and generate benefits for the advancement of the marketplace. The purpose of this study is to obtain an overview of the trends in customer satisfaction supporting criteria in customer social networks, a comparison of trends in these criteria before and during COVID-19 and an overview of user social networks of the marketplace using Twitter data. To achieve the objectives of this study, the researcher uses the Semantic Association Analysis method as a descriptive analysis method and Ego Network Analysis to analyze social networks, both of which are part of Social Network Analysis (SNA). From the results of this study, it was found that the trend of supporting criteria for customer satisfaction in of Shopee was led by the criteria for Delivery Quality and Price. This study recommends some suggestion on how to take advantage of trends in customer satisfaction criteria and customers' social networks on Twitter social media as a focus in developing strategies for the marketplace.

Keywords
Marketplace, Social Media, Social Network Analysis, Semantic Association Analysis, Ego Network Analysis

1. Introduction

The condition of the COVID-19 pandemic has made people all over the world keep their social distance and do a lot of online activities, as well as shopping activities. Based on a statement from the Ministry of Communication and Information (Kominfo) that during the pandemic, the increase in online shopping activity during the pandemic reached 400 percent. Based on the Indonesian marketplace traffic share data from Similarweb, in the first quarter (Q1) of 2021, Shopee being one of Indonesian top marketplaces with 29.73% traffic (117 million visits per month). That amount of traffic share indicating that many people interest with Shopee to do online shopping. To reach many of the customers, Shopee used social media to make interractions with them. That's because social media is one of the methods to be able to listen to comments from various communities or dig up data and the information contained therein (Dietrich and Livingston 2012).

By looking at the many comments related to the largest marketplace Indonesia on social media, such as Twitter, does not guarantee that customers are satisfied with transactions made using the services of the marketplace. For this reason, it is necessary to carry out an analysis that can identify trends criteria that support the customer satisfaction of the marketplace through Twitter social media data. One of the approaches used for analyzing data on social networks is Social Network Analysis. Social Network Analysis is an analytical method that can describe social media data as a network which then performs calculations based on the network formed by bigram data to get information from the network.

In this study, researchers analyze the structure of the social network and textual network of Twitter users in Indonesia regarding Shopee. The author conducts an analysis using the Social Network Analysis (SNA) approach to describe and map customers as a node (point) in the network that is interconnected with other nodes using edges (lines) (Zhao 2009). The use of SNA in this study aims to identify how the structure of the social and textual network is formed by identifying trends in customer satisfaction supporting criteria that are widely discussed and how customers communicate with the marketplace and their influence on the network.

One application of SNA in extracting information in a network is networked data mining (Tabassum 2020). One of the methods used to perform networked data mining in this research is Semantic Association Analysis. The use of Semantic Association Analysis in this study is to be able to analyze trends in customer satisfaction supporting criteria from opinions uploaded by customers on Twitter in the form of tweets related to Shopee marketplaces. Based on research conducted by Lin (2011) and Irawan (2008), the supporting criteria for customer satisfaction used in this study are Product Quality, Price, Service Quality, Delivery Quality, Information Quality, System Quality and Delivery Quality.

In addition to conducting descriptive analysis using Semantic Association Analysis, on text data obtained from Twitter, as was done in the research conducted by Hou (2019), this study also combines research methods by conducting network analysis using Ego Network Analysis which examines network patterns based on values. centrality and betweenness (Watanabe 2021). The analysis was conducted to identify the pattern of interaction between customers and the marketplace and their influence on social networks in Shopee marketplaces. The combination of the analysis carried out by the researcher is expected to become a comprehensive analysis result that can be a reference for developing a marketplace business strategy based on social media data.

1.1 Objectives

The aim of this research is to analyze the trend of supporting criteria in markeplace by the data of social media Twitter. In order to achieve the aim of this research, following objectives were defined

1. Generate a trend picture of customer satisfaction criteria against Indonesia's largest marketplace on social media twitter before and during COVID-19
2. Identify the social network structure between customers and marketplace in social media Twitter
3. Provide recommendations on the use of trends in satisfaction supporting criteria customers and customer social networks on social media Twitter in marketplace business development

2. Literature Review

This research was conducted by taking references from previous studies that have been carried out previously, especially those discussing about Social Network Analysis, Semantic Analysis Network, Ego Network Analysis and analysis on social media. This study uses two models Social Network Analysis approach, namely Semantic Association Analysis and Ego Network Analysis. This study uses in the research by Hou (2019) as a reference for the use of Semantic Association Analysis in analyzing opinion data obtained from social media or online opinion portals.

The use of the Ego Network in this study takes reference to research by Watanabe (2021) and Arnaboldi (2012), the creation of the ego network done after entering all nodes and edges into the Social Network Analysis tools are then filtered using the name of the account you want to use become the center of the network. As for the reference regarding Social The Network Analysis tools used refer to the results of research by Khausik (2018). Gephi tools were chosen because the functions they have are relatively more complete compared to some other social network tools. Gephi has been widely applied to analyze in the field of social network analysis, biology and genomics (Bastian 2009).

3. Methods

To do the research, researcher doing some steps from the collection data, analyzing data untill produces the recommendation based on the result of this research.

3.1 Data Collection Method

To be able to do social network analysis related to Indonesian markertplace in the Twitter network, data collection is needed that obtained from the Twitter network. The marketplace that used in this research is Shopee. Shopee is one of the marketplace leader in Indonesia with the highest traffic share in Q1 2021, with 29.73% traffic share (117 million visits per month). The data used in this study were extracted using Python programming language.

Data collection was carried out in May 2021. The data that been collected are the tweets addressed to official account Shopee (@ShopeeID). The collected data are tweets in Indonesian. In this study, the data focused on the interactions that occurred in Twitter around @ShopeeID account before and during the COVID-19 pandemic. Therefore, the data collected is divided into 2 groups, they are the data before COVID-19 and the data during COVID-19. Data before COVID-19 is the tweets uploaded within 1 year before the COVID-19 pandemic starting from April 1, 2019 until March 31, 2020. While the data during COVID-19 is the tweets that uploaded starting from April 1, 2020 to March 31, 2021.

3.2 Data Analysis Method

Semantic association is an important element of language humans in communicating containing valuable information (Kim 2011). In this study, the Semantic Association Analysis was carried out by using bigram co-occurrence mapping. bigram co-occurrence can helps prevent loss of information and distortion in computing and aggregation of information on linguistic evaluation and making calculation results more accurate (Harrera 2000). Before mapping the textual data into bigram co-occurrence, researcher clean the data through 2 steps, that are Tokenization and Stopword removal. For example, bigram co-occurrence mapping from one of the tweets related to the marketplace as follows:

“Makasih @shopeeID nggak sia2 kan ngumpulin cashback sampe minggu lalu kebeli tas marhenJ pake ovo points”

There are 2 steps for get the data to be ready before data anakyzing :
1. Tokenization : Dividing a sentence to be words or tokens.
Example :
Makasih/shopeeID/nggak/sia2/kan/ngumpulin/cashback/sampe/minggu/lalu/kebeli/tas/marhenJ/pake/shopee/pay
2. Stop word removal : Removing the conjunction words and words that have no singular meaning
Example :
Makasih/shopeeID/nggak/sia2/ngumpulin/cashback/minggu/kebeli/tas/marhenJ/shopee/pay

Table 1. Example for Bigram Co-occurrence phrase mapping

Bigram phrase	Frequency
Makasih-shopeeID	1
shopeeID-nggak	1
Nggak-sia2	1
Sia2-ngumpulin	1
Ngumpulin-cashback	1
Cashback-minggu	1
Minggu-kebeli	1
Kebeli-tas	1
Tas-marhenj	1
Marhenj-shopee	1
shopee-pay	1

After being mapped using bigram co-occurrence, like seen in Table 1, then the researcher accumulates all obtained bigram co-occurrence phrases and make a list of bigram co-occurrence phrases of the marketplace. To mapping bigram data, this research uses bigram tools in the NLTK library with a programming language Pythons. The number of tweet data mapped using bigram co-occurrence in this study makes it possible for bigrams to appear which is reversed as is the case of local-stuff and local-stuff. Due to data generated in the bigram co-occurrence mapping is non-volatile data direction, then in this study treat bigram phrase data with position reversed as phrases with the same meaning (Hou 2019).

Ego Network analysis is done by calculating the value of centrality and betweenness in the network (Watanabe 2021). To perform Ego Network analysis, first determine the focus of a network to be analyzed. In this study, the focus of the Ego Network for each marketplace is the social media accounts of @ShopeeID (Shopee). After determining the focus on each Ego Network, then calculating the Betweenness and Eigenvector Centrality on each of these networks. The visualization that is carried out on the bigram co-occurrence results uses the ForceAtlas 2 layout algorithm. ForceAtlas 2 has the advantage of better measurement quality compared to other algorithms (Jacomy 2014).

4. Result and Discussion

Table 2 shows some differences in the trend of customer comments based on bigram co-occurrence on the Shopee marketplace. In the time period Before COVID-19, some things were discussed and didn't appear on the top 10 bigram phrases during COVID-19 related to respon-dm, beli-barang, nomor-pesanan, ajukan-pengembalian, and sesuai-syarat. while during the time during the COVID-19 pandemic, many phrases discussed and did not appear in the top 10 before COVID-19 were jasa-kirim, vourhcer-gratis, saldo-shopee, and bayar-barang. As for the bigram frasa phrase such as shopee-pay, gratis-ongkir, aplikasi-shopee, and metode-bayar have increase in the number of tweets during the COVID-19 pandemic.

Table 2. Top 10 bigram phrase for Shopee

Before COVID-19		**During COVID-19**	
Bi-gram phrase	**Frequency**	**Bi-gram phrase**	**Frequency**
Shopee-pay	1381	Shopee-pay	3046
Respon-dm	498	Syarat-ketentuan	1285
Gratis-ongkir	387	Gratis-ongkir	1018
Beli-barang	386	Jasa-kirim	908
Nomor-pesanan	342	Voucher-gratis	637
Pengembalian-dana	324	Saldo-shopee	487
Ajukan-pengembalian	302	Aplikasi-shopee	372
Sesuai-syarat	254	Metode-bayar	360
Aplikasi-shopee	244	Opsi-pengiriman	333
Metode-bayar	229	Bayar-barang	330

To be able to read trends in customer satisfaction supporting criteria the largest marketplace in Indonesia on Twitter social media, carried out grouping on the results of the bigram co-occurrence phrase mapping into some criteria. Grouping is done using the top 500 data bigram co-occurrence which is also used in network visualization analysis. Bigram co-occurrence data are then grouped into 6 criteria, namely Information Quality (Information), Service Quality (Service), System Quality (System), Delivery Quality (Delivery), Product Quality (Product) and Price.

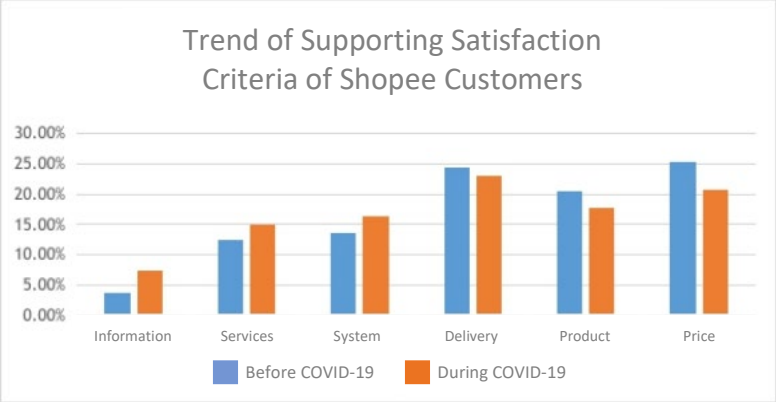

Figure 1. Trend of Supporting Criteria of Shopee Customers Satisfaction

Table 3. Percentage of Supporting Criteria of Shopee Customers Satisfaction

Criteria	Percentage Before COVID-19	Percentage During COVID-19
Information	3.80 %	7.47 %
Service	12.34 %	15.01 %
System	13.49 %	16.24 %
Delivery	24.48 %	**22.92 %**
Product	20.54 %	17.61 %
Price	**25.36 %**	20.76 %

In Figure 1, the trend of criteria for supporting customer satisfaction in the Shopee marketplace is led by Shipping and Price which has a small percentage difference. As for the comparison of trends in customer satisfaction supporting criteria in the Shopee marketplace before and during COVID-19, as shown in Table 3, there was a change in topic trends between data before and during COVID-19. From the results of the grouping of Shopee's co-occurrence bigram data before COVID-19, the trend of criteria at Shopee focuses on the Price criteria with a percentage of 25.36% which is slightly higher than the Delivery criteria which has a percentage of 24.48%. Meanwhile, during COVID-19, despite experiencing a decrease in the percentage of both criteria, the trend of criteria at Shopee was more focused on Delivery with a percentage of 22.92%, while the Price criteria was below it with a percentage of 20.69%.

5. Conclusion and Future Research Directions

There are many research have been conducted to analyze customer satisfaction of marketplace, but still few focus on analysis using customer opinions or responses on social media. This research analyzed tweets data that is part of interaction between Shopee and its customers. This Study categorized the supporting criteria of customers satisfaction, based on Lin (2011), into 6 group that are Information Quality (Information), Service Quality (Service), System Quality (System), Delivery Quality (Delivery), Product Quality (Product) and Price. As the result of this research, Delivery Quality and Price is the most criteria that support customer satisfaction of marketplace in Indonesia. It is almost similar to the research by Lin (2011), that product and delivery is one of the most important in customer satisfaction.

5.1 Practical Implications

This study recommends the marketplace to give more focus on developing strategies according to the results of the trend of customer satisfaction criteria in this study. In Shopee it is recommended to improve strategies related to Delivery Criteria, including Shopee's new feature, like the COD (Cash On Delivery) feature.

5.2 Theoretical Implications

This study points that social network analysis of social media data especially those related to the user or customers comments, can bring insight for the company to choose the direction of development to gain the value of the company.

5.3 Limitation of the Study

Like most researches, this study also have its limitation. One of the most limitation of this research is about the social media that used to get customer's comment is only from Twitter. It hopefully can be more accurate if the data is combined from some social media platform like Facebook and Instagram.

5.4 Future Research

This research can be developed and combined with other methods to get enrich the knowledge around the social media analysis. Because the focus of this research is only around the text mining from comments got from Twitter, It can be expanded also towards analyzing for other social media content or cross social media comment.

References

Arnaboldi, V., Conti, M., Passarella, A. and Pezzoni, F., Analysis of Ego Network Structure in Online Social Networks, *2012 International Conference on Privacy, Security and Trust and 2012 International Conference on Social Computing*, pp. 31-40, 2012.

Bastian, M., Heymann, S. and Jacomy, M., Gephi: An open source software for exploring and manipulating networks, *Proceedings of the third international ICWSM conference*, pp. 361–362, 2009.

Dietrich, G. and Livingston, G., *Marketing In The Round : How to Develop an Integrated Marketing Campaign in the Digital Era*, Que Publishing, Indiana, 2012.

Herrera, F., and Martinez, L., A 2-tuple fuzzy linguistic representation model for computing with words, *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 6, pp. 746–752, 2000.

Hou, Z., Cui, F., Meng, Y., Lian, T. and Yu, C., Opinion mining from online travel reviews: A comparative analysis of Chinese major OTAs using semantic association analysis, *Tourism Management*, vol. 74, pp. 276-289, 2019.

Jacomy, M., Venturini, T., Heymann, S. and Bastian, M., ForceAtlas2 : A continuous graph layout algorithm for handy network visualization designed for the gephi software, *PLoS One*, vol. 9, no. 6, pp. 1–12, 2014.

Kaushik, T., Singhal, S., Mandan,J. and Sharma, K., Social Networking Analysis: A case study in Tools, *International Journal of Engineering and Advanced Technology*, vol. 8, issue 2, 2018.

Kim, K. K., Karunanayaka, P., Privitera, M. D., Holland, S. K. and Szaflarski, J. P., Semantic association investigated with functional MRI and independent component analysis, *Epilepsy and Behavior*, vol. 20, no. 4, pp. 613–622, 2011.

Lin, C., Wu, H. and Chang, Y., The critical factors impact on online customer satisfaction, *Procedia Computer Science*, pp. 276-281, 2011.

Tabassum, S., Pereira, F.S.F., Fernandes, S. and Gama, J., Social Network Analysis: An Overview, *Data Mining and Knowledge Discovery*, vol. 8, issue 5, 2018.

Watanabe, N. M., Kim, J. and Park, J., Social network analysis and domestic and international retailers: An investigation of social media networks of cosmetic brands, *Journal of Retailing and Consumer Services*, vol. 58, 2021.

Yulisetiarini, D., and Prahasta., Y. A., The Effect Of Price, Service Quality, Customer Value, And Brand Image On Customers Satisfaction Of Telkomsel Cellular Operators In East Java Indonesia, International Journal of Science & Technology Research, vol. 8, issue. 3, 2019.

Zhao, X., Gao, X., Guo, J. A., and Tao, N., Research focus analysis based on the frequency of topic words and g-index, *Library & Information Service*, vol. 53, no. 2, pp. 59–61, 2009.

Biographies

Muhammad Iqbal is a Student of Magister for Management of Information Technology, Institut Teknologi Sepuluh Nopember. He received Bachelor of Computer in Informatics Engineering at Brawijaya University in 2017. His main aim of this research was to analyze social media data for generating information that can be some insight through user's comment in social media.

Mohammad Isa Irawan is a Professor of machine learning and he served as a head of Machine Learning and Big Data Laboratory in Mathematics Department, Institut Teknologi Sepuluh Nopember. He holds his Bachelor of Mathematics and Master Degree of Electrical Engineering from Institut Teknologi Bandung, Indonesia. And he holds

his Doctoral of Computer Science from Vienna University of Technology, Austria. His interest of research machine learning, decision science, data mining and bioinformatic.

Rita Ambarwati is the associate professor of Faculty Business Law and Social Science in Universitas Muhammadiyah Sidoarjo. She holds a Master Management degree in Industrial Management from Institut Teknologi Sepuluh Nopember Surabaya, Indonesia and a Doctor of Management Science degree in Management Strategy from Universitas Brawijaya, Malang, Indonesia. She has been recognized as a professional management consultant with over 10 years of experience in working with closely-held businesses. He has taught courses in entrepreneurship, management and corporate entrepreneurship. She has received several research grants from the Ministry of Research, Technology and Higher Education of the Republic of Indonesia.